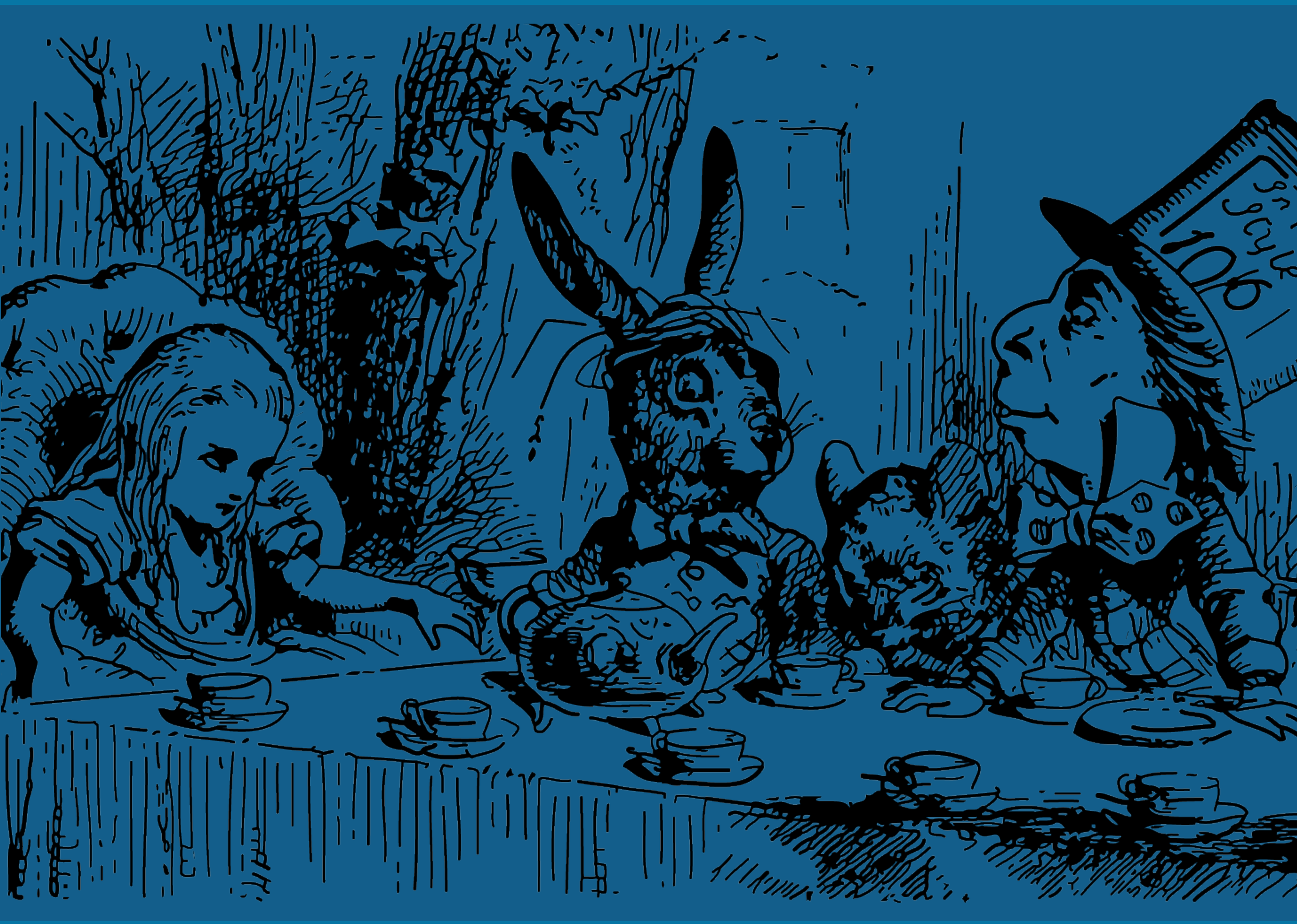
Universitat Oberta de Catalunya



Deeper Down the Rabbit-Hole

Unfolding the Dynamics of Imagination Acts

Joan Casas Roma

PhD Thesis

Universitat Oberta de Catalunya

Department of Computer Science, Multimedia and Telecommunications

---

Deeper Down the Rabbit-Hole: Unfolding the Dynamics of Imagination Acts

---

A DISSERTATION SUBMITTED TO UNIVERSITAT OBERTA DE CATALUNYA IN FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR IN INFORMATION AND KNOWLEDGE SOCIETY

Author:

Joan Casas Roma

Advisers:

Dr. M. Antonia Huertas Sánchez
Dr. M. Elena Rodríguez González

Barcelona

March 11, 2018

The drawings used in the cover of this work are taken from the novel *Alice's Adventures in Wonderland*, originally published in 1865, written by Lewis Carroll and illustrated by John Tenniel.

Follow the white rabbit.

Acknowledgements

Durant la realització d'aquest treball he estat en contacte amb una sèrie de persones que, d'una manera o altra, han tingut un efecte important sobre el resultat d'aquesta recerca, i a qui vull donar les gràcies.

Primerament, voldria donar les gràcies a les dues persones que, des del principi del camí, han estat al peu del canó: la Dra. Maria Antònia Huertas Sánchez i la Dra. Maria Elena Rodríguez González. Des dels inicis pantanosos, on surava entre propostes de recerca, fins als retocs finals de la tesi, on acabava de llimar els últims detalls de fórmules i paràgrafs, totes dues han aguantat estoicament reunions, divagacions filosòfiques i estrès matemàtic en estat pur: és gràcies a la seva paciència i dedicació que he aconseguit dur aquest treball a bon port. Vull agrair especialment a l'Antònia la seva inacabable paciència en guiar-me a través de les jungles matemàtiques més profundes, i a l'Elena la seva entrega i rigor a l'hora d'examinar amb lupa cada paràgraf, frase, paraula i lletra de tot el que jo escrivia: el sobrenom de "el ojo sin párpado" no podria ser més merescut!

I would also like to thank all the people who I met during my two research stays at the Institute for Logic, Language and Computation (ILLC), from the Universiteit van Amsterdam, and with the Logic and Epistemology group, at the Ruhr-Universität Bochum.

From the ILLC, I would like to thank all the students of the Master of Logic who I met while attending the courses, with whom I had very interesting conversations that contributed to my research. Furthermore, I would like to specially thank Dr. Dick de Jongh and Dr. Sonja Smets for hosting me at the ILLC, for their interest and their feedback. Last, but not least, I would like to thank Dr. Franz Berto, with whom I had a very interesting meeting regarding my interest in logic and imagination: his insights pointed me towards the direction that I followed throughout the rest of the present dissertation.

From the Ruhr-Universität Bochum, I would like to thank all the people from the Logic and Epistemology group, both at a professional and personal level, and for the valuable feedback they gave me during my research stay there. In particular, I want to give my special thanks to Dr. Heinrich Wansing, who accepted my request for the research stay, and with whom I had various meetings that allowed me to tackle the philosophical part of this dissertation with renewed insights and energies. I could say, beyond any doubt, that the second part of this dissertation results from the work and feedback I received during my stay there.

D'altra banda, aquesta tesi no hauria estat possible sense el finançament que he rebut, primerament, d'una beca doctoral de la Universitat Oberta de Catalunya. A més, com a part d'aquesta beca doctoral, així com del suport rebut des dels Estudis d'Informàtica, Multimèdia i Telecomunicacions, he tingut la possibilitat d'assistir a congressos, cursos i estades de recerca on he pogut compartir els resultats obtinguts en aquesta tesi amb experts d'arreu del món.

Así mismo, desde el inicio de este proyecto he tenido la oportunidad de colaborar con el proyecto "Lógica Intensional Híbrida", financiado por el Ministerio de Economía y Competitividad (FFI2013-47126-P), y encabezado por la Dra. María Gracia Manzano Arjona. Quiero agradecerle sinceramente la oportunidad de colaborar con este proyecto.

Finalment, i en un àmbit més personal, voldria donar les gràcies a totes les persones que em rodegen i que, d'alguna manera o altra, han contribuït a fer que tant aquesta tesi com un servidor poguéssim arribar al final d'una sola peça. Fer-ne una llista exhaustiva seria massa llarg, però m'agradaria dedicar una menció especial a la gent més propera: a la meva mare, per la seva entrega i els seus ànims; al meu pare, per tot el que em va transmetre i per fer-me una gran part de qui soc (en especial la vessant artística i creativa); als meus germans, pel seu suport; als meus amics més propers, per tenir la paciència d'escoltar els meus deliris de lògica i filosofia; i, molt especialment, a l'Anna, amb qui comparteixo tant, per ser al meu costat a cada moment.

Moltes gràcies a tots!

Joan Casas Roma

Abstract

Imagination has received a great deal of attention by philosophers, cognitive scientists and psychologists. Many of these studies focus on how imagination affects our decision-making abilities, our beliefs or our emotions. Nevertheless, our main interest is not related to how imagination can be used as tool for supporting our understanding of the real world; instead, we want to focus on studying what processes take place, whenever we create and develop imaginary worlds.

In this work, we contribute to a better understanding of imagination by performing a detailed analysis of its dynamic mechanisms at a philosophical, a formal and an applied level: we propose a new theory that accounts for the processes involved in the creation and development of imaginary worlds; we define two different dynamic logics that capture the such processes; we provide a prototype of a computer program that implements the dynamics of one of such logics, and allows to simulate the way human beings perform acts of imagination; in addition, we point out the utilities that our contributions can have in fields such as video games.

We review three influential theories of imagination that identify different mechanisms involved in acts of imagination and, as a first contribution, we show how all of them share a similar structure. Then, we review the existing logics aimed to represent imagination, and we note that none of them can account for our main goal: to capture the dynamic mechanisms involved in an act of imagination.

Our second contribution is the Logic of Imaginary Scenarios, in which we define a layer for imagination acts upon a single-agent epistemic logic. This layer includes a new dynamic operator that calls an algorithm describing how new imaginary worlds are created in the models.

While discussing the properties of the Logic of Imaginary Scenarios, we note that the way imaginary worlds are developed is oversimplified. When revisiting the previously reviewed theories of imagination, we realize that their account for "reality-oriented development" collapses two distinct mechanisms that should be distinguished. A deeper

analysis leads to the third contribution of this work, which is the definition of a new theory specially suited for the dynamics of imagination acts, called the Common Frame for Imagination Acts; furthermore, we define a tool that can be used, together with our theory, to classify imagination acts according to their dynamics: the Rhombus of Imagination.

With this new theory at hand, the fourth contribution of this work is the Logic of Imagination Acts, in which we introduce four different dynamic operators and four different algorithms. In this new approach, an act of imagination is no longer seen as a *single* execution of a kind of brute-force algorithm rendering every possible alternative, but it is rather understood as a sequence of executions of smaller processes. This change not only accounts for the fact that the same imagining can be developed in different ways, but also captures the agentiveness required to decide how to develop a certain imaginary world.

Finally, we discuss and foresee possible applications of our work. As the fifth contribution, we provide an implementation of a prototype of a computer program that captures the algorithms defined by our Logic of Imagination Acts and allows the user to execute them in a formal model. Aside from this implementation, we also consider other areas in which the importance of the creation and development of imaginary worlds is critical, such as in video games. Due to this, we argue how a detailed understanding of the processes involved in imagination acts can be a valuable source of information for studying how players immerse in the virtual worlds they create.

To sum up, we claim that the theoretical and formal contributions we make in this work represent valuable contributions towards a better understanding of the dynamics involved in the creation and development of imaginary worlds, which also lead to a deeper understanding of how imagination works.

Keywords: imagination; dynamic imagination; imaginary worlds; logic of imagination, dynamic logic, algorithms.

Resum

La imaginació ha rebut molta atenció en els àmbits de la filosofia, la ciència cognitiva i la psicologia. Molts d'aquests estudis se centren en els efectes de la imaginació a l'hora de prendre decisions, en les nostres creences o en les nostres emocions. El nostre interès principal, però, no té a veure amb la utilització de la imaginació com una eina per donar suport a la nostra comprensió del món real, sinó que volem centrar-nos en l'estudi dels processos que tenen lloc quan creem i desenvolupem mons imaginaris.

En aquest treball, oferim una anàlisi detallada dels mecanismes dinàmics de la imaginació, tant des d'un punt de vista filosòfic com formal i aplicat: proposem una nova teoria que identifica els processos involucrats en la creació i el desenvolupament de mons imaginaris; definim dues lògiques dinàmiques diferents que capturen la formació de nous mons imaginaris; presentem el prototip d'un programa informàtic que implementa la part dinàmica d'una d'aquestes lògiques i que permet simular la manera com els éssers humans duem a terme actes d'imaginació; a més, identifiquem les aplicacions que les nostres contribucions poden tenir en camps com ara el dels videojocs.

Revisem tres teories influents sobre la imaginació que identifiquen diferents mecanismes involucrats en actes d'imaginació i, com a primera contribució, mostrem com totes elles comparteixen una estructura semblant. A continuació, revisem les lògiques existents encarades a representar la imaginació, i veiem com cap d'elles pot representar el nostre objectiu principal: capturar els mecanismes dinàmics involucrats en un acte d'imaginació.

La nostra segona contribució és la Lògica dels escenaris imaginaris, en la qual, partint d'una lògica epistèmica per a un sol agent, definim una capa per als actes d'imaginació. Aquesta capa inclou un operador dinàmic que crida un algoritme que descriu com es poden crear nous mons imaginaris als models.

Tot discutint les propietats de la Lògica dels escenaris imaginaris, veiem que la manera en què els mons imaginaris es desenvolupen està massa simplificada. En revisar de nou les anteriors teories sobre la imaginació, ens adonem que la manera en què capturen els "mecanismes de desenvolupament basats en la realitat" aglutina dos mecanismes diferents

que caldria distingir. Una anàlisi més profunda ens porta a la tercera contribució d'aquest treball, que és la definició d'una nova teoria especialment dirigida a les dinàmiques dels actes d'imaginació, anomenada Marc comú per a actes d'imaginació; a més, definim una eina que es pot utilitzar, juntament amb la nostra teoria, per classificar actes d'imaginació en funció de les seves dinàmiques: el Rombe de la imaginació.

Amb aquesta nova teoria a les mans, la quarta aportació del nostre estudi és la Lògica dels actes d'imaginació, en la qual introduïm quatre operadors dinàmics i quatre algoritmes diferents. En aquesta nova aproximació, un acte d'imaginació ja no es veu com una sola execució d'un tipus d'algoritme que, per força bruta, calcula cada alternativa possible, sinó que s'entén com una seqüència d'execucions de processos més petits. Aquest canvi no només captura el fet que un mateix acte d'imaginació es pugui desenvolupar de maneres diferents, sinó que també captura l'agentivitat necessària per decidir com desenvolupar un món imaginari concret.

Finalment, discutim i anticipem possibles aplicacions del nostre treball. Com a cinquena contribució, presentem la implementació del prototip d'un programa informàtic que captura els algoritmes definits per la nostra Lògica dels actes d'imaginació, i que permet a l'usuari executar-los en un model formal. A part d'aquesta implementació, també considerem altres camps on la importància de la creació i el desenvolupament de mons imaginaris és crucial, com ara els videojocs. Per això, argumentem com una comprensió detallada dels processos involucrats en els actes d'imaginació pot ser una font d'informació molt útil per a l'estudi de com els jugadors participen en aquests mons virtuals.

En resum, considerem que tant les contribucions teòriques com formals que fem en aquest treball són aportacions de valor per a una millor comprensió de les dinàmiques relacionades amb la creació i el desenvolupament de mons imaginaris, cosa que també comporta una comprensió més profunda del funcionament de la imaginació.

Paraules clau: imaginació; imaginació dinàmica; mons imaginaris; lògica de la imaginació, lògica dinàmica, algoritmes.

Resumen

La imaginación ha recibido una gran atención en los ámbitos de la filosofía, las ciencias cognitivas y la psicología. Muchos de estos estudios se centran en los resultados de la imaginación; esto es, como afecta a la toma de decisiones, a nuestras creencias o a nuestras emociones. Sin embargo, nuestro interés principal no concierne a la utilización de la imaginación como una herramienta para dar soporte a nuestra comprensión del mundo real, sino que se centra en el estudio de los procesos que tienen lugar al crear y desarrollar mundos imaginarios.

En este trabajo ofrecemos un análisis detallado de los mecanismos dinámicos de la imaginación, tanto desde un punto de vista filosófico como formal y aplicado: proponemos una nueva teoría que identifica los procesos involucrados en la creación y el desarrollo de mundos imaginarios; definimos dos lógicas dinámicas diferentes que capturan la formación de nuevos mundos imaginarios; presentamos el prototipo de un programa informático que implementa la parte dinámica de una de estas lógicas y que permite simular la manera como los seres humanos llevamos a cabo actos de imaginación; finalmente, identificamos las aplicaciones que nuestras contribuciones pueden tener en campos como el de los videojuegos.

Revisamos tres teorías influyentes sobre la imaginación que identifican diferentes mecanismos involucrados en actos de imaginación y, como primera contribución, mostramos que todas ellas comparten una estructura similar. A continuación, revisamos las lógicas existentes que representan la imaginación, y vemos como ninguna de ellas puede representar nuestro objetivo principal: capturar los mecanismos dinámicos involucrados en un acto de imaginación.

Nuestra segunda contribución es la Lógica de los Escenarios Imaginarios, en la cual, partiendo de una lógica epistémica para un solo agente, definimos una capa para los actos de imaginación. Esta capa incluye un operador dinámico que invoca a un algoritmo que describe cómo se pueden crear nuevos mundos imaginarios en los modelos.

Al discutir las propiedades de la Lógica de los Escenarios Imaginarios, vemos que

la forma en que los mundos imaginarios se desarrollan está demasiado simplificada. Al revisar de nuevo las anteriores teorías sobre la imaginación, nos damos cuenta que la forma en que capturan los "mecanismos de desarrollo basados en la realidad" aglutina dos mecanismos diferentes que habría que distingir. Un análisis más profundo nos lleva a la tercera contribución de este trabajo, que es la definición de una nueva teoría especialmente dirigida a las dinámicas de los actos de imaginación, llamada Marco Común para Actos de Imaginación; además, definimos una herramienta que se puede utilizar, junto con nuestra teoría, para clasificar actos de imaginación en función de sus dinámicas: el Rombo de la Imaginación.

Gracias a esta nueva teoría, la cuarta aportación de nuestro estudio es la Lógica de los Actos de Imaginación, en la que introducimos cuatro operadores dinámicos y cuatro algoritmos diferentes. En esta nueva aproximación, un acto de imaginación ya no es visto como una sola ejecución de un tipo de algoritmo que, por fuerza bruta, calcula cada alternativa posible, sino que se entiende como una secuencia de ejecuciones de procesos más especializados. Este cambio no solo captura el hecho de que un mismo acto de imaginación se pueda desarrollar de formas distintas, sino que también captura la agentividad necesaria para decidir cómo desarrollar un mundo imaginario concreto.

Finalmente, discutimos y anticipamos posibles aplicaciones de nuestro trabajo. Como quinta contribución, presentamos la implementación del prototipo de un programa informático que captura los algoritmos definidos por nuestra Lógica de los Actos de Imaginación, y que permite al usuario poder ejecutarlos en un modelo formal. Aparte de esta implementación, también consideramos otros campos en donde la importancia de la creación y el desarrollo de mundos imaginarios es crucial, como en los videojuegos. Por ello, discutimos como una comprensión detallada de los procesos involucrados en los actos de imaginación puede ser una fuente de información muy útil para el estudio de cómo los jugadores participan en estos mundos virtuales.

En conclusión, consideramos que tanto las contribuciones teóricas como las formales que hacemos en este trabajo son aportaciones de valor para una mejor comprensión de las dinámicas relacionadas con la creación y el desarrollo de mundos imaginarios, aspecto que también conlleva una comprensión más profunda del funcionamiento de la imaginación.

Palabras clave: imaginación; imaginación dinámica; mundos imaginarios; lógica de la imaginación, lógica dinámica, algoritmos.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

> "Would you tell me, please, which way I ought to go from here?" "That depends a good deal on where you want to get to," said the Cat. "I don't much care where–" said Alice. "Then it doesn't matter which way you go."
>
> —Lewis Carroll
>
> *Alice's Adventures in Wonderland*

Imagination has received a great deal of attention in the areas of philosophy of mind, psychology and cognitive sciences, specially during the last years —as it can be seen in works like [52], [22], [42], or [34], for instance. Our capacity to entertain alternative worlds and states of affairs that are not actual is often studied as being a tool that supports, for example, our decision-making abilities, or our epistemic and doxastic attitudes —like what we know and what we believe in under uncertainty. Through hypothetical and counterfactual reasoning (what would happen if such and such was the case), imaginative rehearsal (imagining oneself carrying out certain activity) or by imagining the contents of someone else's mind (often known as *mindreading* in the literature), we can use imagination as a powerful tool for supporting our interaction with the real world. Many of those works, therefore, focus on studying the *outcomes* that typically follow from imagination, specially when they affect one's beliefs, behavior or emotions.

There is a specific kind of imagination acts that are specially interesting, with respect to our relation with the real world, which are *voluntary* acts of imagination; that is, acts of imagination that are consciously initiated by an agent willing to entertain certain imagining. This kind of acts of imagination differ from those imaginings that pop out into one's mind, with no apparent reason or, at least, without being intentionally initiated by oneself. In voluntary acts of imagination, the agent consciously creates an imagining representing a certain situation she wants to entertain, develop, and maybe even extract conclusions from it. In this work, we focus on this particular kind of imagination acts.

Nevertheless, our main interest in the present work is not related to how voluntary imagination can be used as a supporting tool for our interaction with the real world; instead, we want to study the processes that take place whenever we create and develop an imaginary world —either as a resource for supporting our real-life decisions, or simply as a way of fantasizing, daydreaming or engaging in pretense play.

1.1 Motivations

This work aims to contribute, by deepening our understanding of the dynamic mechanisms involved in the creation and development of imaginary worlds, to the overall understanding of the human capacity to imagine.

Even though it is important to study and understand how we use the conclusions we reach while evaluating how an imaginary world would be like, and how those conclusions affect our real-life decisions, it is equally important to understand how we create and develop the details of such imaginary worlds. Which specific mechanisms do our mind follows, in order to do so? Is there any detailed, identifiable logic behind those mechanisms? If so, could those mechanisms be represented and simulated by using an algorithm? How do our knowledge and beliefs about the real world affect the way in which our imaginings develop? If it tuns out that, in the end, there is indeed an underlying logic governing the creation and elaboration of imaginary worlds, why do different people come up with different imaginings from a similar initial situation? By performing a detailed analysis of the dynamics of voluntary imagination acts, we aim to provide a better understanding about imagination as a whole.

The motivations, in this work, come both from the topic, which is the study of the dynamics of creating and developing imaginary worlds, but also from its inter-disciplinary approach, which benefits and puts together a plurality of disciplines in order to provide a bigger, clearer and more detailed picture of the topic at work. In particular, our approach to the topic at hand will be made from a philosophical, formal and algorithmic point of view: we gather the basics of the dynamics of voluntary imagination from philosophical theories, and we then translate those high-level intuitions into concrete formal languages, while being aided by an algorithmic representation of their dynamic aspects. Furthermore, this interdisciplinary approach builds a feeding channel between all the disciplines involved: the philosophical analysis provides the high-level requirements for the formal system which, in turn, provides, once those high-level requirements are translated into low-level specifications, the basis for an algorithmic approach to the issue at hand. Similarly, the insights gained by the both the formal and the computational approach throw light into the previous philosophical intuitions.

1.2 Objectives

The main goal in this work is to study and analyze how imaginary worlds are created and developed as a result of voluntary acts of imagination, and which are the particular mechanisms involved in doing so. After setting this main goal, we identify the particular objectives embedded into it.

O1. To identify, through a critical review of some of the most influential theories detailing how voluntary acts of imagination work, the mechanisms involved in the creation and development of imaginary worlds.

O2. To define a formal system capable of capturing, through a dynamic process captured by an algorithm, the mechanisms involved in the creation and development of imaginary worlds.

O3. To consider the applications that this formal system and the algorithm defined in it could have, both when implemented as a computer program, and in relation to other fields where imaginary worlds has a critical importance.

From this set of objectives, we expect that our work will contribute to a better understanding of imagination, in general, and to its dynamics, in particular, in three different areas. Our critical review on philosophical works is aimed to provide a detailed definition of the mechanisms involved in creating and developing imaginary worlds. Then, the definition of a formal system will contribute to the topic at hand by providing a more detailed and concrete representation of how those dynamic mechanisms behave. Lastly, the uses of this formal system will be considered, both as a computer-based approach to a representation of the human imagination, and also as a tool that can be used to study this phenomenon in other fields.

1.3 Methodology

The methodology followed in this work combines different methodologies, due to its interdisciplinary approach and to the variety of questions it aims to solve.

Firstly, an important part of the thesis is devoted to the study, the analysis and, further on, the refinement of philosophical accounts of voluntary imagination acts. The methodology followed throughout this part is characterized by a critical analysis of the theories defined in some existing works, followed by philosophical discussion, and motivated by specific examples and intuitions derived from those theories. This philosophical analysis is supported by quasi-formal representations of the contents of such theories, which serve a double purpose:

- To allow for a more clear abstraction of the contents of those theories, providing a better way of comparing them afterwards.
- To serve as a starting point when identifying the mechanisms that a formal logic would require, in order to account for such theories.

As a result of combining this philosophical analysis and discussion with a quasi-formal approach, we argue that the philosophical part of this work also benefits its formal part, and the other way around.

Secondly, another relevant part of this work is devoted to the definition of a formal logical system aimed to capture the dynamics of imagination acts, as recognized in the former philosophical part. The methodology followed in this part comprises two different approaches:

- To translate and express our philosophical, high-level requirements and intuitions behind the dynamics of imagination acts into a precise formal system, using mathematical logic structures (formal syntax and semantics), following a deductive approach (which is the methodology commonly used in the field of formal logic), and using mathematical proofs when needed.
- To support and enhance the previous formal system by combining it with techniques of algorithm design.

By combining these two methodologies, the formal part of this work aims to reproduce the results of the former philosophical part. Furthermore, by merging the usual mathematical approach to formal logics with techniques derived from the definition of algorithms, we argue that our work on formal systems can be a useful way of feeding back the philosophical setting we used to define it.

Thirdly, the last part of our work is focused on developing a computer implementation of the results obtained in the formal part. In this case, we use the Waterfall Model methodology for software engineering; in brief, this methodology involves the analysis of the requirements of the program to be implemented, a design of the elements needed, and an implementation that follows from the selection of the technologies and programming languages better suited for the occasion. Once all this is set, it is a matter of translating, into the appropriate computer languages, the required elements from the formal setting.

1.4 Document Layout

The present work is composed by 9 different chapters, which are then divided and grouped in two different parts, structured as follows:

- The present chapter, corresponding to Chapter 1, introduces the topic, the motivations, the main objectives and the contributions expected to achieve, the methodologies that will be used throughout it and the content of each chapter.

- *Part I: Down the Rabbit Hole.* The first part of this work is devoted to the review and study of existing theories and logics for imagination, and to the definition of a first formal approach to the dynamics of imagination.

  - In Chapter 2 we perform a literature review on the topic of imagination. We start by introducing certain terms and classifications normally used in the topic at hand; then, we present how imagination, as a mental attitude, is often understood within the human mind and in relation to other mental attitudes. Afterwards, we focus on reviewing the influential *cognitive theory of pretense*, proposed by Shaun Nichols and Stephen Stich in [43], and which provides a detailed account of the mechanisms involved in the creation and development of imaginary worlds within the mind of an agent; furthermore, we also review two other theories accounting for the dynamics of voluntary imagination, and which can be related to Nichols and Stich's theory. This underlying similarity, then, is made explicit by identifying a theory-independent structure detailing three different mechanisms that intervene in voluntary acts of imagination.

- In Chapter 3, we start by briefly introducing some basic notions of formal logic: in particular, we review the basics of propositional and first-order logic, and how and why modal logic is needed to account for certain nuances in meaning; then, we introduce the benefits that hybrid logic carries over modal logic, and we provide details on a particular kind of modal logic specially suited to represent knowledge. After that, we revisit existing logical systems capable of accounting for imagination, and we focus on both how they understand imagination, and what they highlight about it. After considering how these logics deal with imagination, we argue that none of them is suitable to account for our main goal: to define a formal system able to capture the dynamics involved in the creation and development of imaginary worlds.

- In Chapter 4, we present our first approach to the definition of a dynamic logic able to capture imagination acts. We identify the requirements that an algorithm for imagination should have, based on the mechanisms recognized at the end of Chapter 2. In order to be able to account for other mental attitudes, we decide to build our logic upon the already existing single-agent epistemic logic, which we already introduce in Chapter 3; after defining the syntax and the semantics of our system, we provide a detailed account of the algorithm for imagination acts, together with a step-by-step example. We devote the last part of this chapter to discuss where our logic succeeds, and where it falls short, with respect to our initial desiderata. We argue that, although the approach is on the right track, there are certain shortcomings that have to be amended in order to increase the level of detail in which we want to capture how imaginary worlds are created and developed.

- In Chapter 5, we make our concerns explicit. Even though defining a dynamic logic of imagination by combining a formal system and an algorithm is a sound approach, the mechanisms represented in our logic still lack depth. As we argue in this short interlude, this realization represents a spin in our work, which makes us reconsider the theories of imagination from their very basis, and which marks the end of this first part of this dissertation.

• *Part II: Deeper Down the Rabbit Hole.* The second part of this work goes over the same structure it already did on the first part, but now, it does with the backpack filled with feedback gained both from the previous theoretical review, and specially from the evaluation of the formal logic defined before.

- In Chapter 6, we argue how, while trying to translate the high-level intuitions of the theories of imagination into low-level algorithmic specifications, we recognized that we needed an even more detailed account of the dynamics of imagination. In particular, we show how the way in which the existing theories define the mechanism devoted to elaborate an imaginary world using reality-oriented rules is not fine-grained enough, and that it actually collapses two different ways of increasing the details of an imaginary world. Due to this, we engage in a philosophical analysis that goes deeper down the dynamics of imagination acts, and we propose a new theory specially suited to account for them. Furthermore, we argue how the precise distinction that we make in our theory regarding the dynamics of imagination allows us to draw a distinction, at a procedural level, between kinds of imagination acts that are indeed different, but which could not be distinguished by the previous theories. In order to aid this procedural analysis of imagination acts and make it more straightforward, we define a tool aimed to compute the "dynamical blue-print" of any act of imagination.

- In Chapter 7, we take our own theory for voluntary imagination acts as the underlying setting for a new approach to a formal logic of imagination. After introducing the logic's language and semantics, we define, this time, four different algorithms responsible for capturing the different processes involved in the creation and development of imaginary worlds; then, we provide a detailed example showing how each of these algorithms work, and we explain how the mathematical models of this logic could be used to automatically compute the dynamical blue-print of their corresponding imagination act, by using the tool defined in the previous chapter. While discussing the features of this new approach to a formal logic, we argue how the refinement in the dynamics of creating and developing imaginary worlds has led to a much better representation of voluntary imagination acts, and how this new modular approach to their dynamics defines a formal system much more suited to account for our goal than our previous logic.

- In Chapter 8, we move on the third and last topic involved in our work: applications for our theoretical and formal contributions. Due to the spin that took place in Chapter 6, and which fueled this whole new part, we did not consider the applications that the logic defined in Chapter 4 could have. Now, after having provided the definition of a new, more modular approach to a formal system for the dynamics of imagination, it is time to do so. We present an implementation of a computer prototype that captures the four algorithm defined by our latter logic; this prototype can be used both to test how the algorithms in that logic behave, and thus it is a valuable source of feedback for the logic itself, but it can also be used as a way of simulating how human beings create and develop imaginary worlds within their mind, as define by the theory we present in Chapter 6, and by the logic capturing this theory in Chapter 7. After that, we discuss how our contributions to the understanding of the dynamics of imagination can be also valuable in the field of video games. In particular, a detailed account of how imaginary worlds are created and developed could be a useful setting for studying the immersion of players within the virtual worlds defined in some genres of video games.

- In Chapter 9, we summarize the conclusions of our work. Firstly, we argue that we have indeed managed to achieve our initial objectives. Then, we present the different contributions that result from this dissertation (including a list of all the dissemination that derives from these contributions), we discuss our conclusions, specially regarding the interdisciplinary approach we followed in this work, and we also point out to some interesting lines of future research.

Part I:

Down the Rabbit-Hole

In another moment down went Alice after it, never once considering how in the world she was to get out again.

—Lewis Carroll
Alice's Adventures in Wonderland

Chapter 2

Philosophy and Imagination

In this chapter, we settle the theoretical background we are going to use throughout the rest of this work. We start by briefly reviewing how terms like "imagination", "conception" or "pretense" are usually understood by different authors, and we set our stance regarding how we use them. Once the terminology is clear, we place imagination within the mind, and we discuss its relation with other mental attitudes such as belief or desire. Then, we introduce and review an influential theory detailing how the dynamics of imagination acts work, and we compare it to two other recent theories that, we argue, coincide in great measure in their procedural analysis of imagination. We end this chapter by identifying how, according to these theories, acts of imagination work: the mechanics of this process set the ground of our desiderata, and we move on the the next chapter to start reviewing the formal background needed throughout the rest of this work.

2.1 Terms and Classifications

"Imagination" is a term that is often used as a generalization of other more fine-grained mental activities, specially in informal discussions. Some authors in the literature use this term in a very specific way and require that it fulfills a certain set of properties, whereas others use it in a more general way. Before digging deeper into how we use the term "imagination" in this work, we provide a quick overview of how different activities often referred to as "imagination" can be distinguished; then, we set our stance regarding the use of terminology and we set how we are going to use this and other terms throughout this work.

The term "imagination" is often used to refer to a set of mental activities whose input does not depend on the current, actual stimulus environment or reality (that is: it is not usually triggered by our sensory perceptions). When used in this general way, it embraces a wide variety of mental activities, such as forming mental images of things that may or may not exist in reality, counterfactual reasoning involving hypothetical situations, reasoning about alternative pasts and possible futures, daydreaming and fantasizing, pretending and games of make-believe, understanding and predicting others' actions and mental activities (often called *mindreading* in the literature) and mental rehearsal of activities, among others. When fine-graining the term, "imagination" is sometimes distinguished by some authors from terms like *conceiving* or *supposing* on the grounds that "imagination" is sometimes required to involve the use of "mental imagery", whereas the others do not. One of the examples commonly used (as in [21]) is the one saying that, although you can *conceive* a chiliagon (a polygon with a thousand sides), you cannot really form a mental image of it, and thus you cannot *imagine* it. However, and although this distinction could be clearly drawn when referring to physical entities, it is not so clear with *propositional imagination*; that is, "imagining that there is a chiliagon", instead of simply "imagining a chiliagon". In fact, as it appears in [42], recent discussions on propositional imagination claim that this distinction is just terminological, and contemporary cognitive accounts of imagination tend not to treat it as imagistic. It is also worth noting that some authors speak of imagination and pretense interchangeably (as in [47]), whereas others take the former to be more *mentalistic*, and the latter to be more *behavioral* (see [22]). Table 2.1 briefly summarizes the distinctions that are sometimes required to distinguish different terms related to imagination.

Term	*Related to*
Imagining	Forming a mental image
Conceiving	Comprehension of concepts
Supposing	Temporarily assuming something to be the case
Pretending	Results in external, observable behavior

Table 2.1: Distinction between terms related to imagination.

The "content" of what is being imagined is also a way of classifying its different uses. In this sense, many philosophers distinguish between *propositional* imagination and *non-propositional* imagination; the latter can then be divided into *sensory imaginaning* or *objectual imagination* (see [35] and [58], respectively) and *active imagining* (see [52]).

More specifically, propositional imagination (imagining that φ) involves creating a mental representation of some state of affairs described by a certain proposition, which, nevertheless, does not require us to "picture" it within our mind; sensory or objectual imagination (imagining a) involves creating a mental, quasi-sensory representation of a certain entity or situation, which *does* require us to form a mental image of how this entity or situation would look like; active imagination (imagining α-ing) involves creating a representation of oneself carrying out a certain activity or undergoing a certain experience, and it implies a first-person mental relation to some behavior or phenomenological experience. Table 2.2 summarizes the distinction of the different uses of imagination, according to what its content is.

Use of imagination	*Represents*	*Expressed as*
Propositional imagination	State of affairs	Imagine that φ
Sensory imagination	Imagistic representation	Imagine an a
Active imagining	Carrying out an activity	Imagine α-ing

Table 2.2: Distinction between uses of imagination according to its content.

Once the terminology of imagination has been briefly presented, let's set our stance regarding them. Our interest in this work is towards specifying and understanding, with the help of formal tools, how people carry out acts of imagination; now, by "acts of imagination", we refer to any kind of mental action that involves creating a mental representation of something which is not actual, nor initiated by any perceptual stimulus. Therefore, and as we do not focus on any particular, fine-grained use of imagination, we will be speaking of "imagining", "conceiving", "supposing" or "pretending" interchangeably, without purporting any relevant difference between one or another, at least up to the point that concerns our work.

Regarding the "content" of those mental representations, our interest lies on propositional imagination: that is, "imagining that φ is the case". We focus on this kind of imagination because the formal approach we take later on this work towards imagination is through propositional logic, and so interpreting imagination in a propositional way gives us the appropriate setting for a rather straightforward translation into a logical language. Furthermore, and as authors like Wansing in [54] do, we will interpret expressions of the form "Alice imagines a unicorn" (which refer to a more objectual or sensory use of imagination) as "Alice imagines that there is a unicorn" (rendering it propositional). Similarly, we can interpret statements of the form "Alice imagines herself running" (which would correspond to an active use of imagination) as "Alice imagines that she is running". Our stance regarding imagination, thus, can be summarized by the following statement:

> *We use terms such as "imagine", "conceive", "suppose" or "pretend" interchangeably in order to refer to any kind of mental activity that involves creating a mental representation of a certain state of affairs (usually different from the actual state of affairs). Moreover, we assume that this state of affairs can be expressed by propositions describing it.*

Now, having set the way we use different terms and the content of the acts of imagination we are interested in, let's see how imagination fits within the theory of mind.

2.2 Overview on the Theory of Mind

Once we have settled the grounds regarding terms and taxonomies, let's move to the next topic, which concerns the architecture of the mind. How is the mind structured, regarding different mental attitudes and processes? Where is imagination, in particular? And, moreover, how does it interact with other mental attitudes? Is it functionally similar to other cognitive processes? Does it share some of its gears with other attitudes? Works in philosophy of mind have deeply explored the cognitive architecture of imagination and, in particular, how imagination differs from and resembles other mental states. In the following subsections, we provide an overview of some of the topics that concerns us the most.

2.2.1 Imagination and Other Mental Attitudes

In philosophy of mind, mental attitudes are often divided in two different groups (as in the Introduction of [34]), depending on how the content of such attitudes is related to "the outer world'.

Cognitive attitudes are those in which the direction of fit is said to be "mind to world", like in *believing*: one's beliefs are considered true whenever they match what is the case in the real world. If it turns out that the content of one's beliefs do *not* match what is the case in reality, then one's beliefs are considered to be false; in that case, beliefs that turn out to be false should be typically updated in order to match what is the case in the world for them to be considered true. In a nutshell, cognitive attitudes are those whose content must coincide with what is the actual case in order to be true.

Conversely, the direction of fit in *conative attitudes* is just the opposite, like in *desiring*: one's desires are fulfilled (or made true) when what is the case in the real world matches the content of one's desires. In this case, desires do not typically adapt to what is the actual case (although someone who could do that would always end up being content with her life), but the other way around: the world should adapt to one's desires to make them true or fulfilled. Moreover, and even though it is common for beliefs to be revisited and updated when the world does not match them, it is seldom the case that desires change and reshape themselves according to the way the world is (unless they are clearly seen to be impossible to satisfy, in which case the saddened bearer of these desires may decide to give up and look for new ones). In a nutshell, conative attitudes are those in which the actual case must coincide with their content in order to be true. A brief schema showing the direction of fit upon which these two groups of mental attitudes are evaluated can be seen in Figure 2.1; then, Figure 2.2 provides two examples[1] in which a belief and a desire are not fulfilled, and shows what and how should change in order for them to be fulfilled.

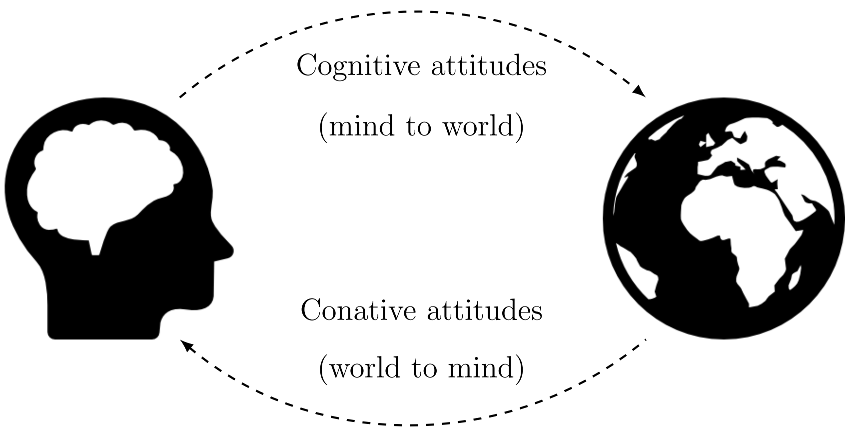


Figure 2.1: Direction of fit of cognitive and conative mental attitudes.

Specifically, imagination is considered to be a cognitive attitude (still in [34]), but with one particularity: in the case of imagination, the "relevant world of fit" is not the

[1]Icon acknowledgement: All icons have been taken from www.flaticon.com under a Creative Commons License CC 3.0 BY; the "brain", "Superman", "carrot" and "question mark" icons have been made by Freepik; the "Earth" icon has been made by Designerz Base; the "plane" icon has been made by GraphBerry; the "candy" icon has been made by Smashicons.

actual, real world, but rather an imagined or fictional world. Moreover, imagining is often distinguished from other cognitive attitudes, like *perceiving* or *believing*, on the grounds that imagining something does not require for it to be the case in the real world, whereas the former cases do (or, at least, they require that the subject *believes* that it is the case). In much the same way, imagination is also distinguished from conative attitudes like *desire* on the grounds that imagining something does not require the subject to wish for it to be the case in the real world.

Example: The agent *believes* that she sees a plane, but she is actually seeing Superman; therefore, her belief should be updated in order to match what is the case in reality.

Example: The agent *desires* to eat a candy, but she only finds a carrot; therefore, she must keep looking until reality matches what is the case in her desire.

Figure 2.2: Examples of unfulfilled (or false) cognitive and conative attitudes.

However, we think that imagination somehow challenges its classification into a cognitive attitude. Take, for instance, those acts of imagination that are initiated by someone willing to entertain a certain imaginary scenario: in those cases, we could distinguish between two different states within the act of imagination.

Firstly, acts of imagination are initiated by creating a new imaginary world. Recall that, when considering imagination, the world of fit is said to be, precisely, the imaginary world itself. Now, when an agent decides to entertain a certain imagining, the imaginary world created as a result of this is actually shaped according to a *desire* of the agent to imagine such and such; considering this, imagination seems to have more in common with conative attitudes in this case. In other words, the agent *wishes* to imagine something, and thus an imaginary world is created and shaped according to this wish of the agent: it is not that the imaginary world "accidentally" matches what the agent wants to imagine, but is precisely the *desire* of the agent to imagine such world what makes the imaginary world be like it actually is[2]. In this process, the direction of fit can be considered as being "(imaginary) world to mind", in the sense that the imaginary world must coincide to the wish the agent has for entertaining it; moreover, the imaginary world is actually created for that purpose.

Once the imaginary world has been created, then imagination may indeed be more closely related to cognitive attitudes. For instance, if the agent entertains the imaginary world in order to evaluate whether a specific course of events would probably take place in there, then beliefs also come into play. In other words, if the imaginary world has been created as a way of checking whether a certain outcome may be the case, then the direction of fit becomes "mind to (imaginary) world" again, as in cognitive attitudes. Unlike in the previous case, depending on how the agent believes that the imaginary world would move forward, and depending on how it actually does, it may be the case that the agent's beliefs do *not* match what she is imagining.

In this sense, and if we look closely to the process involving creating and entertaining an imaginary world, we could recognize how imagination can be classified as both kinds of mental attitudes, depending on the phase we are evaluating within the whole act of imagination. We will not pursue this issue further, as it deviates from the main goals of this work, but we believe that it is something worth taking into account when considering how imagination together with other mental attitudes within the mind. Our claim is that, depending on whether the imaginary world is being created, or already develop, then the way imagination works resembles more to how desires work, or how beliefs work, respectively.

2.2.2 Comparing Imagination and Belief

As we have already said, in this work we focus on propositional imagination, which brings our interest towards imagination even closer to beliefs. In particular, and regarding beliefs, it would be strange to say "Alice believes a unicorn"; rather, we would probably say "Alice believes that there is a unicorn". Similarly, we could express a more natural belief-

[2]See Chapter 10 in [42] for a discussion on cases of imaginative resistance in which, due to various reasons, it can be difficult or even impossible for an agent to imagine something she is asked to.

sentence like “Alice believes in unicorns” as “Alice believes that there exist unicorns”. Due to this, and when focusing on their propositional use, it is widely accepted that both imagination and belief range over the same domain, which is that of all understandable propositions (see [25]). Nevertheless, and although the range of both attitudes may be the same, imagination is often seen as departing from belief, in the sense that it is widely accepted that we can imagine things that we do *not* actually believe (and, conversely, we can also imagine things that we believe; see Chapter 1, Section 3.2 in [42]). Another way in which imagination strongly differs from belief concerns its intentionality: we can decide to imagine something, while we cannot usually decide what we believe in (although sometimes we may wish to believe something, but we still cannot, except in cases of self-deception).

Aside from these differences, imagination is strongly related with beliefs when it comes to defining how an imaginary world would look like. In particular, there are two features that connect imagination and beliefs and that seem to be present at almost every imaginative episode (see [25]): *mirroring* and *quarantining*. *Mirroring* refers to how imaginary worlds, when not willingly different, are usually close to the way the real world would be like, if the imagined situation was the actual case. More specifically, when certain details of an imaginary situation have not been explicitly settled by the content of the imagination attitude, people tend to infer these details by following the same rules and fact that would apply in the real-world analogue of the imagined situation. On the other hand, *quarantining* refers to the fact that what happens within an imaginary scenario does not usually affects one’s beliefs. A simple example of quarantining would be someone who imagines (for instance in a story-telling) that the world has been overrun by zombies: although that person could then imagine feeling scared in that imaginary world, she will not usually get scared in the real world... unless quarantine fails, which takes us to our next topic.

Mirroring and quarantining usually guarantee that our imaginings will resemble how the actual world would look like and, at the same time, they guarantee that whatever happens in an imaginary episode will not stain our beliefs. But can these mechanisms fail? As almost everything in the human nature, yes: from time to time, they also fail. In those cases, each mechanism gives rise to an “anomaly”. When the rules followed in an imagined scenario do not match those of its real-world analogue, mirroring gives way to *disparity*. In particular, imagined scenarios can be *incomplete* (there might be details of the imaginary world that are not specified at all during an imaginative episode) or even *incoherent* (elements of the imagined scenario might behave in ways they would never do in the real world, for instance when a coffee machine is used as a time-freezer in a game of make-believe). On the other hand, a failure of quarantine gives rise to *contagion*, which is characterized by the fact that things happening in an imaginary world affect our current beliefs: for instance, when someone imagines that there is a monster under her bed and she suddenly becomes scared in the real world. An overview of these two phenomena and the consequences of their failure is summarizes in Table 2.3.

	Mirroring	*Quarantining*
Ensures	Imagining is similar to reality	Imagining is not believed
Related anomaly	Disparity	Contagion
Anomaly leads to	Incoherent scenarios	Imagining affect our beliefs

Table 2.3: Features relating imagination and belief.

2.3 Existing Theories on Imagination

The main theory that we use as the background of our work is the influential *cognitive theory of pretense*, from Shaun Nichols and Stephen Stich, which provides a detailed account of the mechanisms involved in the creation of imaginary scenarios in pretense play. While reviewing other theories of imagination, we notice how some of them also identify certain dynamic mechanisms of voluntary imagination acts that are, in turn, closely related to Nichols and Stich's theory.

After reviewing those theories, we show how they account for the processes of imagination in a very similar way, which we summarize as a general mechanism responsible for carrying out acts of imagination. This mechanism is, precisely, what we aim to capture, using a formal system, in the next chapters of this work.

2.3.1 Nichols and Stich's Cognitive Theory of Pretense

Shaun Nichols and Stephen Stich develop in [43] a theory of pretense that aims to identify which mechanisms, within our minds, are involved when engaged in a pretense. Even though Nichols and Stich devote a part of their theory to understand and explain the behavior involved in an episode of pretense, our interest is limited to how episodes of pretense work on a representational level: namely, how a cognitive agent can go about imagining a world which is different from the actual[3]. Before going into the details of their theory, it is worth introducing two framework assumptions about the mind upon which they build their theory.

The first one is called the *basic architecture assumption*. It claims that a well-known commonsense account of the architecture of the cognitive mind is largely correct, but far from complete: namely, that the mind contains two quite different kinds of representational states, which are *beliefs* and *desires*. These two kinds of mental states are functionally different, in the sense that they interact with other components of the mind in different ways. A deeper look into this basic architecture assumption reveals an intricate system of procedures, mechanisms and "storage units" that, when put together, are used to create, manage, update and delete beliefs and desires. Some pieces of this architecture are devoted to tasks like processing perceptual stimulus (which directly affect our beliefs), monitoring the body (which, together with other mechanisms, affect our desires), or decision-making and action control systems (which are affected and affect both our beliefs and desires, and determine in turn our behavior).

We focus on two components that are specially relevant in Nichols and Stich's approach to pretense: the *Belief Box* and the *Inference mechanisms* (see Figure 2.3). The *Belief Box* should be intuitively understood (as the name strongly suggests) as a "box" in which the representation tokens of the agent's beliefs are stored; note, nevertheless, how on page 121 of [43], the authors point out that the box analogy is not be misinterpreted as implying some sort of spatial distinction within the brain, but it is merely used as a way of classifying and keeping apart representation tokens (such as propositions) related to different mental attitudes. The *Inference mechanisms* are a set of processes that, among other things, update and interact with the agent's beliefs —for instance, by removing inconsistent beliefs or sorting already existing beliefs when new information is acquired.

The second assumption, which they call the *representational account of cognition*, maintains that mental attitudes such as beliefs or desires are relational states. Therefore, having a belief or a desire with a particular content amounts to having a representation token of that content stored in the appropriate "box" within the mind. So, in order for someone to believe that she is going to win the lottery, she must have a representation token whose content is "I am going to win the lottery" in her Belief Box; similarly, when

[3]It is worth mentioning that in [43] (page 127) the authors note how their theory, without taking behavior into account, stands for a theory of imagination. In this sense, and recalling what we said regarding terminology in Section 2.1, Nichols and Stich distinguish "pretense" as being more behavior-oriented than "imagination". As said by the end of that section, and keeping in mind that our concern only refers to the representational level of these phenomena, we use both terms indistinguishably.

someone *desires* to win the lottery (which might be more often than believing so), a representation token containing the previous proposition is stored in her *Desire Box*.

After having introduced these basic assumptions, Nichols and Stich introduce three further components in the cognitive architecture of the mind that are needed in order to account for pretense episodes: the *Possible World Box*, the *UpDater* and the *Script Elaborator*. The full picture of the mind can be seen in Figure 2.3, which has been taken from [41].

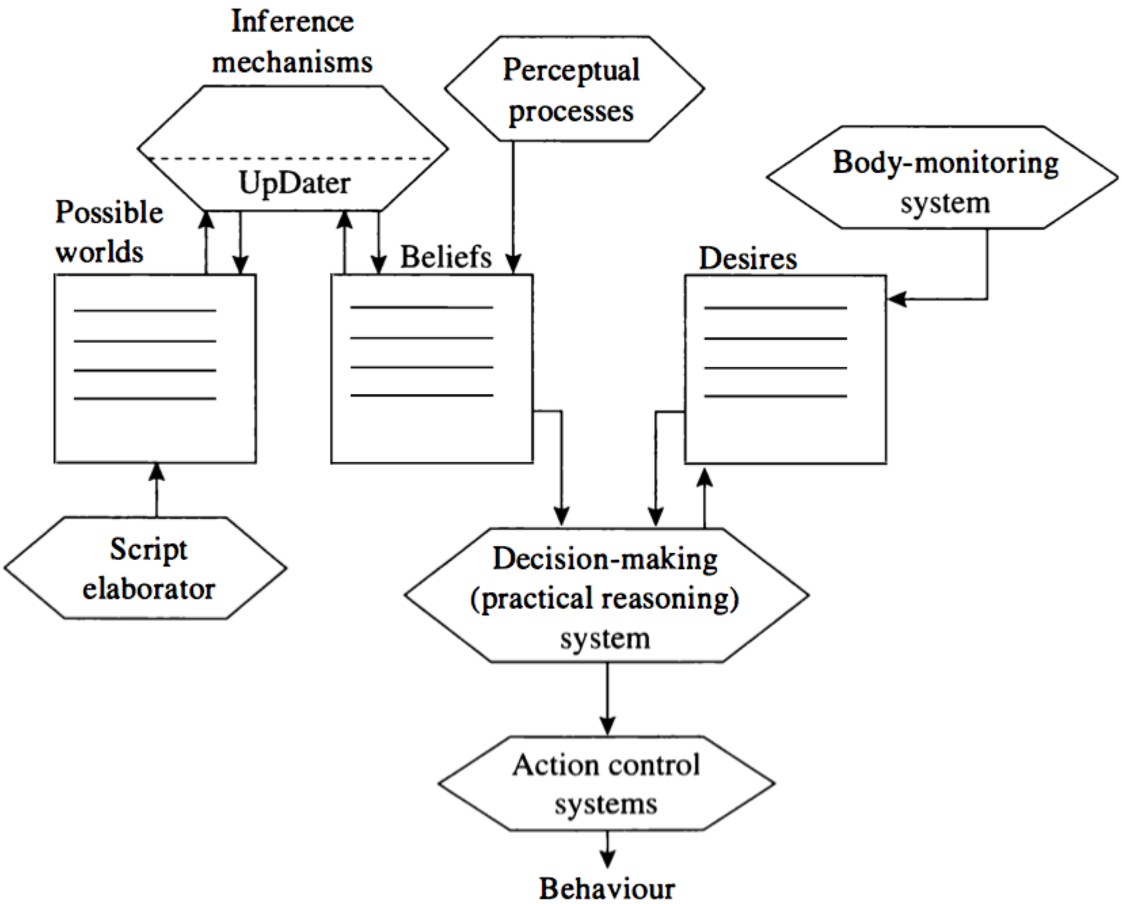


Figure 2.3: Nichols and Stich's architecture of the cognitive mind.

The *Possible World Box* (which will be called PWB onwards) is structurally and functionally very similar to the Belief Box and the Desire Box, and it also contains representation tokens. However, its main difference with respect to these other two boxes is that the PWB's job is not to represent the world as it is, nor as the agent wishes it to be, but rather to represent what the world would be like, given a certain set of initial premises[4]; similarly, these initial premises need not be true, nor wished to be true. Aside from being used in pretense episodes, the PWB is also key for a other tasks like mindreading, strategic reasoning or empathy, to name a few. Following Nichols and Stich, typical episodes of pretense are initiated with a pretense premise (determining what the pretense will be about) which is placed into the PWB workspace. So, for instance, when a pretense episode is initiated with the premise "this banana is a telephone", a representation token whose content is "this banana is a telephone" is placed into the PWB.

Once the initial pretense premise has been added to the PWB, the cognitive system starts to fill up the PWB with a detailed description of how the world would be like, if the pretense premise were true. As Nichols and Stich point out, the inference mechanism that takes care of this process is the same one that is used in the formation of our beliefs. So, in a pretense episode, new representations get added to the PWB by being inferred from representations that are already there, and by following the same inference mechanisms that would be used when reasoning about our beliefs. Nevertheless, there is only a limited amount of things one can infer from a certain set of premises, so what happens with everything else? What happens with details concerning things the initial premise is *not* about? All this extra information is taken from the agent's Belief Box; so, in this sense, initiating a pretense episode sets certain initial details about the scenario, but, regarding everything not explicitly set or inferable from the premises, there is no *a priori* reason to think it should be different to what the agent actually believes. However, there seems to be something problematic with this claim: if I *know* that the banana on the table is not a telephone, then what happens when I try to add this information to the pretense episode, as it would directly contradict the initial pretense premise? In order to explain this, the authors call for the *UpDater* mechanism.

The *UpDater* mechanism is a sub-system embedded in the *Inference mechanisms* earlier introduced, and which works both over the contents of the Belief Box and the PWB. The UpDater mechanism starts working as soon as a new premise is added into the PWB: it takes the agent's beliefs, dumps them into the PWB, and then it goes through all the representations in the PWB, while deleting those that are not compatible with the new premise that has just been added, and updating the information in those that are kept. It is worth remarking that, during this process, the initial premises have priority over what the UpDater is trying to import from the Belief Box: if a pretense episode is initiated by a premise saying "this banana is a telephone", beliefs contradicting this assumption

[4]Note that what Nichols and Stich refer to as *initial premises* is, in fact, equivalent to what Williamson refers to as *initial conditions*; we stick to each author's terminology in each section.

would not be added to the PWB. Following the terminology used by the authors in [41] (on page 32, when introducing their own cognitive theory of pretense), the initial premises in a pretense episode are "clamped" into the PWB. It is worth noting, though, that this "clamping" effect lasts only until the UpDater mechanism has finished working: after that, a new input in the pretense episode may override what was previously stated by a previously added pretense premise, and so initiate all over again the updating process.

Nichols and Stich point out how the contents of a pretender's Belief Box are not only representations about "facts" (say: "bananas are yellow", or "the Moon is made of cheese"), but also clusters or packets of representations constituting *scripts* or *paradigms* detailing how certain situations typically unfold. These scripts play an important role, both when guiding and constraining how situations in pretense episodes are filled up with details, and in determining how a pretense episode could advance (in the sense of what kind of things would usually happen, or how people involved in the pretense would typically behave in these sort of situations). Therefore, these scripts play an important role when trying to explain why pretenders behave in the way they do, when engaged in a pretense episode.

Aside from the way things may typically unfold, there are still many ways in which a pretense episode could develop that are neither specified by the initial pretense premises, nor by the pretender's beliefs, nor by any set of the sort of "default" scripts handled by the UpDater[5]. Although the authors do not provide a detailed account of its structure and the way it should work, they claim that there exists *some* mechanism (their emphasis) subserving this process, which they call the *Script Elaborator*. In a nutshell, its job is to fill the pretense episode with those details that cannot be inferred from the pretense premises, the contents of the Belief Box, the default scripts, or from what has happened earlier in the pretense episode. In other words, the Script Elaborator is responsible for handling any outcome, in an imaginative episode, that deviates from what one would typically expect. As we mention in the following subsection, authors like Peter Langland-Hassan express their concern with the way Nichols and Stich introduce the Script Elaborator: although they claim that it is responsible for handling a wide variety of imaginative outcomes, they introduce it as a kind of black box, and thus provide no insight whatsoever as to how it actually works. Langland-Hassan, on the other hand, provide an account as to how the Script Elaborator could work.

[5] Actually, the authors themselves point out in a footnote on page 127 in [43] that these script constrains are only "soft" constrains, and that they can be (and are) violated in pretense episodes quite dramatically.

Figure 2.4 represents[6] how Nichols and Stich's theory account for the way pretense episodes are created and elaborated. A new pretense episode is initiated by putting a pretense premise φ into the PWB, and then the scenario is filled up with details taken from the agent's beliefs regarding reality-oriented facts and scripts. Then, the Script Elaborator may also come up with a new pretense premise ψ which, similarly to the way the initial pretense is created, gets clamped into the pretense episode.

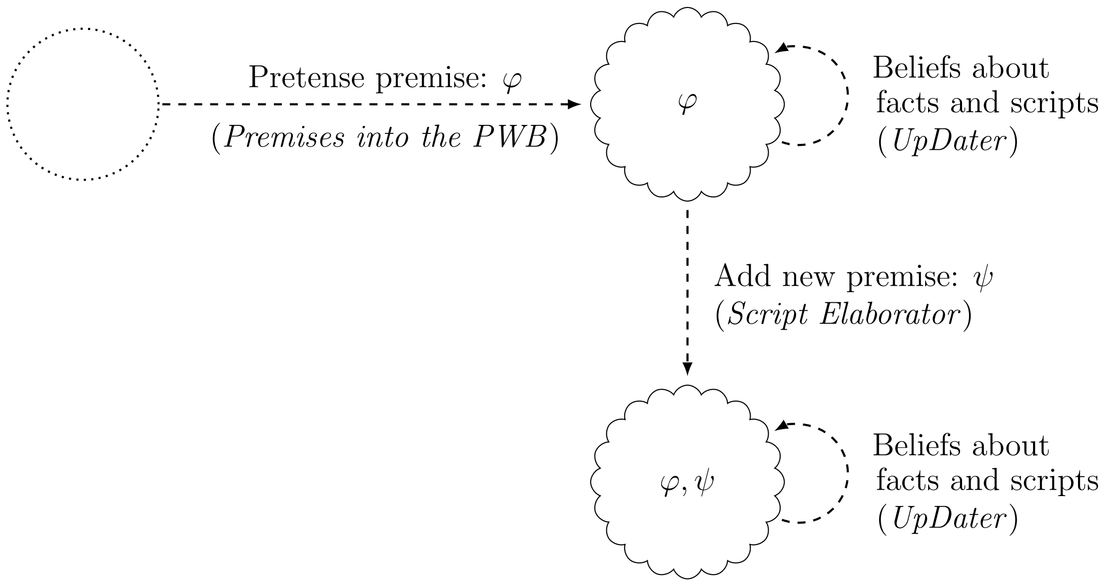


Figure 2.4: An act of imagination according to Nichols and Stich's theory.

2.3.2 Williamson's Modes of Imagination

In [56], Timothy Williamson argues against the apparent opposition that exists between knowledge and imagination, and that often appears in the literature; according to this distinction, knowledge is related to facts, whereas imagination is related to fiction[7]. Still

[6]We want to stress that, although our graphical representation resembles a Kripke model for Modal Logic (see Section 3.1.2 on Chapter 3 for more details), it is not such thing; rather, it should be simply understood as a sort of schema or flowchart representing certain actions that affect certain states of the system.

[7]It is worth noting that, although Williamson talks about knowledge, his work also refers to beliefs. The author analyzes the relation between imagining, in the sense of creating mental representations of non-actual states of affairs, and knowing, in the sense of having mental representations that describe how reality is, or how one *thinks* reality is. In this sense, imagination can be used as a tool to acquire new knowledge or new beliefs, or even to update previous beliefs. There is a vast amount of literature devoted to the definitions on knowledge and belief (see [55], for instance); in our work, nevertheless, we use both

driven by this distinction, imagination is often given the role of "raising possibilities", rather than assessing truth-values to propositions; thus, imagination seems to belong to the *context of discovery*, whereas rationality belongs to the *context of justification*. However, and as Williamson claims, imagination can be, when used properly, a powerful tool to acquire new knowledge. In fact, in most cases of uncertainty, imagination is the only tool we have available to form new beliefs and decide which course of action to take. However, in order to be used in such a way, imagination must be constrained by "reality-oriented rules and facts" that make imaginary scenarios unfold the same way they would do, if they were real. Therefore, in order to be regarded as a suitable tool that can be used in acquiring or improving one's knowledge and belief, imagination should not be completely independent from what one knows or believes about the world.

The author proposes a distinction between two different modes in which imagination can work: *voluntary* and *involuntary*. When in voluntary mode, one sets certain initial conditions that describe an imaginary scenario; for instance, if one wants to assess whether she would be able to jump a certain mountain stream, the imaginary scenario should describe the closest possible scenario matching that particular mountain stream, plus any other detail that might be relevant for the assessment of the person jumping over it.

Once these initial conditions have been set up, imagination starts running in involuntary mode and unfolds the consequences that would likely follow from the initially set scenario, if it was real. The key to the epistemic value of these imaginary exercises lies in the "if it was real": of course, one can choose to imagine (almost) anything she wants; so, even if the mountain stream is seven miles wide, one can imagine jumping, growing a pair of wings and flying swiftly to the other side of the stream. However, and entertaining as this imagining may be, this development of the initial scenario has little epistemic value, if any. This is, precisely, because it has not been developed by following the rules that it would presumably follow, if it was real. Therefore, in order for imagination to be used as a tool to acquire knowledge, the unfolding of the initial imaginary scenario should be made by following reality-oriented rules and facts[8]. So, in the mountain stream example, one should use her knowledge and beliefs about previous experiences involving jumps, while taking in account the particular details of the situation as well: do the rocks look slippery? Is there someplace where it looks safe to land after the jump? How tired is the person trying to assess whether she would be able to jump? Thus, one can only get

to refer to mental representations aimed to represent the actual states of affairs, and that are either the case, or thought by the agent to be the case.

[8]Note that this corresponds, precisely, to the *mirroring* effect we introduced in Section 2.2.2.

epistemically relevant results from these kind of imaginary experiments when one feeds involuntary imagination with reality-oriented rules and facts.

Nevertheless, and as Williamson points out, setting the right initial conditions and using the right rules in the involuntary unfolding of the scenario do not guarantee that the imagined outcome will actually match the real one: imagination is also fallible, as any other human cognitive capacities are, like perception or memory. But still, and also mimicking other human cognitive capacities, using reality-oriented imaginary experiments to form new beliefs or knowledge about is a useful tool of the human mind that, although being fallible, is also useful in many situations.

In Figure 2.5, we schematize how acts of imagination work, according to Williamson: the dotted circle at the left hand of the figure represents the fact that there is no imaginary situation at that moment, and it is not until performing an act of voluntary imagination (with certain initial conditions φ) that an imaginary world is created; then, the imaginary scenario is developed by the involuntary imagination, which follows reality-oriented rules and facts.

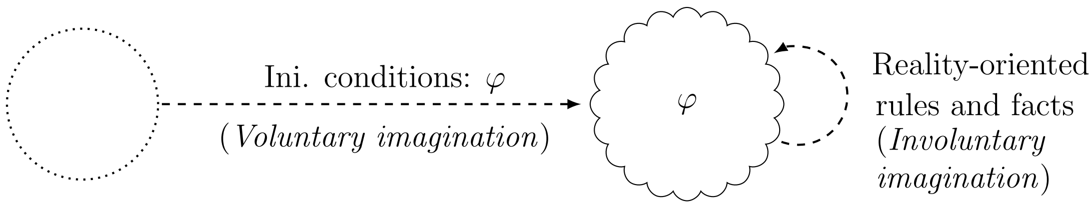


Figure 2.5: An act of imagination according to Williamson's theory.

When comparing Williamson's distinction of the two modes of imagination with Nichols and Stich's theory, in Figure 2.4, we can already anticipate how Williamson also recognizes, in the dynamics of voluntary imagination acts, an initial voluntary action that sets the conditions of the imaginary world, and a sort of automated, inferential-driven mechanism that elaborates on the details of such initial conditions. Nevertheless, and due to the fact that Williamson's interest does not concern additions that go beyond reality-oriented rules and facts, his theory lacks a mechanism matching Nichols and Stich's Script Elaborator.

2.3.3 Langland-Hassan's GC Imaginings

In [35], Peter Langland-Hassan analyses how it is possible that voluntary acts of imagination can be used to improve our epistemic state. In other words: if what we imagine is (in certain cases) determined by what we *want* to imagine, then how can it be that we can learn anything new at all, if that imagining was initiated by our own intentions?

In his work, Langland-Hassan distinguishes between four kinds of imaginings, sorted by two different properties: imaginings can be either *chosen* or *unbidden*, where the former are initiated voluntarily by the agent setting the initial content of the imagining, and the latter simply pop out into our minds; imaginings can also be either *guiding* or *misguiding*, where the former provide useful insights regarding how imaginary situations would be like, or how one should act in a certain situation, and the latter lead us to wrong or useless inferences that cannot (or should not) be used to aid our decisions. By combining these properties of imaginings, the author distinguishes between *Guiding Chosen* imaginings, *Guiding Unbidden*, *Misguiding Chosen* and *Misguiding Unbidden*, as shown in Table 2.4. The author focuses mainly on *Guiding Chosen* imaginings, which are also the ones that concerns our work the most. As we have already mentioned before, we are interested in those acts of imagination that are initiated by a voluntary action of the agent, who decides to imagine such and such: in this sense, the kind of acts of imagination we are interested in are chosen. As we also want to focus on those acts of imagination that are belief-like (meaning that their default unfolding follows what would be the case, were they real), our interest lies in imagining that are also guiding. Langland-Hassan focuses on this kind of imaginings for similar reasons: he wants to explore how imagination works, when initiated by a voluntary action and developed by reality-oriented rules and facts.

	Voluntary	*Involuntary*
Useful inferences	Guiding Chosen	Guiding Unbidden
Useless inferences	Misguiding Chosen	Misguiding Unbidden

Table 2.4: Langland-Hassan's distinction on types of imaginings.

Before going into detail with *Guiding Chosen* (GC) imaginings, the author establishes an analogy with respect to bodily actions in comparison: as he states ([35], page 62), "if a bodily action derives from an explicit intention to carry out that action, it clearly counts as chosen". However, there are many actions that, although not in an explicit or conscious way, are also chosen by the subject; say, when someone decides to go get a glass of water, the actions involved in going to the kitchen, opening the cupboard, reaching for a glass, etc. are not explicitly chosen one by one (say, as unitary actions), but are part of the more "general", explicitly chosen action to go get a glass of water. These sort of action-initiating mental states are called *intentions in action*, and they are distinguished from *prospective*, or "prior" intentions (such as the intention to have pizza for dinner, or to go to Indonesia on summer vacation), in that the former sort accounts for a particular movement being voluntary and willful, whereas the latter refers to a form of plan that does not involve immediate bodily actions.

By drawing on this analogy, Langland-Hassan claims that the kind of imaginings that can contribute in advancing one's interests (or, similarly, improving one's epistemic states) are those that are chosen, as it happens with bodily actions. Now, in order to understand GC imaginings, there are three general features of their architecture that need to be accounted for: the initial involvement of "top-down" intentions for initiating imaginings, the use of "lateral" constrains in the development of an imagining, and the cyclical involvement of top-down intentions throughout the course of an imagining. This assumptions render a picture of GC imaginings as a kind of continuously guided conditional reasoning.

Langland-Hassan identifies paradigmatic episodes of imagination as a sequence of mental states $i_1 \cdots i_n$, called an "imaginative episode", and in which each i_x is called an "imaginative state". An imaginative episode is initiated by an intention, which is to be regarded as an intention in action; in those cases, the imaginative episode counts as an *action* (a *mental action*, specifically), whereas imaginative episodes that are not chosen do not count as mental actions, on the grounds that they are not decided, similarly to the previous analogy with bodily actions. With respect to the "content" of the imaginings, Langland-Hassan distinguishes between *propositional* imaginings and *sensory* imaginings; the basic difference is that the former kind of imagining does not require the use of mental imagery[9], whereas the latter does. In the remainder of this section, we will only focus on Langland-Hassan's position regarding propositional imagination, as it is the kind of imagination we are interested in this work.

When considering up to which point imaginings can be seen as chosen, Langland-Hassan claims that, when setting the initial content of an imagining, top-down intentions (that is: voluntary, conscious actions of the agent regarding *what* to imagine) are used. In his work, nevertheless, the author focuses mostly on how imaginative episodes "advance" from one state to another (in other words, how a certain imaginative state leads to another specific imaginative state), rather than analyzing how an imaginative state is filled up with details not initially specified by the top-down intention. So, although the

[9] It is worth noting that Langland-Hassan supports this claim by considering the case of pretense: as the author claims, if a person can pretend that she is a tiger, or a mobster, or a snowflake, without using sensory imagery at all, then this seems to be a reason supporting that propositional imagination is not imagistic.

voluntary nature of the agent's intentions when initiating an imaginative episode is clearly present in his theory, the author's worries concern whether (and up to which point) the "advancement" of an imaginative episode is based on further top-down intentions. As he claims, if *only* top-down intentions were used in the advancement of the episode, then learning via imaginings would be pretty useless, as their advancement would already be determined by our will to make the episode evolve in such a way.

In order to overcome the issue of using only intentions in an imaginative episode, Langland-Hassan calls for a set of *lateral constrains* which, after the conditions of the initial imagining have been set, would encode the set of norms, logic or algorithm governing how the imaginative episode unfolds. Langland-Hassan sees each of the imaginative states i_x in an imaginative episode $i_1 \cdots i_n$ as a result of an inference step, given what is the case in the previous scenario i_{x-1}, and following certain rules determining how i_{x-1} leads to i_x. During his analysis of how imaginative episodes move forward, Langland-Hassan draws a comparison between his work and Nichols and Stich's proposal. In particular, and while revisiting Nichols and Stich's theory ([35], page 68), the author points out to Nichols and Stich's *inference mechanisms* as the responsible of determining, through inference rules, how a state i_{x+1} following a certain i_x would be. Therefore, Langland-Hassan's lateral constrains, when compared with Nichols and Stich's theory, seem to be more closely related to what Nichols and Stich identify as the facts and the default scripts that are used to determine how an imaginary scenario can be elaborated, rather than to the UpDater's role on consistency checking, belief-import and such. In particular, Langland-Hassan's account describes the unfolding of an imaginative scenario as a sort of step-by-step process guided by implications; for example, given an imaginative state i_x with content p and a set of lateral constrains containing a norm of the form $p \to q$, the next imaginative state i_{x+1} would have content q.

Langland-Hassan refers to the "deviance" objection when taking into account how, even in cases of GC imagining, the way an imaginative episode advances usually deviates from anything we would likely expect to happen, given the initial conditions. In this case, and still comparing his own proposal to Nichols and Stich's work, the author relates the deviance objection to both Nichols and Stich's default scripts (which are part of the UpDater mechanisms) and their Script Elaborator. Briefly recalling what we introduced in Section 2.3.1, default scripts encode different ways a situation could typically evolve, but without forcing any particular one: there are various possible outcomes, and it is up to the agent to choose which one she wants to follow. When things evolve in an "atypical" way, then the Script Elaborator is the key: beyond details that cannot be inferred from the pretense premise, the pretender's beliefs, and what happened earlier in the pretense, the Script Elaborator is responsible for coming up with unexpected, atypical plot twists. Langland-Hassan summarizes his view on the Script Elaborator as follows (see [35] page 73; the author's emphasis): "In essence, the job of the Script Elaborator is to account for whatever stages there may be in the sequence $i_1 \cdots i_n$ that would *not* have been inferred if i_1 were believed". The author sees in the Script Elaborator the key to understanding how much alike inferences in imagination and beliefs are; for, as he claims, if the Script Elaborator interferes with those belief-like inferential patterns, then it might be that inferences in the reign of imagination are not that much belief-like as some authors claim them to be. Therefore, until the Script Elaborator has been properly defined, and thus until it has been understood up to which point it interferes with belief-like inferences, it cannot be said whether (or how much) imaginings are belief-like.

Digging a bit deeper into the role of the Script Elaborator, and following from Langland-Hassan's "deviance" objection, the author suggests that, when our imaginings deviate from the "usual" patterns proposed by the lateral constrains, it is because the person intentionally intervenes in the imagining by stopping the undergoing process of the lateral constrains and inserting a new premise. This, in turn, involves a cyclical process that resets the whole cycle of initiating a new imaginative state (although already populated by whatever happened earlier in the imaginative episode) and calling again for the lateral constrains. Why, then, would the person entertaining the imagining intentionally intervene? Simply because she *wishes* to; she wishes for something unexpected to happen, for instance in an improvisation episode involving theater actors. According to Langland-Hassan, then, desires come at play into the picture. If this is how these "off-script" interventions take place, then the role Nichols and Stich attribute to the Script Elaborator can be reduced to a combination of the person's desires to make the imaginative episode evolve in such-and-such way, and a cyclical call to the whole process of GC imaginings. Following these considerations, Langland-Hassan claims how the cyclical processing of top-down intentions and the lateral constrains are enough to accomplish the work Nichols and Stich set out for the Script Elaborator.

The whole process of GC imaginings can be summarized by the schema in Figure 2.6. Note how the schema represents the fact that, whenever the scenario deviates in a "non-scripted" way through desire-driven premises the agent decides to add in the imagining, this deviation is identified as not being "typical" development: in particular, this corresponds to what Langland-Hassan identifies as the cyclical involvement of top-down intentions. In the schema, we represent this by the fact that the state at the upper-right corner of the figure is identified as i_x and, through applying the lateral constrains, it unfolds a next i_{x+1}; however, and in order to stress out that a new top-down intention with a new premise will alter the imagining in atypical ways, the state at the lower-right corner of the figure is identified by i_y, instead of i_{x+n}. This represents that voluntarily adding a new premise to the imagining would not lead, as Langland-Hassan points out, to a scenario typically following from the previous one, but rather alter it in ways that deviate from the usual patterns captured by the lateral constrains. Once the new imagining has been set, though, lateral constrains will indeed compute the next i_{y+1} by following the norms, rules and algorithms that describe how the imagining could typically unfold.

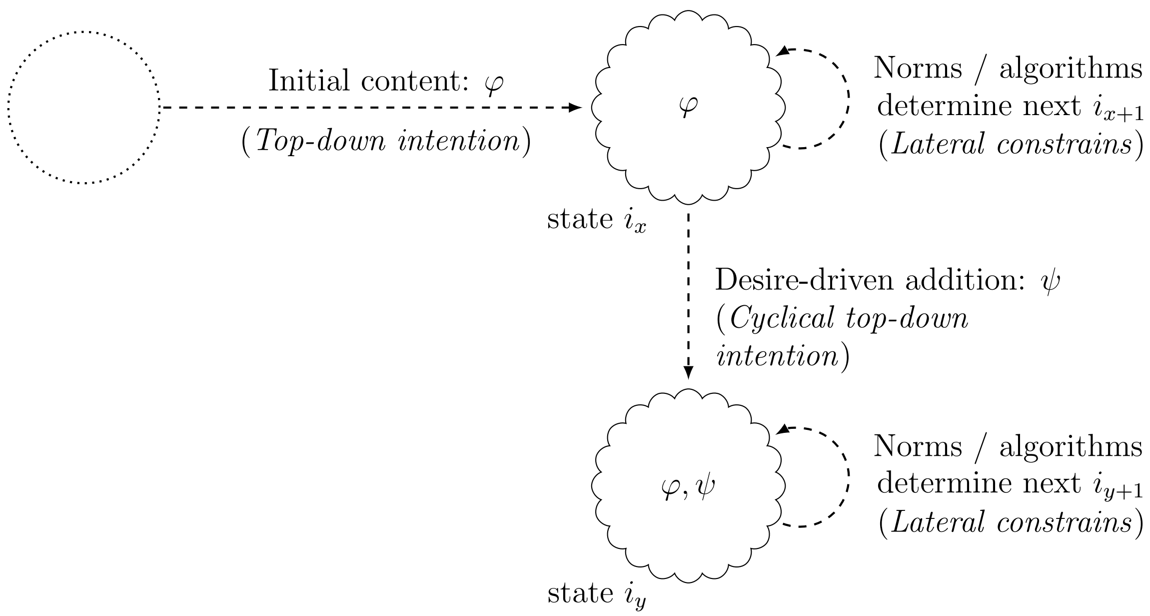


Figure 2.6: An act of imagination according to Langland-Hassan's theory.

When comparing the structure drawn by the mechanisms identified by Langland-Hassan's theory, we can see how they coincide in great measure with those identified by Nichols and Stich, as shown in Figure 2.4. Although it is true that the authors define their respective mechanisms differently in some ways, they nevertheless represent a similar underlying composition of mechanisms.

2.4 Distilling the Theories

We have reviewed the influential cognitive theory of pretense, from Nichols and Stich, which provides a detailed analysis of the mechanisms involved in the creation and development of imaginary worlds; furthermore, we have also reviewed two recent theories of voluntary imagination that coincide, in their procedural analysis, with the cognitive theory of pretense.

Even though each of these theories has certain differences, the analysis they do of voluntary acts of imagination have a lot in common. In particular, the underlying mechanics involved in any act of imagination is somehow captured by every theory. Therefore, we will not focus on just one of them, but rather we will distill the process of imagining as a whole from all the theories, and we will aim to capture it by using a formal system in Chapter 4.

As we have already said, our interest, as well as the interest of the previous theories, lies on *voluntary* acts of imagination. We do not want to represent an agent who is suddenly "surprised" by an imagining popping out into her mind, but rather an agent who voluntarily decides to entertain an imaginary scenario. Every act of imagination characterized as such must start, according to each one of the previous theories we reviewed, with a *voluntary action* that creates or initiates the imaginary scenario. This action is characterized by carrying with it a certain argument, namely a set of initial premises or initial conditions, that characterize the imaginary scenario that is to be created. Therefore, this initial action is the one responsible of creating, where there was none, a brand-new imaginary world characterized by its initial premises. Following Nichols and Stich's terminology (see Section 2.3.1), those initial premises are "clamped" in the imaginary world: they have preference over everything else, as they are precisely what will make the imaginary world be the way it will be. This action corresponds, precisely, to Nichols and Stich's action of putting a set of premises into the PWB, to Williamson's voluntary imagination, and to Langland-Hassan's initial top-down intentions.

Once the imaginary scenario has been created, and the premises characterizing it have been "clamped", the imaginary scenario should be elaborated. So, aside from what the agent has "forced" to be the case in the scenario, what else would likely be the case in it? The source of such development are, again according to all three theories, the agent's beliefs. The underlying idea is that an imaginary world will, at least initially, be expected to look like and behave in the same way it would, if it was real. Therefore, the rules that determine what else would be the case in the imaginary scenario are the same rules that the agent would expect to hold if the scenario was real. We refer to this process either as *reality-oriented development* (following Williamson), or *belief-like development* (following the ideas behind Nichols and Stich's approach). One way or another, we refer to the fact that imaginings will develop by mirroring what the agent believes would happen in that scenario if it was real. So, after having defined the initial imaginary scenario, there is some kind of "belief-import" process aimed to make the imaginary scenario as reality-like as it can be. This mirrors Nichols and Stich's UpDater mechanism, Williamson's involuntary imagination, and Langland-Hassan's lateral constrains.

The imaginary scenario has been created and characterized by a set of initial premises, and it has been filled up by using the agent's beliefs about the real world. What is next? More often than not, games of make-believe or imaginative episodes develop in highly unexpected ways: when listening to a story-teller, when seeing or acting in a theater play, or simple when pretending to be in a tea-party with our little niece, things turn out in ways radically different that they would, if they were real. Why does this happen, and where do these plot-twists come from? The agent, by following a certain desire (according to Langland-Hassan), wishes for something unexpected to happen, and thus she decides to stop mirroring what would happen in that situation if it was real, and starts feeding new premises into the imaginary world. These premises are fed in a similar way as they are in the initial voluntary action: they come from the agent's will to do so, and they are clamped into the imagining. Only two of the three previous theories account for this action: as Williamson is only interested in reality-oriented imagination, his work does not account for this phenomenon. Nevertheless, both Nichols and Stich's theory, through the Script Elaborator, and Langland-Hassan's theory, through the cyclical involvement of top-down intentions, do take this into account. Table 2.5 summarizes the relation between the identified components and each of the previously reviewed theories.

	Nichols / Stich	*Williamson*	*Langland-Hassan*
Voluntary initiation	Premises into PWB	Volunt. imag.	Top-down intention
Real.-orient. develop.	UpDater	Involunt. imag.	Lateral constrains
Atypical develop.	Script Elaborator	-	Cyclical top-down int.

Table 2.5: Distilling a general algorithm for imagination acts.

After distilling the three reviewed theories, we can identify three main components within the action of performing an imagination act: a voluntary action of the agent that creates a new imaginary scenario according to a set of initial premises, a belief-importing process that develops the imaginary scenario by using the agent's beliefs about the real world, and a voluntary action of the agent to add something new to the imaginary scenario that does not follow from the belief-importing process.

After setting the theoretical grounds of our work, we now move on to a more technical part. In Chapter 3, we briefly introduce the basics of formal logics, review some existing logic systems aimed to represent a logic for imagination, and point to some important caveats they have, with respect to our goals. Then, in Chapter 4, we define a logic aimed to capture and represent an agent performing an act of imagination, as we have identified such process in this last subsection. After doing so, we prove its soundness and completeness, and discuss its properties to see how good it fulfills our desiderata. After reviewing how our formal proposal behaves, and as we explain in the Brief Interlude (Chapter 5), the analysis made of imagination acts by the previous theories is not fine-grained enough to define a logic that captures imagination acts in the way we want to. In other words: the theoretical background we now take as the basis for our work is not enough, when moving to more formal grounds. Therefore, after discussing the issues we recognize in the system proposed in Chapter 4, we move to the second part of this work and take on a new approach to the initial problem of analyzing acts of imagination and representing them in a formal system.

Chapter 3

Formal Logic and Imagination

In Chapter 2 we have reviewed three relevant theories that philosophers and cognitive scientists have developed regarding imagination. In those works, their authors analyze and discuss how imagination (and, in particular, the processes involved in carrying out an act of imagination) works. Now, we move to more formal grounds. We begin this chapter by providing a brief introduction to formal logic; then, we review different logics for imagination. After that, we point out an important feature we think they all lack: they can represent an agent who is *already* imagining something, but they cannot capture how the act of imagination, in itself, takes place; moreover, they throw no light over the processes involved in creating new imaginary worlds.

3.1 One Further Stop: Some Formal Background

Before moving on to review the already existing logics of imagination, we briefly introduce the main formal details of some logical systems; specifically, we begin by introducing propositional and first-order logic in order to show some caveats that required a new way of handling formal languages; then, we follow this path and introduce modal, hybrid and epistemic logic.

We provide this brief introduction as an aid for those readers that are not acquainted with those systems. However, the formal complexity of chapters 4 and 7 assume that the reader is already familiar with modal logic and its semantics. Otherwise, we refer to the citations provided in the following pages, as providing a complete and comprehensive introduction to such logics would require almost a whole book.

3.1.1 Propositional and First-Order Logic

Propositional logic allows us to translate natural language sentences into a formal language in order to analyze their structure or to use certain formal methods to determine the validity of an argument composes by various formalized sentences[1]. For instance, consider the following sentence:

> "If I go to the desert and I don't have a map, I get lost"

If we want to capture the underlying structure of the sentence by using a formal language (instead of our powerful but rather vague natural language), we can do so in propositional logic in the following formula (where d stands for "I go to the desert", m for "I have a map", and l for "I get lost"):

$$d \wedge \neg m \rightarrow l \tag{3.1}$$

In the previous formula, symbol $\wedge$ stands for "and" (meaning conjunction), symbol $\neg$ stands for "not" (meaning negation), and symbol $\rightarrow$ stands for "if $\cdots$ then $\cdots$" (meaning implication or consequence). By using the language of propositional logic, we have been able to code the information conveyed by the original sentence, expressed in natural language, into a formal expression.

However, we may want to convey this piece of advice not only to ourselves, but also to everyone else facing the same situation: after all, navigating through a desert on your own is not an easy matter for anyone. Thus, we may be interested in formalizing sentences concerning various individuals, and say that these individuals fulfill a certain property, or are affected by some state (say, getting lost). Now, in order to formalize properties or quantify over a domain of elements, propositional logic is not expressive enough. Consider a more generalized version of the previous sentence:

> "Anyone who goes to the desert and does not have a map, gets lost"

This time, if we want to capture the meaning of this sentence in a formal language, we will need to use first-order logic[2] as follows (where $D(x)$ stands for "x goes to the desert", $M(y)$ for "y is a map", $H(x, y)$ for "x has y", and $L(x)$ for "x gets lost"):

$$\forall x(D(x) \wedge \neg\exists y(M(y) \wedge H(x, y)) \rightarrow L(x)) \tag{3.2}$$

[1] For a thorough introduction to propositional logic, see [31].
[2] Again, we refer to [31] as well for an introduction to first-order logic.

In this example, the symbol $\forall x$ stands for "for all elements x", and $\exists y$ stands for "there exists an element y" (where x and y are names given to variables); the other formal symbols are as in the first example.

However, consider now that what we want to say is not that every desert travelers who does not have a map will get *inevitably* lost, but only that she *may* get lost. Therefore, what we want to formalize an advice to desert travelers that points to something that can *possibly* happen, like in the following sentence (which we express again in first person):

"If I go to the desert and I don't have a map, I *may* get lost"

If we try to formalize this using propositional logic, we will end up having the same expression as in Formula 3.1, and so we will be saying that, if I go to the desert without a map, I will get *inevitably* lost, no matter what. Similarly, if we try to formalize the corresponding quantified sentence by using first-order logic, we again end up with Formula 3.2, which also collapses the modal expression of our conclusion (i.e.; "I may get lost"), into something that cannot be prevented (i.e.; "I get lost"). How then, could we capture this shade in the previous sentence, without losing its full meaning? And, more specifically, could we do so by using propositional or first-order logic? The answer is no[3]. In order to grasp the full meaning of modal expressions, we need to use modal logic.

3.1.2 Modal Logic

Modal logic was defined in order to amend some caveats propositional and first-order logics were facing when trying to capture the meaning of *modal expressions* in our everyday language. Following the definition given by [24], "a modal is an expression (like 'necessarily' or 'possibly') that is used to qualify the truth of a judgement.". For instance, consider the following sentence, which we already introduced in the previous section:

"If I go to the desert and I don't have a map, I may get lost"

As we have argued, the formal languages we presented in the previous section cannot fully account for what we want to say, and they can just account for rather simplified versions of this sentence. Specifically, what we want to express is that, when going to the desert without having a map, it is *possible* to get lost: that is, getting lost may follow from going to the desert without a map, or it may not. Thus, when using modalities in

[3] There actually exists a strong relation between modal and first-order logic, and it is possible to translate certain fragments of both languages; for more on this topic, see [7].

our everyday language (such as *possible*, *necessary*, *eventually*, *previously*, etc.) we are not conveying information about the actual state of affairs, but rather about a different state of affairs that is somehow related to the current one.

In order to capture the modal meaning of someone *possibly* getting lost in the desert, we can use the language of propositional modal logic[4] as follows (we use the same *atomic variables*, represented by lowercase letters, as in the first example):

$$d \wedge \neg m \rightarrow \Diamond l$$

In this example, symbol $\Diamond$ is to be read as "it is possible that"; the other symbols are likewise to the other examples.

Similarly, we can use modal logic to formalize something not being just possible, but also *necessary*; we use the symbol $\Box$, standing for "necessarily", and expressing that no state of affairs resulting from the current one could prevent what it is necessary to happen. Take, as an example, the following sentence:

"When one introduces modal logic, it is necessary to talk about modalities"

The previous sentence is represented in the language of modal logic as follows (where l stands for "one introduces modal logic" and m stands for "one talks about modalities"):

$$l \rightarrow \Box m$$

After this brief introduction to the language of modal logic, let's get into detail. We devote the rest of this subsection to introducing how modal logic works (specially regarding its semantics); for a more extensive introduction, we refer again to [6].

First of all, let's define the language of modal logic. We start with a set ATOM of *atomic formulas* (represented by lowercase letters $p, q, r...$, and standing for facts we want to speak about), a set of propositional connectives $\neg, \wedge, \vee, \rightarrow$ (standing, respectively, for "negation", "conjunction", "disjunction" and "material implication") and a modal operator $\Diamond$ (standing for "it is possible that"). We call FORM the set of well-formed formulas of the language of modal logic, and we define them as follows:

$$p \mid \neg\varphi \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid \varphi \rightarrow \psi \mid \Diamond\varphi$$

[4]See [6] for an extensive introduction to modal logic.

where $p \in$ ATOM and $\{\varphi, \psi\} \in$ FORM. We define a derived operator $\Box$ (standing for "it is necessary that") as a dual of $\Diamond$ in the following way:

$$\Box\varphi \equiv \neg\Diamond\neg\varphi$$

So, intuitively (and before presenting the formal definition of the satisfiability for $\Diamond$), φ is necessary if and only if it is *not* possible for $\neg\varphi$ to be the case.

In order to give an interpretation of the formulas of the modal logic language (and thus be able to assert whether they are true or not), we need to use the so-called *Kripke models* or *relational models* (see again [6] for a more detailed account of different kinds of relational models). We define a model $\mathcal{M}$ as a structure $\mathcal{M} = \langle W, R, V\rangle$ formed by the following elements:

- W is a set of states called "possible-worlds" or "states of affairs". We use the lowercase letters $w, v, u, \cdots$ to refer to elements of W.

- $R \subseteq W \times W$ is a binary relation over elements of W called the "accessibility relation". Intuitively, this relation establishes which possible worlds can be accessed from other possible worlds, and which not. We use pairs of the form $(w, v), (v, u), \cdots$ to refer to elements of R.

- $V :$ ATOM $\to \mathcal{P}(W)$ is a function from atomic formulas of the language to elements of W, called the "valuation function". Intuitively, it sets which atomic formulas are true at each possible world.

We evaluate the formulas of the language in a world w of a model $\mathcal{M}$, and we use $\mathcal{M}, w \vDash \varphi$ express that formula φ is true at world w within model $\mathcal{M}$; conversely, we use $\mathcal{M}, w \nvDash \varphi$ to express that formula φ is *not* true at world w within model $\mathcal{M}$. We recursively define the semantics for each kind of well-formed formula as follows (we include the semantics of both $\Diamond$ and $\Box$ for the sake of clarity, but defining the semantics for one would be enough, as one can be derived from the other; we use the expression "iff" as an abbreviation for "if and only if"):

$$
\begin{array}{lll}
\mathcal{M}, w \vDash p & \text{iff} & w \in V(p) \\
\mathcal{M}, w \vDash \neg\varphi & \text{iff} & \mathcal{M}, w \nvDash \varphi \\
\mathcal{M}, w \vDash \varphi \wedge \psi & \text{iff} & \mathcal{M}, w \vDash \varphi \text{ and } \mathcal{M}, w \vDash \psi \\
\mathcal{M}, w \vDash \varphi \vee \psi & \text{iff} & \mathcal{M}, w \vDash \varphi \text{ or } \mathcal{M}, w \vDash \psi \\
\mathcal{M}, w \vDash \varphi \to \psi & \text{iff} & \mathcal{M}, w \vDash \neg\varphi \text{ or } \mathcal{M}, w \vDash \psi \\
\mathcal{M}, w \vDash \Diamond\varphi & \text{iff} & \text{there exists a world } v \in W \text{ such that } (w, v) \in R \\
 & & \text{and } \mathcal{M}, v \vDash \varphi \\
\mathcal{M}, w \vDash \Box\psi & \text{iff} & \text{for every world } v \in W \text{ such that } (w, v) \in R \text{, it is} \\
 & & \text{the case that } \mathcal{M}, v \vDash \varphi
\end{array}
$$

Note how the truth-value of the propositional formulas (that is, those formulas not involving operators $\Diamond$ or $\Box$) only concern the world where the formulas are evaluated (namely, w), whereas the modal formulas (those involving either $\Diamond$ or $\Box$) need to account for worlds that are distinct from w but which are, nevertheless, *accessible* (that is, related via the accessibility relation) from the world of evaluation w: this is, precisely, what characterizes the modal character of this kind of logics.

Figure 3.1 represents a simple Kripke model encoding a few basic pieces of advice for desert travelers (in which we use the same propositional variables we used in the previous examples, plus a new o, standing for "I arrive to an oasis"). For the sake of clarity, we also include the relevant negated formulas at each world in order to make the examples following the figure more understandable.

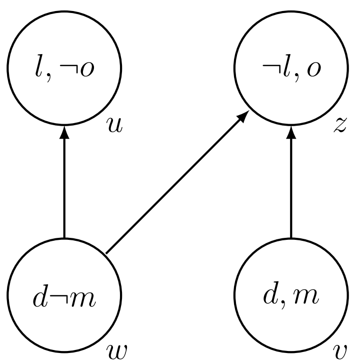


Figure 3.1: A Kripke model for propositional modal logic.

Now, just to put the logic in motion and show how the language is interpreted with respect to the model, let's check whether certain sentences (in natural language first, and then translated to our logic) are true in certain possible worlds (we use the symbol $\Rightarrow$ as a metalanguage symbol to express that we translate the natural language sentence to our logic and we evaluate it in the model: note how the world of evaluation is necessary to establish whether the sentence is true or not, and how it changes in some of our evaluations); note how the world at which we evaluate each formula, represented by the letters $\{w, v, u, z\}$, is crucial then determining whether the formula is true or not; if we were to evaluate each of these formulas at different worlds, their truth-value would probably change from one to another:

"If I go to the desert and I don't have a map, I may get lost"	$\Rightarrow$	$\mathcal{M}, w \vDash d \wedge \neg m \rightarrow \Diamond l$
"If I go to the desert and I have a map, it is impossible that I get lost"	$\Rightarrow$	$\mathcal{M}, v \vDash d \wedge m \rightarrow \neg\Diamond l$
"If I go to the desert and I have a map, I'll surely get to an oasis"	$\Rightarrow$	$\mathcal{M}, v \vDash d \wedge m \rightarrow \Box o$
"If I get lost, it is impossible that I can get to an oasis"	$\Rightarrow$	$\mathcal{M}, u \vDash l \rightarrow \neg\Diamond o$
"If I get lost, it is not possible to get to an oasis"	$\Rightarrow$	$\mathcal{M}, u \vDash l \rightarrow \neg\Diamond o$

3.1.3 Hybrid Logic

Although being a rather expressive language, modal logic still has some limitations. In order to highlight them, let's take an example based on temporal logic, in which possible worlds in the model represent different days, and therefore the accessibility relation between them expresses how the state of affairs could change, as days go by. By giving a temporal reading to modal logic, one can express things like "something could be the case at some future time", or "something has always been the case on the past", or even "something is, will always be, and has always been the case".

However, let's now try to represent, by using a temporal structure, the way time behaves in the movie *The Groundhog Day*[5]: briefly recalling, the movie is about a journalist who finds himself suddenly stuck in a kind of "temporal loop" in which, each morning, he

[5]The reader who has not seen the movie yet may want to skip the remaining of the section, as it may spoil some important details on the movie's plot.

wakes up in the same day he just was (this is just the phenomenon we want to represent right now, so we do not focus on how the movie goes any further, nor on how does it end). We could represent this in a temporal model by having a possible world identifying "today" (which we call w and which is the only possible world we have in this model; that is, $W = \{w\}$) and a single accessibility relation pointing to itself (that is, $R = \{(w, w)\}$): in this very simple temporal model, moving forward in time from day w takes one back to the same day, namely w, as it can be seen in Figure 3.2.

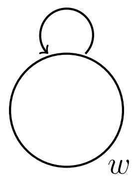


Figure 3.2: A temporal model for *The Groundhog Day* scenario.

Now, in the temporal logic language, we substitute the modal operators we had in standard modal logic by different symbols: we use the symbol F to represent "at *some* time in the future" (roughly corresponding to the $\Diamond$ in modal logic), and the symbol G to represent "at *every* time in the future" (roughly corresponding to the $\Box$ in modal logic)[6]. Now, we want to say something like "at some time in the future it will be today" (recalling that we identify "today" with world w); in fact, we want to say something even stronger: "at every time in the future it will be today". Well, the question is: could we even do that, using the language we have introduced so far?

If we stick to standard modal (or temporal) logic, then the answer is "no": in these languages, we can formalize facts and complex formulas that "take place" (or "are true at") certain possible worlds, but, besides that, we cannot talk *about* the possible worlds themselves. We can say what happens *in* them, but not what happens *about* them. Nevertheless, there is a new layer we can add to modal logic that greatly increases its expressive power, allowing us to talk about those possible worlds that are part of the model: hybrid logic (see [5] for a thorough introduction to this kind of logic).

Hybrid logic adds a set of *nominals* NOM to the modal logic language, and a function $N : \text{NOM} \to \mathcal{P}(W)$ to the model which assigns, to each nominal $i \in \text{NOM}$, a single possible world $w \in W$; in other words, $N(i)$ is a singleton, for any $i \in \text{NOM}$. Intuitively, nominals play the role of unique names identifying possible worlds. However, nominals

[6]As we are only concerned with the example, we do not include any operator regarding the past: again, we refer to [6], or Chapter 10 in [26] for a comprehensive introduction to temporal logic.

are also treated as formulas within the language, but with one particularity: they are *only* true when evaluated at the possible world they identify. Besides, hybrid logic introduces an operator @ according to which a formula $@_i\varphi$ (note that the operator requires using a nominal as a subscript) is interpreted as "at the world identified by i, it is the case that φ". The formal definition of the semantics of the operator, therefore, is as follows (given a model $\mathcal{M}$ and a possible world w):

$$\mathcal{M}, w \vDash @_i\varphi \quad \text{iff} \quad \text{there exists a world } v \in W \text{ such that } N(i) = v \text{ and } \mathcal{M}, v \vDash \varphi$$

Now, recalling that nominals are also formulas of the language, here is how, by using hybrid logic, we can formalize the temporal anomaly depicted in the movie *The Groundhog Day*, as we wanted to (we take i is the nominal identifying world w):

$$\begin{array}{lcl} \text{"At some time in the future it will be today"} & \Rightarrow & @_iFi \\ \text{"At every time in the future it will be today"} & \Rightarrow & @_iGi \end{array}$$

3.1.4 Epistemic Logic

As we have said in Section 3.1.2, Kripke models and modal logics have been used to formalize a wide variety of modalities, including the concepts of *knowledge* and *belief*. In [28], Jakko Hintikka proposes a modal language and a semantics to formalize these concepts, and discusses the properties they should fulfill. As we are interested in representing imagination (which is, as discussed in Chapter 2, a mental attitude like believing or knowing), it seems natural to take, as a starting point, a system already capable of handling other relevant mental attitudes. In the following paragraphs, we introduce the most basic system: the single-agent epistemic logic (see [50] for a more detailed account of different logic systems dealing with knowledge and belief).

The logic we introduce here allows to represent an agent who may know how certain things are, and may at the same time ignore other things. Thus, this logic allows us to formalize sentences such as "the agent knows that φ is the case", or "the agent does not know φ to be the case"; furthermore, it can express second-order knowledge, like "the agent knows that she does not know φ", for instance. The language of epistemic logic is almost the same propositional language already introduced in Section 3.1.2, but, in this case, the modal operator is usually referred to as K, where $K\varphi$ is interpreted as "the agent knows φ". The definition of its semantics, roughly corresponding to the $\Box$ operator in standard modal logic, is:

$$\mathcal{M}, w \vDash K\varphi \quad \text{iff} \quad \text{for every world } v \in W \text{ such that } (w, v) \in R, \text{ it is the case that } \mathcal{M}, v \vDash \varphi$$

Similarly, we can define a new operator M, which can be derived from operator K as $M\varphi \equiv \neg K \neg \varphi$, and which is often read as "the agent considers φ to be possible" (in the sense of it being plausible according to her knowledge, rather than the sense of possibility we discussed in Section 3.1.2). The semantics for this derived operator are similar to the semantics for the $\Diamond$ operator in modal logic:

$$\mathcal{M}, w \vDash M\varphi \quad \text{iff} \quad \text{there exists a world } v \in W \text{ such that } (w, v) \in R, \text{ and it is the case that } \mathcal{M}, v \vDash \varphi$$

Regarding the semantics, epistemic models are as standard Kripke models of modal logic, but with one particularity: the accessibility relation, which now aims to represent whether the agent can distinguish, according to what she knows, different possible states of affairs, has certain restrictions imposed. Namely, the accessibility relation (which we will now refer to as "indistinguishability relation", and call R_K) is required to be *reflexive* (for every w, it is the case that w is related to w), *transitive* (if a state w is related to a state v, and v is related to u, then w is also related to u) and *symmetric* (if w is related to v, then v is also related to w). Binary relations that fulfill such properties are known as *equivalence relations*. These restrictions aim to capture some of the properties that had been attributed to "knowledge", as it was defined and analyzed in the literature (see again [28]). In particular, each one of these properties can be expressed using formulas of the epistemic logic (recall that we defined $M\varphi \equiv \neg K \neg \varphi$); whenever a model for epistemic logic fulfills any of these properties, the formula that captures is becomes *valid* in the model (i.e.; it is true at *every* possible world within the model, no matter what):

$$\begin{array}{lcl} \text{Reflexivity} & \Rightarrow & K\varphi \to \varphi \\ \text{Transitivity} & \Rightarrow & K\varphi \to KK\varphi \\ \text{Symmetry} & \Rightarrow & \varphi \to KM\varphi \end{array}$$

Although this topic falls outside the scope of our work, it is worth mentioning that attributing such strong properties to knowledge leads to a rather idealized notion of it. Specifically, forcing the indistinguishability relation to be an equivalence relation leads to an agent who is logically omniscient: namely, an agent who automatically knows every logical consequence of what she knows and every logical tautology, for instance. We will not dig deeper into this topics, but there is a vast amount of literature on the subject, and a lot of logical systems aimed to avoid logically omniscient agents (see, for instance: [51], [32], or [2]).

Regardless of the problem of logical omniscience, the interpretation of the "knowing" attitude within an epistemic model is different from the way we interpreted "possible" and "necessary" in a standard modal model. In epistemic logic, each possible world represents a different state of affairs that could be the case; the relation R_K is used to connect those states of affairs that, according to what the agent knows, she cannot tell apart. Therefore, an epistemic model relating two states w and v, and wherein state w it is the case that φ, whereas in state v it is the case that $\neg\varphi$, represent that the agent does *not* know neither φ, nor $\neg\varphi$. As far as she knows, the *actual* state of affairs could be any of those two. According to this, the smaller the model (in terms of number of worlds related by the indistinguishability relation), the more knowledge the agent has; particularly, an epistemic model with a single possible world would represent absolute knowledge of the current state of affairs. In order to further clarify this notion, take the epistemic model in Figure 3.3 as an example.

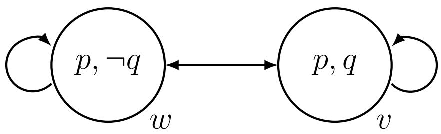


Figure 3.3: A Kripke model for epistemic logic.

This model represents an agent who knows p (as p is true at every possible world accessible by the agent), but does not know q, nor $\neg q$ (as there exists an accessible world in which q, and also another one in which $\neg q$).

Our interest in epistemic logic stays in its single-agent setting, as we are only interested in representing how a single agent performs acts of imagination. However, epistemic logic is normally defined in a multi-agent settings (see Chapter 1 in [50] for more on this topic), which allows to study the dynamics of information exchange, knowledge and beliefs. This topic concerns how, either through acts of communication or by internal reasoning steps, an agent or a set of agents can come to know (or believe) something new, and how this process of updating one's knowledge and beliefs should be accounted for in formal systems. For a comprehensive introduction to some of the most relevant works on these topics, we refer again to [50]; for a more specific work on the dynamics of belief revision, see [1]; for formal systems representing internal reasoning steps of an agent, see [51].

At this point, and having already introduced the appropriate background, we present some of the most relevant works that aim to represent imagination via logical systems. After reviewing them, we discuss an important caveat we think they all have with respect to our interests, which is what, in turn, fuels our approach when defining a formal logic in Chapter 4.

3.2 Existing Logics of Imagination

Although formal works on other mental attitudes like believing or knowing have received a great deal of attention, imagination has received little. Still, there are some interesting work that we want to briefly present in here, while sketching some of their most important points. While reviewing these approaches, though, we will not focus on the technical details, but rather on the way imagination works in them. We are interested in seeing what these authors highlight about imagination and, as a result of this, what is their contribution in the understanding of the way imagination works. Then, we argue in Section 3.3 how these systems do not capture, precisely, what we aim to represent: the dynamics involved within an act of imagination.

3.2.1 Lewis: Counterfactuals

In [36], David Lewis provides an extensive study of how counterfactual reasoning works, and defines a logic capable of accounting for it.

At the very beginning of his work, Lewis introduces a logical operator $\Box\!\!\rightarrow$, in which $A \Box\!\!\rightarrow B$ stands for "if it were the case that A, then it would also be the case that B"; similarly as how modal operators are related in modal logic (see Section 3.1.2), he also defines a weaker notion of "counterfactual possibility", expressed by operator $\Diamond\!\!\rightarrow$, and in which $A \Diamond\!\!\rightarrow B$ stands for "if it were the case that A, then it might be the case that B". After briefly introducing this pair of operators, Lewis discusses and characterizes strict conditionals, counterfactual conditionals and the implications of requiring *ceteris paribus*[7] to hypothetical reasoning.

[7] The expression *ceteris paribus* comes from the Latin language: it literally means "other things equal", and it is usually translated to the English language as "all other things being equal". The expression is commonly used when analyzing relations between two states of affairs (either empirical, causal or logical) in order express the requirement or assumption that, aside from the relevant variables being changed,

Regarding the *ceteris paribus* clause, the author points out how problematic can be to try to stick too hard to it. By following the same example the author uses, let's consider a world that is exactly as our own world, but in which kangaroos have no tails. Following Lewis' argument, *if* everything else, besides kangaroos having no tails, have to stay exactly the same as it is in our world, then the trace kangaroos leave when walking on the sand should still be the same; however, how could it be the same, if they had no tails to leave the same kind of tracks they leave in our actual world? In that case, one should then assume that the way those specific tracks are produced is different from the way they currently are, but nonetheless result in the exact same kind of track. In short, the underlying conclusion is that looking too hard for a strict *ceteris paribus* clause may result in putting in motion a series of causal relations that end up changing more things than what one actually intended to.

In order to represent counterfactual reasoning, Lewis introduces a notion of "closeness" between worlds. Intuitively, the less things one that are different in an hypothetical world, the closer it will be the actual one. To account for that, the author proposes a system of spheres (which are, roughly speaking, sets and subsets of worlds) representing this notion of closeness: the closer (or more similar) a world w is to a world v, the closer the sphere containing w would be to the sphere containing v.

According to this, a counterfactual formula $\varphi \mathrel{\Box\!\!\rightarrow} \psi$ is considered to be true at a world w if and only if either φ is *not* true in any world v within certain spheres of similarity related to w, or if φ holds in at least one world within a sphere S close enough to the one containing w, and $\varphi \rightarrow \psi$ holds at every world contained in that sphere S. Note how the first case corresponds to a sort of vacuous case: either φ is not the case in any world contained within certain spheres of similarity, or it is true in some sphere which we are not considering (for being not close enough to the world w, in which we are evaluating the counterfactual formula); in this case, one may say that φ is *not entertainable* at world w (because it depicts a situation which is not close enough to the current state of affairs defined in w). The second case represents the main case for counterfactual reasoning: φ is true at some world within a sphere S which is close enough to the sphere containing w, and the implication $\varphi \rightarrow \psi$ holds at every world contained in that sphere S.

everything else in the state of affairs stays the same. Particularly on the topic we are discussing, *ceteris paribus* is used to require that, when engaging in hypothetical or counterfactual reasoning, everything else besides what it is intentionally being changed stays the same as it is in the base situation.

3.2.2 Niiniluoto: Imagination and Fiction

In [44], Ilkka Niiniluoto develops a systematic framework for studying the syntax and semantics of imagination, using possible worlds semantics and following Jaakko Hintikka's approach on propositional attitudes like knowledge or perception (as in [29] and [30]). Niiniluoto writes $I_a p$ to stand for the expression "a imagines that p" (where a is an agent and p a proposition), and takes it as the basic form of other expressions such as "a imagines b" (where both a and b are agents), or "a imagines b as an F" (where F is the name of a predicate).

As it can be seen, Niiniluoto is not only interest in the propositional use of imagination (namely, "imagine that"), but also in objectual (or sensory) and active imagination, which involve both imagining other entities (rather than simply states of affairs), or imagining entities performing certain actions (see Section 2.1 in Chapter 2 for more details on different types of imagination). It is worth noting that the way imagination is treated by the author often involves some sort of mental imagery (which, as argued in Chapter 2, is something we do not require in our work). Later on, the author identifies and discusses a set of syntactic properties that the newly defined operator I_a should (or should not) satisfy.

Once the syntactic grounds have been set, Niiniluoto's work goes more in-depth on an analysis of how people normally use imagination, and how this can be represented in a formal way. When we imagine something, the author argues, the depicted situation lacks many of their relevant features; for instance, and following the same example used by the author, when "I imagine Ingmar dancing with a blond girl", there are many things that are left unspecified: how is Ingmar dressed? Who is the blond girl? Are they dancing a waltz or a tango?

By following these lines, the author defends that imagined scenarios must not be *complete*, in the sense that every proposition must be specified (or, more technically, assigned a truth-value) in there: in order to account for this, the author uses what Hintikka calls *small worlds* (which are, precisely, worlds in which some propositions fail to have any true value assigned). Regarding what else is, nevertheless, specified in the imagined scenario, Niiniluoto follows a first-order approach: therefore, a sentence such as "I imagine of Anna Karenina that she has green eyes" would be formalized as $\exists x(x = \text{Anna Karenina} \wedge (x \text{ has green eyes}))$.

Regarding that, the author defends how imagination can indeed be directed towards real and fictional entities alike. As expected, though, the problem of intensionality appears as a result of his first-order approach to (specially fictional) entities: how could we know whether the Donald Duck you and me are imagining is the same? How could our acts of imagination be directed towards the same entity? And what properties do our Donald Ducks actually share? The author proposes a solution to this issue based on drawing a distinction between private and public fictional entities: although I can still imagine a new character I have just invented (with anyone else having “access” to this private, fictional character that only exists in my imagination), by making a fictional entity public (like in the case of Donald Duck), people can then imagine about that entity and direct their imagination towards it (even if the entity does not always share the same properties when imagined by different people). One of the most interesting uses of these public acts of imagination, states the author, would be to understand a work of fiction (such as a novel) as an announcement of its author publicly stating “let us imagine that T” (being T the set of all fictional facts, entities and actions happening in the novel).

3.2.3 Costa-Leite: Logical Properties of Imagination

Alexandre Costa-Leite takes in [21] the work of Niiniluoto as the starting point of his own and defines a logic that explores the interactions between “imagination”, “conception” and “possibility”.

Regarding the distinction between these concepts (and recalling what has been said in Section 2.1 of Chapter 2), Costa-Leite does require for “imagination” to be related to some sort of mental imagery, whereas “conception” is instead associated with comprehension of concepts, but without the need to involve imagery at all. Through this distinction, the author argues how imagining can be considered an act of conceiving, while not the other way around. Regarding “possibility”, Costa-Leite distinguishes between two different kinds: on the one hand, *empirical possibility* depends on a given context X (like a scientific area) that allows to define whether something is X-possible, depending on whether it contradicts the underlying empirical theory (be it biology, physics...); on the other hand, *logical possibility* concerns whether something is possible with respect to logical consistency (that is, whether something does not contradict a set of logical rules). The latter kind of possibility is the one Costa-Leite is interested in his work, and, specifically, the one he aims to combine with imagination and conception.

Once the distinctions between these three terms has been set, Costa-Leite draws on theories by Descartes and Hume to define how those notions interact with each other. In particular, Descartes argues that imagination implies conception, while not conversely;

Hume, instead, argues that both imagination and conception imply possibility.

When defining a formal language, Costa-Leite keeps Niiniluoto's notation of $I_a\varphi$ standing for "agent a imagines φ"; then, he similarly defines $C_a\varphi$ as "agent a conceives φ", and uses the usual alethic modality $\Diamond\varphi$ for "it is possible that φ". Contrary to Niiniluoto, Costa-Leite focuses only on propositional imagination. Therefore, and in order to collapse sentences expressing attitudes towards objects (for instance), the author points out how expressions about objects can be reformulated in a propositional way. For example, instead of saying "I imagine Manhattan", one could say "I imagine that there exists Manhattan". By doing so, any other kind of imagination can be reduced to propositional imagination.

Once the theoretical and philosophical motivations have been set, and once the new operators have been defined, the author defines a different axiomatic system for each of the three notions introduced in his work. By themselves, these three systems share the same axiomatic and semantic rules (which are based on the standard treatment of modal operators in standard modal logic; see Section 3.1.2 for more details); then, in order to define the interactions between the three concepts, Costa-Leite merges the three axiomatic systems and uses (his words, page 110 in [21]) "basic philosophical intuitions to determine which are the interesting axioms to be added". These philosophical intuitions seek to capture the intuitions of Descartes and Hume, introduced earlier in Costa-Leite's work.

The main three interactions capturing the philosophers' views are the so-called Descartes-Vasiliev law (stating that imagination implies conception), and the so-called laws of Hume (relating imagination, conception and possibility), which Costa-Leite formalizes in the following way:

$$I\varphi \to C\varphi$$
$$C\varphi \to \Diamond\varphi$$
$$I\varphi \to \Diamond\varphi$$

The logical system that follows from merging the three axiomatic systems defined for I, C and $\Diamond$, plus the three previously stated laws, is called IMAG, and it is proved to be sound and complete. When analyzing the properties of IMAG, the author notes how notions such as "imaginability" ($\Diamond I$) or "conceivability" ($\Diamond C$) cannot be reduced to simply "imagine" or "conceive"; moreover, the author claims that, as this logic account for the standard metalogical properties, it can be very useful in order to settle disputes regarding the interactions of the three notions it account for.

On one of his final remarks, it is worth noting that, although the author accepts Descartes' ideas according to which imagination and conception are two kinds of weak possibility, he rejects Hume's approach in which imagination and conception can be collapsed into a single concept.

3.2.4 Wansing: Doxastic Control Through Imagination

In [54], Heinrich Wansing starts by pointing out a rather important feature of imagination: unlike other mental attitudes like belief, imagination is *agentive*; i.e., the agent can decide what she wants to imagine. Following the ideas of [56] (see Section 2.3.2 in Chapter 2 for details), the author explains how, taking into account that imagination can be used to affect our belief and even to form new ones, it can be also understood as a way of indirect doxastic control over the agent's beliefs and knowledge.

In order to capture the agentive character of imagination, and after briefly discussing Niiniluoto's work, the author presents an alternative that aims to capture this feature by using mechanisms from STIT logics (referring to "Seeing To It That" something is the case, and aiming to represent an agent *choosing* to do something) and neighborhood semantics (as they fall outside the scope of our work, we do not provide the technical details of these mechanisms: we refer to Wansing's own work for details). It is worth noting that, although Wansing does highlight the relation between imagination and belief as one of the main motivations of his work, he chooses to focus only on the imagination attitude (propositional imagination, specifically), instead of working in a multi-modal setting involving both imagination and belief.

When discussing Niiniluoto's work, one of the main problems Wansing points out is the well-known problem of logical omniscience in epistemic logic (see Section 3.1.4 for more on this topic), which resurfaces in Niiniluoto's work as the problem of "logical omni-imagination". Although Niiniluoto proposes, in his work, the use of non-normal or impossible worlds[8] to overcome this issue, Wansing sees the proposal as a rather drastic one. His approach is based on proposing another semantics for imagination ascriptions and using neighborhood functions (which do not define binary relations between individual possible worlds, but rather between sets of worlds and families of sets of worlds). Without digging into the technical details of the formal expression, the semantics of Wansing's main

[8] In a nutshell, states of affairs that can be either "inconsistent" (making a formula φ both true and false at the same time) or "incomplete" (making a formula φ lack any truth-value); see Section 3.2.5 about Franz Berto's approach to imagination, in which he uses this kind of worlds.

operator can be defined as follows (where formula $I_j A$ stands for "agent j imagines A", expression $||A||_{\mathcal{M}}$ refers to the set of possible worlds in model $\mathcal{M}$ in which formula A holds, and $N_j(w)$ is the neighborhood function determining, for agent j, the sets of worlds that are accessible from a specific world w):

$$\mathcal{M}, w \vDash I_j A \text{ iff } ||A||_{\mathcal{M}} \in N_j(w)$$

Informally, the previous expression can be read as "$I_j A$ is satisfied in a world w of a model $\mathcal{M}$ if and only if the proposition expressed by A belongs to the set of propositions contained in j's mental image at w in model $\mathcal{M}$" (note how, according to this interpretation, the sets of worlds accessible through the neighborhood function correspond to the mental images agent j can access from world w).

Then, and in order to interpret imagination as an action that the agent decides to carry out, neighborhood semantics are combined with the *dstit*-operator from STIT-theory and branching-time structures (which are tree-like structures formed by different instants of time branching towards the future). By combining these mechanisms, Wansing interprets $I_j A$ as "agent j sees to it that j's mental image contains the proposition expressed by A", thus capturing the choice agent j makes when she *decides* to imagine A.

At the end of his work, the author points out how this approach to imagination not only captures the desired agentive flavor, but it also provides a useful framework for taking into account possible uses of strategic imagination; that is, performing different acts of imagination to successfully reach a desired, imaginary state.

3.2.5 Berto: Conception, Impossible Worlds and Aboutness

The works of Francesco Berto on imagination and conception use non-normal or impossible worlds in [2] and [4], and then an alternative possible-world approach in [3]. In each of these works, Berto presents the semantics of three different logical approaches to imagination, and then discusses some properties of such systems.

His way of understanding the notions of "imagination" and "conception" is related to the way Chalmers characterize the notion of *positive conceivability*: when one positively conceives that p, one does not just assume or suppose that p, but rather she represents in her mind a scenario or state of affairs (that is, a configuration of objects and properties) truthfully descibed by p. Therefore, the author uses both notions indistinguishably in his work.

Regarding the kind of semantics Berto uses in the earlier works, the reasons for using non-normal worlds are two-folded. Firstly, non-normal worlds allow to represent agents that are not "ideal", in the sense that they are not logically omniscient (see Section 3.1.4 for more details on this issue). Secondly, and as Berto makes explicit in [2] (page 103), he wants to represent an agent who can imagine both what is possible and what is impossible; by 'impossible', the author means what obtains at no possible world at all, and not only what could contradict, for instance, the actual physical laws of our world. The author calls *glutty* worlds those worlds in which some formula may be, at the same time, true *and* false (thus representing inconsistent or contradictory states of affairs), and he calls *gappy* worlds those worlds in which some formula may be neither true *nor* false (thus representing "incomplete" worlds where some formulas lack any truth-value).

When using non-normal worlds in [2] and [4], the models defined by Berto use different kinds of worlds; namely, possible and impossible worlds. Whereas in [2] the author distinguishes between *intensionally impossible* (in which intensional operators such as implication and modal operators can behave anarchically) and *extensionally impossible* worlds (in which even the extensional vocabulary can behave anarchically), in [4] he just distinguishes between possible and impossible ones, without making any further distinction regarding impossible worlds. Intuitively, possible worlds cannot be neither glutty nor gappy, whereas impossible ones can.

In order to be able to represent these two kinds of impossibilities, the author defines two different valuation functions: a "positive" valuation $\Vdash^{+}$, stating which formulas are true at which worlds, and a "negative" valuation $\Vdash^{-}$, stating which formulas are false at which worlds. This way, a formula can be both assigned to be true *and* false at a certain world, or may be assigned to be neither true nor false at a certain one (recalling that, while working with logics that allow inconsistencies, the negation of a formula is not usually defined as "the formula not being true", as it normally is in classical logics, and as we did in Section 3.1.2 when defining the semantics of classical modal logic).

Once the apparatus needed to represent both possible and impossible worlds has been set, Berto defines the semantics of a new modal operator standing for conceiving or imagining. In [2], the author starts by proposing an new operator ®, called the *naive representation* operator, and according to which ®A stands for "it is represented that A", or "it is conceived that A". Later in the same paper, and also in [4] and [3], the author defines another modal operator $[A]$ (or, rather, a set of modal operators, as each one is signed with a specific formula A), according to which $[A]B$ is interpreted as "it is imagined in the act whose explicit input is A, that B". Intuitively, this operator expresses whether, when imagining something with an explicit content A, another formula B also holds in the imaginary situation; in this sense, the content of an imagination act plays a similar role to the antecedent of a conditional.

After having set the formal grounds of his proposal, Berto discusses the properties these operators should account for, regarding how imagination and conception should behave, and while basing his intuitions on the literature regarding *ceteris paribus* reasoning. Among these properties, the author discusses both what imagination should account for, and also what it should not.

In [3], Berto proposes an alternative semantics for imagination in which he does not use impossible worlds, but only possible ones; besides, he also introduces 'aboutness' into his system. In short, 'aboutness' is "the relation that meaningful items bear to whatever it is that they are *on* or *of* or that they *address* or *concern*" (taken from [3] itself, page 1). By introducing aboutness into the system, the author wants to address the issue of determining what should be specified or set into an imaginary world, and what not, according to what the agent is explicitly imagining. For instance, when conceiving something about "Sherlock Holmes", one may be entitled to import whatever she knows about London, as the fictional character is explicitly set to live in there, but one should not be entitled to import whatever she knows about quantum mechanics, as they has nothing to do with the fictional character.

Aboutness is introduced in the logic as a syntactic mechanism that determines what a formula is about, according to the atomic propositions that are part of that formula. This mechanism is then used to redefine operator $[A]B$ by requiring not only that B obtains in those worlds accessible through the imagination relation with content A, but also requiring B to be something A is about. This approach allows to filter up what is being detailed or imported when the agent conceives A, and helps preventing a sort of "over-specification" in which the conceived world ends up being "too detailed", concerning things that are not even related to what the agent is imagining.

3.3 Towards a New Approach

The main goal of our work is to perform a systematic analysis of acts of imagination, understood as dynamic processes; one of the tools we want to use for aiding us in such analysis is a formal logic allowing us to represent *how* imaginary worlds are created and *what* processes take place, within the mind of the agent, when creating such worlds. In the previous pages we have briefly reviewed some of the main contributions that relate logic and imagination, as summarized in Table 3.1. Although we have skipped the formal details of those works, the review we provide is enough to see how their authors understand and characterize imagination acts, and, more importantly, what did they choose to highlight when they defined a logic for imagination. We also emphasize, in the same Table 3.1, how these systems fail to represent some of the main features we want to capture in our analysis of imagination acts as dynamic actions.

As we can see, each one of these logics is missing a crucial part of what we want to represent: the dynamics involved in acts of imagination, understood as processes involving a change in the content of the mind of the agent. The previously reviewed logics are all based on static, predefined models in which a certain set of imaginary scenarios are considered, but which are nonetheless limited to those: the acts of imagination involved in creating those imaginary scenarios have already been performed, the new imaginary content has already been created an added to the mind of the agent, and the agent can already "access" these imaginary worlds. The process involved in imagining something and, in particular, in creating a brand-new imaginary world, is somehow previous to the definition of the model, and therefore hidden in those systems.

When reviewing Nichols and Stich's, Williamson's, and Langland's theories, we have seen how the authors analyze and dissect *how* an agent engages in acts of imagination, and also what happens within the mind of the agent, on a representational level, when she engages in an act of imagination. However, the logics we have just reviewed in this section do not account for this, or at least not fully. Wansing's proposal, for instance, does capture the agentive nature of acts of imagination by introducing the STIT mechanics and neighborhood semantics: in his logic, the agent can indeed be seen as voluntarily choosing what she wants to imagine. Through the notion of "closeness" between counterfactual worlds and the real one, Lewis' work can be interpreted as capturing the fact that imaginary worlds are similar (or "close") to our real world in many ways, which is related to the mirroring effect and the reality-oriented elaboration of imaginary worlds. Similarly, Berto's proposals move towards capturing how the "content" of acts of imagination characterizes what would be the case in the resulting imaginary worlds; specifically, his work dealing with "aboutness" captures relevant notions of the reality-oriented development of imaginary worlds.

Nevertheless, none of the logics we have reviewed can account for the dynamics of imagination acts, and so we must look for an alternative approach. Therefore, the path we take in the remaining of this work goes in a different direction than the system we have

	Main goal	*Formal Framework*	*Agentive / Voluntary*	*Reality-oriented development*	*Dynamic*
Lewis	*Ceteris paribus* hypothetical reasoning	Modal system, spheres	No	Yes	No
Niiniluoto	Imagination as a propositional attitude	Modal system, first-order discussion	No	No	No
Costa-Leite	Relation between imagination, conception and possibility	Multi-modal system	No	No	No
Wansing	Agentive character of imagination	Neighborhood semantics, STIT logic	Yes	No	No
Berto	Characterize the outcomes of imagination	Normal and non-normal worlds, aboutness	No	Yes	No

Table 3.1: Main features of the existing logics of imagination.

just reviewed: on Chapter 4 we focus on the dynamics of acts of imagination by defining a logical system in which it is possible to execute an action that creates and defines new imaginary worlds in real-time; specifically, this is done by expanding the initial model through an algorithm encoding the processes presented in the theories of imagination reviewed earlier.

Chapter 4

The Logic of Imaginary Scenarios

In this chapter, we present a formal system aimed to capture how voluntary imagination acts work, according to the theories reviewed in Chapter 2: the Logic of Imaginary Scenarios. We define its language and semantics, and we discuss the system's strong and weak points.

4.1 Towards an Algorithm for Imagination Acts

Before diving into the deep waters of formality, we want to recap what is it that we are looking for. Our main interest, and so the main goal of the present work, is to unfold, in a systematic way and aided by formal tools, the dynamics involved in voluntary imagination acts; in other words, what happens when we create and develop an imaginary world that was not there before, and that we know it is neither real, nor believed. We are not concerned with *what* results from an act of imagination, but rather with *how* this act of imagination is performed.

In this first section, we discuss and sketch the intuitive mechanics of what we call the *Imagination Algorithm*. This algorithm will then be captured in a formal way when defining the Logic of Imaginary Scenarios, and it will be the process governing how such imaginary scenarios are created and developed.

As we have already seen at the end of Chapter 2 (specifically in Section 2.4), the reviewed theories of imagination have an underlying structure that points out to a sort of common algorithm determining how imagination acts work. Specifically, performing an act of imagination involves the following mechanisms:

1. Create a new representation of an imaginary scenario characterized by a certain

initial premise or condition.

2. Elaborate on the details of the imaginary scenario by mirroring reality-oriented rules and facts that are either known or believed by the agent, and that provide details about "how the scenario would be, if it was real"[1].

3. Add new premises to the scenario that, although they would not usually follow from it (based on reality-oriented mirroring), they are intentionally added by a decision of the agent.

Now, the questions holding the key to the algorithm's nuts and bolts are: how is the new imaginary scenario created, and how do the initial premises characterize it? Where do the agent's knowledge and beliefs come from, and how do they affect what is already the case in the imaginary scenario? How do new premises get added to an imaginary scenario that already exists, and that is already determined by a certain state of affairs?

4.1.1 Creating a New Imaginary Scenario

When an agent decides to perform an act of imagination, the overall algorithm governing such decision must start by finding a spot within the mind[2] of the agent in where to create the imaginary scenario she is going to entertain.

The content of this new imaginary scenario must be defined, at the beginning, only by the initial premise used to create it. The reason why we require this to be the case follows, for instance, from what Nichols and Stich define in their cognitive theory of pretense (see Section 2.3.1 for details): the initial premise defining how an imaginary scenario should be have preference over anything else and, therefore, it is "clamped" into such scenario. This is done in order to be able to conceive anything that may go against our knowledge or beliefs: if we could not clamp the initial premise characterizing an imaginary scenario, then any premise contradicting our beliefs would be instantly overridden when merged with them. Without prioritizing the initial premise over our knowledge and beliefs, imagining something differing from them would be impossible.

[1] Keep in mind that, by saying how an imaginary world would be "if it was real" or "in a reality-oriented way", we mean "according to what the agent knows and believes". By using this kind of expressions we refer to the way the agent thinks the real world would behave, according to what she knows and believes about it, if the situation represented in the imaginary world was the case. Therefore, when using these kind of expressions, we also mean "related to the agent's knowledge and beliefs".

[2] Similarly to [43] (in page 121), we want to stress the fact that we use expressions like "finding a spot", or "place within the mind" without implying, nor defending the existence of any kind of separate, specific physical place within the agent's mind.

Take, for instance, the example in which I want to imagine that I can fly, and suppose that I do *not* clamp this initial supposition into the imaginary scenario. As soon as I start elaborating on it, I would inevitably stumble upon my knowledge that, in fact, I *cannot* fly in the real world, and so this fact will override the premise saying that I can do it. The imagining, initially intended to represent a world in which I could actually fly, would lose all its reason to be and would end up being, simply, the same representation of the real world, as determined by my knowledge and beliefs.

Figure 4.1 represents what we require of this first mechanism: when executing an act of imagination, the first thing we must do is to create a new imaginary scenario in which we can clamp the initial premise φ that characterizes it.

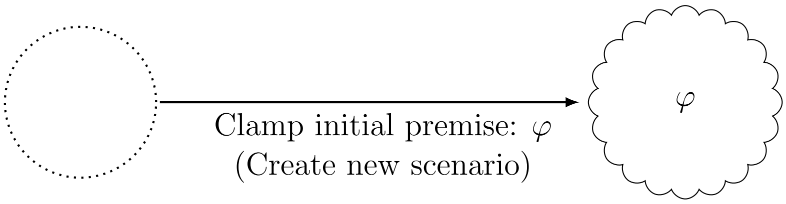


Figure 4.1: Creating a new imaginary scenario clamps the initial premise.

4.1.2 Importing Knowledge and Beliefs

Once the new imaginary scenario has been created, and the initial premise has been clamped in it, it is time to elaborate on the details of the scenario. Typically, and as we have seen when reviewing the theories of imagination in Chapter 2, the way imaginings develop follow the same reasoning mechanisms used for our beliefs; moreover, we usually fill in the blanks of an imaginary scenario using whatever we know and believe about reality, and we expect the imaginary scenario to develop in the same way it would do, if it was real. This phenomenon accounts for both what we labeled as the "reality-oriented development" (in Section 2.4, and following the theories previously reviewed in Chapter 2), and for the so-called "mirroring" effect that relates imagination and belief (as explained in Section 2.2.2).

When integrating this mechanism into the overall algorithm, we must look back to where the initialization of the imaginary scenario took place. In the previous section we have created an imaginary scenario and clamped the initial premise in it: now, in order to fill in the details of the scenario that are not explicitly set by the initial premise, we must look into the agent's knowledge and beliefs. We must take into account, nevertheless, that the initial premise should still be considered clamped, and so it must have priority over what the agent knows or believes in the real world; otherwise, we would risk losing up, precisely, what makes the imaginary scenario different from reality.

In order to do so, the agent should check all her knowledge and beliefs about the real world and, if they do *not* contradict the premise clamped into the imaginary scenario, then import them into it. For instance, by taking again the imagining in which I can fly, I would import my knowledge about the color of my hair, the way I am dressed and the way my house looks like (among many other things). When considering the fact that I know I cannot fly, however, I should *not* import it into the imaginary world, as it would be in contradiction with the initial premise clamped in it. Therefore, this fact should be skipped when considering what to import into the imaginary scenario in which I can fly.

Figure 4.2 represents the way this mechanism works. Note how we draw this mechanism upon the previous Figure 4.1, in order to show how both mechanisms work together. In Figure 4.2, the circle on the left side contains what the agent knows and believes about the real world (in this case, $\neg\varphi$ and ψ); note how, when importing the agent's knowledge and beliefs, as $\neg\varphi$ would contradict φ (which is the initial premise that was clamped into the imaginary scenario, and so it has preference), it is *not* imported.

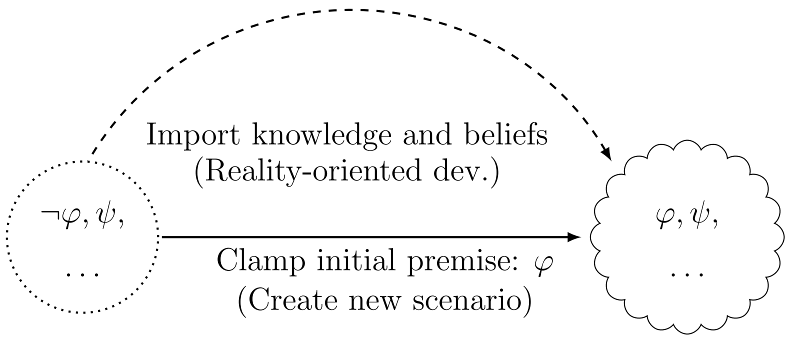


Figure 4.2: Import facts about the real world.

4.1.3 Adding New Premises

The last mechanism involved in an act of imagination is very similar to the first one; in fact, and as Langland-Hassan suggests in his theory (see Section 2.3.3), this voluntary addition of new premises into the imagining can be seen as a cyclical process that begins anew the whole cycle. We follow Langland-Hassan in this mechanism, and thus treat this addition of new premises as a sort of new imagining-initiating mechanism; there are, nevertheless, some important remarks we want to make.

1. Firstly, this addition of new premise differs from the initial creation of the imaginary scenario in an important way; namely, the initial premise used to characterize a new imaginary scenario were clamped on a blank, brand-new imagining, whereas the new premise that must be added in this process should be clamped into an imagining which has been already characterized in a previous step, and also elaborated through importing the agent's knowledge and beliefs. Therefore, the new premise that is to be added in this mechanism may be in conflict with something that is already the case in the imaginary scenario. As this premise should be clamped into the imagining as well, it should be prioritized over what is already the case in there.

   In order to illustrate this, suppose that I am still entertaining the scenario in which I can fly, and that the scenario has already been elaborated by importing my knowledge and my beliefs. Now, suppose that one of the facts that I imported described the way I am actually dressed; say, jeans and a white shirt. However, and as I am enjoying this imaginary scenario, I now want to picture myself dressed as a kind of superhero, and so I want to add a new premise representing the fact that I am dressed in a sort of black, cool-looking armor. As this new premise has preference over everything else, the previously imported fact about my jeans and shirt should be removed from the imagining, as it would conflict with the new, prioritized premise.

   Moreover, I could even want to add a new premise saying that, suddenly, I lose my power of flying. Although in a previous step I created an imaginary scenario with the premise "I can fly", and that premise was clamped into the scenario, it would now conflict with the new premise I want to add. Following Nichols and Stich's theory (see Section 2.3.1), as soon as the process of importing knowledge and beliefs is over, the initial premises lose their "privileged status" of being clamped, and so they can be removed when new premises come into play. In such situation, the premise "I can fly" would be removed from the imaginary scenario in order to accomodate the new "I cannot fly" premise.

2. Secondly, there is another important fact concerning the reality-oriented development that is relevant here. When defining how knowledge and belief import works, in the previous section, we were supposing that a new imaginary scenario had been

just created, and so that it was empty (besides from the initial premise) and ready to be filled with our knowledge and beliefs about the real world. However, when we add a new premise to an already existing imaginary scenario, this new premise elaborates not upon reality, but upon the scenario we were already imagining. For instance, when adding the premise about the way I dress into the already created imagining of me being able to fly, I should not elaborate on this scenario by importing my knowledge about the fact that I cannot fly in the real world; rather, I should import the previously added premise saying that I can fly, even if it conflicts with my real-world knowledge.

Therefore, when importing the agent's knowledge and beliefs after adding a new premise to an imaginary scenario, the "world" in which the agent should look for her knowledge and beliefs should not be the real world, but rather the imaginary one. Not only this represents the fact that, when adding a new premise, we build upon the imaginary scenario, but it is also consistent with the fact that, whichever knowledge and beliefs were consistent with the previous premises, they have already been imported to the imaginary scenario that is being built upon. According to this, the reality-oriented development will still work as expected when importing knowledge and beliefs from the previous imaginary scenario, except for those facts that would contradict something that was already introduced into the imagining.

3. Finally, note that adding certain premises may actually amount to "retrieving" certain facts about the real world that were not initially imported due to conflicts with previous premises. In the scenario where I imagine that I can fly, I do *not* import my knowledge about the fact that I cannot fly, because it would contradict the initial premise. Now, in every elaboration of this scenario in which I still imagine that I can fly, my real knowledge keeps being "blocked" by the initial premise. However, if I now add to that imagining that I suddenly lose the ability to fly, the fact that "I cannot fly" will be added to the scenario. This will not only override a premise that was previously added to the imagining, but it will add a new fact into it that, aside from not having been there before during the whole imagining, it will now coincide with my knowledge about the fact that, in the real world, I cannot fly.

Figure 4.3 includes this last mechanism upon the previous Figure 4.2, and thus represents the whole Imagination Algorithm. Concerning this last mechanism, note how the process of importing the agent's knowledge and beliefs is now based upon the previous imaginary world, rather than the real world that was used before. In this example, the new premise γ added in the voluntary addition does not contradict any of the facts already present in the imaginary world where the new premise must be added. In particular, the initial premise φ, which was previously used to characterize the first imaginary world, does not conflict with the new premise γ, and so it is also imported into it by the reality-oriented development mechanisms.

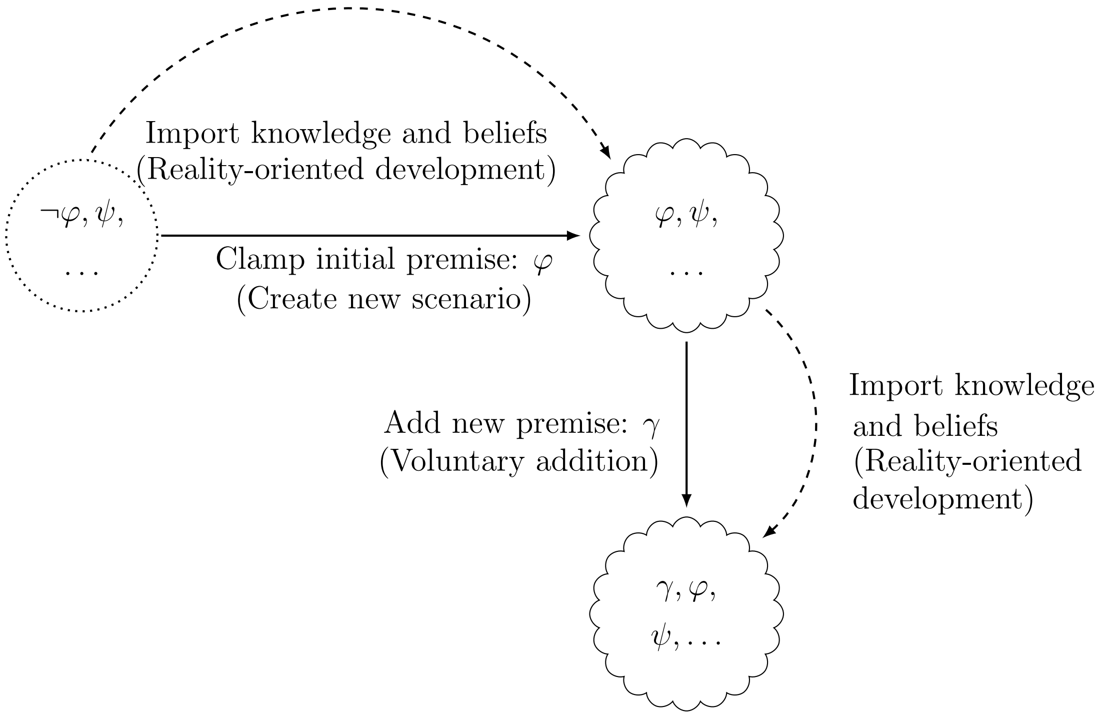


Figure 4.3: The full Imagination Algorithm.

4.1.4 Wrapping Up the Imagination Algorithm

The previous sections provide insights about the different mechanisms that take part in an execution of the Imagination Algorithm. Figure 4.3, specifically, represents the whole cycle of executing the Imagination Algorithm. In fact, it represents more than that, as the addition of a new premise into an already existing imaginary scenario begins the process anew, and so corresponds to a new call to the algorithm.

Therefore, and by taking into account how we understand the voluntary additions discussed in the previous section, we intuitively define an execution of the Imagination Algorithm as follows:

1. The execution requires an initial premise φ characterizing the initial scenario, and a *world of reference* upon which the agent bases her imagining.

2. The algorithm creates a new imaginary world and clamps φ into it.

3. The algorithm goes over what the agent knows or beliefs in the world of reference, and, if it does not contradict φ, imports it into the imaginary world.

4. The imaginary world is related to the world of reference through an act of imagination executed by the agent, with initial premise φ.

Note how, by introducing the notion of "world of reference", we already account for both those acts of imagination used to create a new imaginary world (and thus the ones that take the real world as the world of reference), but also for those acts of imagination used to add new premises into an already existing imaginary world (in which case, the world of reference is not the real world, but also an imaginary one).

Figure 4.4 highlights, using the same example as the one depicted in Figure 4.3, the two different acts of imagination that take place in there: the one used to create a new imaginary world from scratch by using an initial premise φ (and importing the agent's knowledge and beliefs from the real world), and the one used to add a new premise γ into an imaginary world that already exists (importing knowledge and beliefs from that imaginary world, in this case). For the sake of readability, we omit the labels explaining the relations.

After having discussed and specified how the Imagination Algorithm should capture the theories reviewed in Chapter 2, we have settled our desiderata regarding how acts of imagination should behave in our formal proposal. In the following sections, we introduce all the formal details needed to define the Logic of Imaginary Scenarios, including a formal definition of the Imagination Algorithm, which defines how new imaginary worlds are created and elaborated, according to our logical system.

4.2 Syntax

As explained in Chapter 3, we do not build our proposal from scratch. As imagination is related to other mental attitudes, we want to define our system upon a logic already able to handle, at least, some of those mental attitudes; however, we also want to build our proposal step by step, and without being overwhelmed by technical difficulties inherited

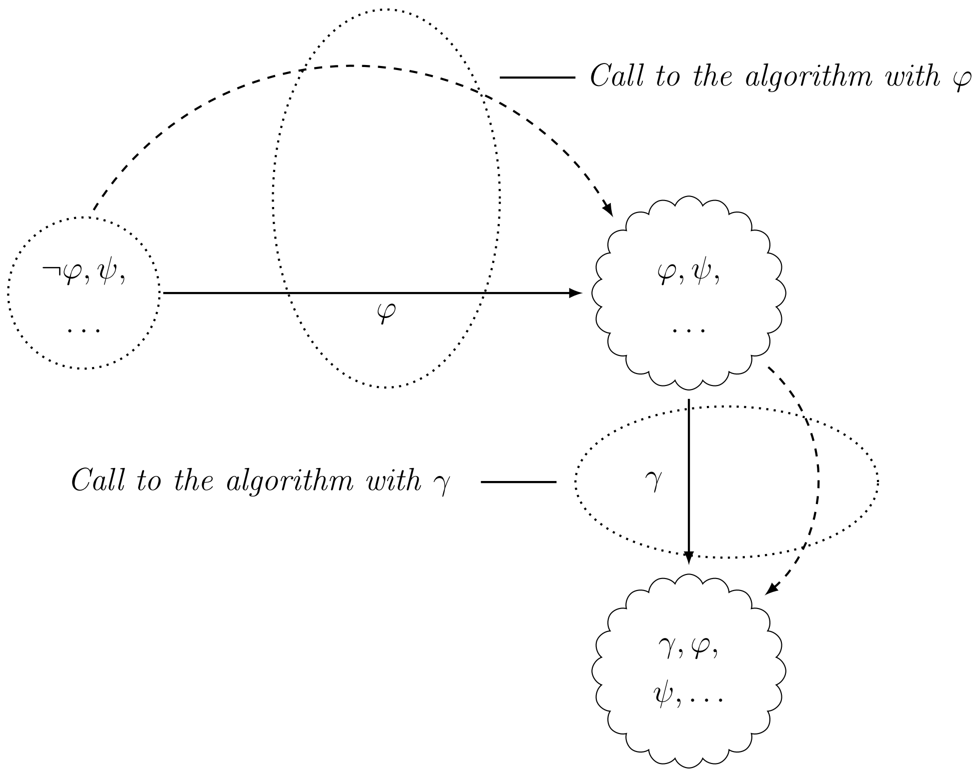


Figure 4.4: Two different executions of the Imagination Algorithm.

from the background system used. Therefore, we have build our proposal upon the *single-agent epistemic logic* (which we have briefly introduced in Section 3.1.4). Furthermore, we also add some features of *hybrid logic* (which we introduced in Section 3.1.3) into our initial mix.

It is worth noting that, while presenting the language and semantics of our logic, there will be some elements which we will need to mention before introducing: this is because the language, the models and the dynamic part of this proposal (handled by the Imagination Algorithm) are closely related between them. However, we will try to give an intuitive understanding of each notion before formally defining it.

The language of the *Logic of Imaginary Scenarios* is formed by a countably infinite set of *atomic formulas*, called ATOM, and represented by the lowercase letters $p, q, r \dots p_1, p_2 \dots$; besides, there is a countably infinite set of *nominals* (taken from hybrid logic), represented by the lowercase letters $i, j, k \dots i_1, i_2 \dots$ and called NOM.

We use the standard propositional operators $\neg, \wedge, \vee, \rightarrow$ (standing for "negation", "conjunction", "disjunction" and "material implication", respectively); besides, we include the unary "knowledge" operator K, taken from epistemic logic, and the unary "at" operator @ taken from hybrid logic.

Furthermore, we also introduce two new operators: a dynamic unary operator $\text{Img}(\delta)$ called "dynamic imagination" and an unary modal operator $\langle I(\delta)\rangle$ called "static imagination"; both operators are signed with a formula δ of a special kind —which we introduce in the following lines.

We use bracket symbols (, [,),] as usual, and usually omit them when the context is clear. Now, the well-formed formulas of the language are:

$$i \mid p \mid \neg\varphi \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid \varphi \rightarrow \psi \mid K\varphi \mid @_i\varphi \mid \text{Img}(\delta) \mid \langle I(\delta)\rangle\varphi$$

where $i \in \text{NOM}$, $p \in \text{ATOM}$, $\{\varphi, \psi\} \subseteq \text{FORM}$ and $\delta \in \text{FORM}^*$. We define $\text{FORM}^* \subset \text{FORM}$ as follows:

$$p \mid \neg\varphi \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid \varphi \rightarrow \psi$$

where $p \in \text{ATOM}$ and φ either in ATOM or in FORM^*. That is: FORM^* is the propositional fragment of FORM. From now on, we use variables $\delta, \gamma, \ldots$ to refer to elements of FORM^*, in order to distinguish them from formulas belonging to FORM.

We also introduce two symbols $\top, \bot$ to refer to *truth* and *falsity*, respectively, and we define them as follows (for $p \in \text{ATOM}$):

$$\top \equiv p \vee \neg p$$
$$\bot \equiv p \wedge \neg p$$

The new operators introduced to represent imagination are $\text{Img}(\delta)$ and $\langle I(\delta)\rangle\varphi$. The intuitive, informal reading of the former operator would be "the agent creates an imaginary scenario using premise δ", whereas the latter stands for "in an imaginary scenario initiated by premise δ, it is the case that φ".

Following the theories previously reviewed in Chapter 2, and, in particular, the intuitive understanding of the Imagination Algorithm introduced in Section 4.1, the dynamic operator $\text{Img}(\delta)$ is intended to entail a call to the Imagination Algorithm: in particular, it captures the fact that the agent decides to initiate an imaginary scenario characterized by an initial premise δ. Aside from what we informally introduced in Section 4.1, we formally develop, in Section 4.4, an algorithm that describes *how* an imaginary scenario must be created.

Regarding the static operator $\langle I(\delta)\rangle$, note how it corresponds to a sort of static evaluation of an act of imagination that has *already* been performed: in this sense, this operator does not aim to represent any of the mechanisms involved when performing an act of imagination, but rather to evaluate an imaginary scenario, once it has already been created. This static operator, therefore, is more in accordance to the kind of operators used in the existing logics of imagination reviewed in Chapter 3.

It is worth noting that, although we build our proposal upon an existing logic, we want to keep this underlying system as simple as possible: as a consequence, we do not have an explicit representation of *beliefs* in our logic. Therefore and for the moment, we use a derived operator M to represent a weak form of belief, and we interpret it as a complementary of knowledge, following what we saw in Section 3.1.4:

$$M\varphi \equiv \neg K \neg \varphi$$

Intuitively, if the agent does *not* know $\neg\varphi$, then it is because she considers that φ could be the case as well: therefore, we could say that the agent believes φ (understanding this notion of "believing" as considering it possible to be the case, as far as the agent knows). Although this is a rather simplified account of beliefs, it will allow us to concentrate, at this stage, on our main goal, which concerns the dynamics of imagination acts. In Section 4.8.2 we discuss the consequences of this decision, and we point to ways of amending it in an extended version of the logic we propose here.

4.3 The Models for Imaginary Scenarios

Similarly to what we do with the language of our logic, we build our models upon a standard model of single-agent epistemic logic, plus the elements introduced by hybrid logic. We take this model as basic, and we add a new accessibility relation upon it in order to account for imagination acts.

It is worth stressing the fact that, unlike most logic systems that represent static scenarios, our models are intended to represent the change involved in performing an act of imagination; therefore, they are *dynamic* by definition.

Typically, our models will be initially defined as being single-agent epistemic models, without any act of imagination represented in them yet. The interest of our proposal is, precisely, to allow for these acts of imagination to "happen" within our model, and thus expanding our model when they happen. Therefore, we require every element of the relation R_{Img} (introduced in the following lines, and representing an act of imagination) to be created by explicitly following our Imagination Algorithm (formally defined in Section 4.4), thus ensuring that both the accessibility relation, and the imaginary worlds created by the algorithm, fulfill the conditions imposed by the way the algorithm behaves.

Definition 4.3.1. *We define a Model for Imaginary Scenarios as a structure* $\mathcal{M} = \langle W, R_K, R_{Img}, V, N\rangle$ *formed by the following elements (aside from* R_{Img}*, they are all taken from single-agent epistemic logic and hybrid logic):*

- W *is a non-empty set of elements called "possible-worlds" or "states of affairs". We use the lowercase letters* $w, v, u, \dots w_1, w_2, \dots$ *to refer to elements of* W*.*

- $R_K \subseteq W \times W$ *is a binary relation over elements of* W *called the "indistinguishability relation", and which we require to be reflexive, transitive and symmetric (due to restrictions typically imposed to knowledge: see Section 3.1.4 for details). Intuitively, this relation establishes which possible worlds the agent thinks that can be the actual case, as far as she knows. We use pairs of the form* $(w, v), (v, u), \dots$ *to refer to elements of* R_K*.*

- $R_{Img} \subseteq W \times W \times FORM^*$ *is a ternary relation called the "imagination relation". Intuitively, an element* (w, v, δ) *captures how, by performing an act of imagination with content* δ*, and by taking* w *as the world of reference (in terms of being the possible world the agent considers to represent the actual case), an imaginary world* v *is created. We use triplets of the form* $(w, v, \delta), (u, z, \gamma), \dots$ *to refer to elements of* R_{Img}*.*

- $V : ATOM \to \mathcal{P}(W)$ *is a function from atomic formulas of the language to subsets of the power set of* W*, called the "valuation function". Intuitively, it keeps track of which atomic formulas are true at which subset of possible world.*

- $N : NOM \to W$ *is an exhaustive function setting, for each element of NOM, a possible world in* W*. Intuitively, this function sets which nominal is used to identify each world.*

By the way imaginary worlds are created by the Imagination Algorithm, a Model for Imaginary Scenarios represents different "clusters" of possible worlds; later on, in Section 4.6, we present a detailed example that shows how the system works.

4.3.1 A Brief Remark on Terminology

At this point, and after having introduced the formal elements of our logic, it is important to stress the difference between certain terms that we will be using from now one in a specific way:

- We use the terms *imagination act* or *act of imagination* to refer to the outcomes of a single execution of the Imagination Algorithm; this outcomes include all the new imaginary worlds, the new imagination relations R_{Img}, and also the execution of the Imagination Algorithm itself. In this sense, and following Langland-Hassan's intuitions (as presented in Section 2.3.3), we interpret the execution of the Imagination Algorithm over an already existing imaginary world as a cyclical involvement of the whole process of imagining.

- We use the terms *real possible world* or *real world* to refer to a possible world $w \in W$ that may either be the origin of an element of the imagination relation (that is: $(w, v, \delta) \in R_{Img}$), or may not appear in R_{Img} at all, but which is *never* the destination of an element of R_{Img}. By using this term we refer to one of the possible worlds or states of affairs that the agent considers that could represent the actual world, and thus are not a product of any imagination act.

- We sometimes use the term *reality* to refer to the set of real possible worlds the agent considers that can be the actual case. Similarly, we often refer to such set of worlds by using expressions like "how things actually are", or "the actual state of affairs", or even "the real world" (in a more loose sense than in the previous case).

- Conversely, we use the terms *imaginary possible world* or *imaginary world* to refer to a possible world $v \in W$ which is the destination of, at least, an element of the imagination relation $(w, v, \delta) \in R_{Img}$ (and which may also be the origin of another element of R_{Img}, if it has been used as the new world of reference for another imagination act). Therefore, by using this term we refer to a specific possible world or state of affairs that was defined as a result of an act of imagination.

- We use the terms *imaginary scenario* or *imagining* to refer to the whole set of possible worlds $\{v_1, \ldots, v_n\} \in W$ that result from a *single* execution of the Imagination Algorithms, and that have have $\{(w, v_1, \delta), \ldots, (w, v_n, \delta)\} \subseteq R_{Img}$ (for the same w and δ). Therefore, when we talk about an imaginary scenario or an imagining, we are referring to any of the imaginary possible worlds that have been defined during a particular act of imagination. The difference between the way we use the terms imaginary scenario or imagining, and imagination act, is that the former terms are used to refer explicitly to the set of possible worlds created by the Imagination Algorithm, whereas the latter refers to both the execution of the algorithm, and every outcome that results from it.

4.4 The Imagination Algorithm

In the present section, we are going to define how the Imagination Algorithm, which we intuitively introduced in Section 4.1, works, with respect to the language and the models we have already presented in Section 4.2 and Section 4.3, respectively. Let's start by recalling which are the mechanisms embedded in the Imagination Algorithm:

1. The execution requires an initial premise δ characterizing the initial imaginary scenario, and a *world of reference* upon which the agent creates her imagining.
2. The algorithm creates a new imaginary world and clamps δ into it.
3. The algorithm goes over what the agent knows or beliefs in the world of reference, and, if it does not contradict δ, imports it into the imaginary world.
4. The imaginary world is related to the world of reference through an act of imagination executed by the agent, with initial premise δ.

Now that we have already defined the kind of models we will use in the Logic of Imaginary Scenarios, we can be more specific regarding our previous intuitions.

From now on, we use the term `ImgAlg` as a way of referring to the formal Imagination Algorithm defined within the Logic of Imaginary Scenarios. We have said that an execution of the algorithm requires an initial premise, and a world of reference: we will refer to the initial premise as δ (taking into account that, while talking about the formal language, symbols φ and ψ are normally used as variables for formulas in FORM, whereas δ refers to a formula in FORM*), and to the world of reference as w^R. Therefore, a call to the algorithm is expressed as follows:

$$\texttt{ImgAlg}(\delta, w^R)$$

We already know, at an intuitive level, what is the role of δ. When translating its role into the formal approach, δ is a formula that must hold (i.e.; it must be "clamped") at the imaginary world created by the corresponding imagination act. In other words, δ must be used to determine the atomic valuation of the world that will be created by the execution of $\texttt{ImgAlg}(\delta, w^R)$, as it is precisely the atomic valuation of the resulting world which will determine whether δ holds in there.

Now, as we have already specified in Section 4.2, $\delta \in \text{FORM}^*$, which means that δ belongs to the propositional subset of formulas of the language; in other words, there is no occurrence of $K, @, \langle I(\delta)\rangle$ nor Img within δ. Why is that so?

Regarding the restriction on modal operators, it is due to a lack of expressive power of our language. If we wanted to allow our agent to imagine that her knowledge is somehow different, our algorithm would not need to create just a new imaginary scenario (maybe formed by different imaginary worlds, but all of them aiming to account for the same initial premise), but rather a whole relational structure formed by different imaginary worlds, not aimed to represent different alternatives to the same premise, but rather to representing how the agent imagines her knowledge would be. However, if we wanted to do so, we would need to be able to quantify over nominals in order to express things such as "every world that the agent considers possible", or "there exists a world such that". As our interest lies in the process of imagining something, and thus we want to start by focusing on that topic, we want to keep our initial setting simple: adding quantification involves a series of technical difficulties which we will leave for and extended version of the present system.

Regarding the restriction imposed on hybrid operators, it is due to philosophical reasons. It does not seem to make sense for an agent to imagine something *about* a specific, already made and defined world. The agent must indeed take one world as the "world of reference" for her imaginings, but she cannot imagine that *that* specific world changes; rather, she must create a new imaginary world which, although being based on that one, will still be different. Even if the "content" of the imaginary world ends up being exactly the same of the original world, the simple fact that the imaginary world is not real will already make it different. Due to this, imagining that "at a certain world something happens" would not capture what we really do when we imagine.

There is still one further notion that we need to introduce. During an execution of the `ImgAlg`, the Model for Imaginary Scenarios $\mathcal{M}$ upon which the algorithm is executed gets expanded: this is, precisely, what we want to capture by understanding an act of imagination as an action that creates new imaginary worlds. Therefore, during the process of executing the `ImgAlg` within a model $\mathcal{M}$, we end up having more possible worlds and more accessibility relations than we had just before executing the algorithm. In order to refer to this version of the model $\mathcal{M}$, that is being expanded in the current execution of the `ImgAlg`, we introduce one further concept: the *expanded model* of $\mathcal{M}$, to which we refer to as $\mathcal{M}^+$.

During an execution of the `ImgAlg`, the elements of the model $\mathcal{M}$, upon which the algorithm is executed, are expanded; from now on, we may refer to either the whole expanded model $\mathcal{M}^+$, or to any of its elements as follows: $\mathcal{M}^+ = \langle W^+, R_K^+, R_{Img}^+, V^+, N^+\rangle$. We provide the formal definition of each of these expanded elements in the following paragraphs.

Now, the following steps define how the `ImgAlg` works, with respect to a Model for Imaginary Scenarios $\mathcal{M}$, a formula $\delta \in \text{FORM}^*$ and a possible world of reference $w^R \in W$:

1. The algorithm `ImgAlg` starts by being called with arguments δ and w^R. If formula δ is contradictory (that is: if $\delta \equiv \bot$), the execution of the `ImgAlg` ends at this point[3]. We do *not* allow our agent to imagine contradictory scenarios.

2. In order to handle the formula in an efficient way, we compute the *Disjunctive Normal Form*[4] (DNF from now on) of δ, to which we refer as $\text{DNF}(\delta)$. In the following steps of the `ImgAlg`, we refer to the clauses that form the $\text{DNF}(\delta)$ as follows: $\text{DNF}(\delta) = \delta_1 \vee \ldots \vee \delta_n$.

3. The `ImgAlg` must create a new imaginary world for each possible clause satisfying formula δ. Therefore, and recalling that $\text{DNF}(\delta) = \delta_1 \vee \ldots \vee \delta_n$, the `ImgAlg` must create n new imaginary possible worlds, to which we will refer as $w_1, \ldots, w_n$, and add them to W. These new possible worlds define the *expanded set of possible worlds* as follows:
$$W^+ = W \cup \{w_1, \ldots, w_n\}$$

4. Once the new possible imaginary worlds have been created, the `ImgAlg` must create new imaginary relations expressing that, when imagining formula δ at the world of reference w^R, the agent can access the new imaginary worlds $w_1, \ldots, w_n$. This

[3]As `ImgAlg` is executed on a model $\mathcal{M}$, in case δ is contradictory the algorithm does not expand $\mathcal{M}$ in any way; therefore, we can consider that, if δ is contradictory, the algorithm returns $\mathcal{M}^+ = \mathcal{M}$.

[4]In a nutshell, the Disjunctive Normal Formal of a formula δ of propositional logic corresponds to an equivalent formula $\delta_1 \vee \ldots \vee \delta_n$ expressed as a disjunction of clauses δ_i, which are, in turn, conjunctions of literals, which are either atoms (called positive literals), or negations of atoms (called negative literals). For a more comprehensive explanation of how the DNF of a formula can be computed, see [38]. For the present case, it suffices to say that every formula of propositional logic can be expressed in its equivalent DNF formula by following a simple algorithm.

defines the *expanded set of imaginary relations* as follows:

$$R^{+}_{Img} = R_{Img} \cup \Big(\bigcup_{i=1\ldots n} \{(w^R, w_i, \delta)\} \Big)$$

5. As any of the new imaginary possible worlds satisfies what the agent is imagining (specifically, δ), they should all be epistemic alternatives to the other imaginary worlds considered in this execution of the `ImgAlg`; in other words, the agent must consider them all as a possible way of representing an imaginary world satisfying δ. This defines the *expanded set of epistemic indistinguishability relations* as follows:

$$R^{+}_{K} = R_{K} \cup \Big(\bigcup_{\substack{i=1\ldots n \\ j=1\ldots n}} \{(w_i, w_j)\} \Big)$$

6. After that step, the structure of the expanded model has been defined. Now, the `ImgAlg` must add a set of new nominals to refer to the newly created imaginary worlds. This defines both the *expanded set of nominals*, by adding one new nominal k_i for each new possible world w_i created during the current execution of the `ImgAlg`, and the *expanded nominal function*, which is a functional extension of N relating the new pairs of nominals and possible worlds:

$$\mathrm{NOM}^{+} = \mathrm{NOM} \cup \{k_1, \ldots, k_n\}$$

$$N^{+} = N \cup \Big(\bigcup_{i=1\ldots n} \{(k_i, w_i)\} \Big)$$

7. Last but not least, the `ImgAlg` must expand the valuation function to account for the atomic propositions holding at the new imaginary possible worlds. In order to do so, the algorithm must account for both the literals that appear in each δ_i, and also for the atoms that are true in the world of reference w^R and which should be imported to the new imaginary worlds, provided they do not appear in δ; this is so because any atom appearing in δ has preference over the atoms of the world of reference (the agent "clamps" the initial premise into the imaginary worlds). Therefore, the definition of the *expanded valuation function* involves two different phases:

   (a) Firstly, the `ImgAlg` must set the new valuation functions according to the atoms

p appearing[5] in δ_i, for each new imaginary possible world w_i:

$$V_1^+(p) = V(p) \cup \Big(\bigcup_i \{w_i \mid p \text{ is a positive literal appearing in } \delta_i\}\Big)$$

(b) Then, it must import all the atoms that are true at the world of reference w^R, provided they do not appear in δ_i, for each new imaginary possible world w_i:

$$V^+(p) = V_1^+(p) \cup \Big(\bigcup_i \{w_i \mid w^R \in V_1^+(p) \text{ and } p \text{ is not a literal of } \delta_i\}\Big)$$

8. The `ImgAlg` has finished its execution: a new set of imaginary possible worlds satisfying δ has been created, these worlds are accessible through the imagination relation R_{Img} from the world of reference w^R, and they are epistemically indistinguishable by the agent in the corresponding imaginary scenario that results from imagining δ.

The Imagination Algorithm has been formally defined as the `ImgAlg`, and it can now be executed to expand a model $\mathcal{M}$ into a model $\mathcal{M}^+$, which includes a set of new imaginary possible worlds that were not there before, and that result from the agent performing an act of imagination with an initial premise δ.

As it can be seen in the previous specification of the `ImgAlg`, the only restriction we put on the content of the act of imagination is that it can be expressed in the propositional fragment of our logic (that is, we require it to belong to FORM*), and we do not allow our agent to imagine contradictory premises. Aside from that, our `ImgAlg` provides the required mechanisms to allow the agent to imagine whatever she wants to, and expands the model in consequence by creating new imaginary possible worlds.

4.5 Semantics

Similarly to what we have seen in Section 3.1.2, we evaluate a formula φ of the Logic of Imaginary Scenarios at a world $w \in W$ of a model $\mathcal{M}$. We use symbol $\vDash$, which we call *local consequence*, and we write $\mathcal{M}, w \vDash \varphi$ to express that φ is true at w in model $\mathcal{M}$; conversely, we write $\mathcal{M}, w \nvDash \varphi$ to express that φ is not true at w in model $\mathcal{M}$. Furthermore, we write $\vDash \varphi$ to express that φ is true at every world of every model; i.e., φ

[5]Recall that, as δ has been converted to DNF, δ_i is a clause of DNF(δ), which means that δ_i is a conjunction of literals of the form $l_1 \wedge \ldots \wedge l_m$; each one of these literals l_j can either be a positive literal, meaning that it is an atom $p \in$ ATOM, or a negative literal, meaning that it is the negation $\neg p$ of an atom $p \in$ ATOM.

is a *validity*. Besides, we write $\Gamma \vDash \varphi$ (for Γ being a set of formulas) if, for every model $\mathcal{M}$ and world w such that $\mathcal{M}, w \vDash \Gamma$ (that is: every formula in Γ is true at world w of $\mathcal{M}$), it is the case that $\mathcal{M}, w \vDash \varphi$. In other words: any model satisfying the set of formulas Γ would also satisfy formula φ. In this case, we say that φ is a *semantic consequence* of Γ.

$\mathcal{M}, w \vDash i$	iff	$N(i) = w$ and, for every $v \in W$, if $\mathcal{M}, v \vDash i$, then $v = w$
$\mathcal{M}, w \vDash p$	iff	$w \in V(p)$
$\mathcal{M}, w \vDash \neg\varphi$	iff	$\mathcal{M}, w \nvDash \varphi$
$\mathcal{M}, w \vDash \varphi \wedge \psi$	iff	$\mathcal{M}, w \vDash \varphi$ and $\mathcal{M}, w \vDash \psi$
$\mathcal{M}, w \vDash \varphi \vee \psi$	iff	$\mathcal{M}, w \vDash \varphi$ or $\mathcal{M}, w \vDash \psi$
$\mathcal{M}, w \vDash \varphi \rightarrow \psi$	iff	$\mathcal{M}, w \vDash \neg\varphi$ or $\mathcal{M}, w \vDash \psi$
$\mathcal{M}, w \vDash K\psi$	iff	for every world $v \in W$ such that $(w, v) \in R_K$, it is the case that $\mathcal{M}, v \vDash \varphi$
$\mathcal{M}, w \vDash @_i\varphi$	iff	there exists a world $v \in W$ such that $N(i) = v$ and $\mathcal{M}, v \vDash \varphi$
$\mathcal{M}, w \vDash \mathrm{Img}(\delta)$	iff	δ is not contradictory ($\delta \not\equiv \bot$) and either there already exists $v \in W$ such that $(w, v, \delta) \in R_{Img}$ or, after executing $\texttt{ImgAlg}(\delta, w)$, $\mathcal{M}$ is expanded into $\mathcal{M}^+$
$\mathcal{M}, w \vDash \langle I(\delta)\rangle\varphi$	iff	$\mathcal{M}, w \vDash \mathrm{Img}(\delta)$ and there is some $v \in W^+$ such that $(w, v, \delta) \in R^+_{Img}$ and $\mathcal{M}^+, v \vDash \varphi$ and $\mathcal{M}^+, v \vDash \delta$

Recall that, during an execution of the `ImgAlg` in a model $\mathcal{M}$, we define the expanded model $\mathcal{M}^+$; elements W^+ and R^+_{Img} belong to this expanded model.

It is worth stressing the fact that, both during and just after each new execution of the `ImgAlg`, we distinguish the model used at the beginning of the execution (that is, the model before getting expanded) as $\mathcal{M}$, whereas we refer to the model that result from such execution (the expanded model) as $\mathcal{M}^+$; therefore, elements $\mathcal{M}^+$, W^+ and R^+_{Img} in the satisfiability definition of operator $\langle I(\delta)\rangle$ are meant to be understood as the expanded model corresponding to the execution of the `ImgAlg` related to the evaluation $\mathcal{M}, w \vDash \mathrm{Img}(\delta)$, whether it is the last expansion of the model that has been done or not. However, whenever we perform a new execution of the `ImgAlg`, we refer to the current state of the model as $\mathcal{M}$, no matter how many times it has been expanded[6], and to the

[6]This practice is similar to the way variables are usually handled in programming languages. For instance, is we had a certain object *Obj* in our program, and we wanted to expand such object by, say, adding a new element a to it, we could directly overwrite *Obj* with its expanded version by writing something such as $Obj := Obj + a$ (abusing notation). What we do with $\mathcal{M}$ and $\mathcal{M}^+$ is similar, although

model resulting from that particular execution as $\mathcal{M}^+$; intuitively, as soon as we need to expand the model again, we use our current model as the "unexpanded" model for that particular new call to `ImgAlg`.

It is worth devoting a few lines to clarifying how the dynamic operator $\text{Img}(\delta)$ works. As we have already explained, this operator has the particularity of representing a voluntary action, performed by the agent, to imagine something (δ, specifically). The aim of this operator, therefore, is to call the `ImgAlg` with parameters δ and w (being w the world where the formula is evaluated and, thus, the world the agent takes as the reference to carry out such act of imagination). Providing satisfiability conditions for this operator, then, is basically a technical requirement, and thus it has little to do with the operator's main goal: this operator is *not* intended to be used in order to check whether it is true or false, but rather it is intended to call a procedure that changes the current model into its expanded version.

Regarding the satisfiability conditions of the $\text{Img}(\delta)$ operator, note how they are, in fact, quite trivial. The reason why we impose the condition of δ not being contradictory is two-folded. Firstly, because our philosophical approach to imagination does not allow to imagine contradictory premises (that is, formulas from which a contradiction can be derived: $\delta \equiv \bot$); we do not believe that we can actually do that, and so we do not want to allow our formal system to do so. Secondly, because our formal logic is classical: in classical logic, consistency is required in order to avoid triviality. In particular, the following axiom of classical logic, $\varphi \wedge \neg\varphi \to \psi$, states that, from a contradictory pair of formulas, it is possible to derive any other formula; obviously, we do not want this to happen in our system. The rest of satisfiability conditions for operator $\text{Img}(\delta)$ are trivial by the way the `ImgAlg` itself works. If δ is not contradictory, then either there will already exist an imagination relation R_{Img} leading to an imaginary world, and meaning that the agent has already performed the act of imagining δ at the world of evaluation, or the `ImgAlg` will be executable, and thus able to expand a model $\mathcal{M}$ into model $\mathcal{M}^+$.

So far, the Logic of Imaginary Scenarios provides the appropriate framework to account for our goals and desiderata. In order to get a better overview of the logic, let's follow a detailed example showing how the `ImgAlg`, which is the key part of our system, works.

we sometimes distinguish between both by using these different terms, when required.

4.6 An Example of an Act of Imagination

In the present section, we provide a full example detailing how an execution of the `ImgAlg` works, given an initial Model for Imaginary Scenarios. Before any execution of the `ImgAlg` (or, in other words, before performing any act of imagination), a Model for Imaginary Scenarios looks like a standard, single-agent epistemic model. Particularly, the imagination relation R_{Img} is empty, as no act of imagination is represented in there yet.

Let's take, as the basis of our example, the initial Model for Imaginary Scenarios represented in Figure 4.5; in this model, the agent knows r, but does not know whether p or q holds. In fact, the agent knows that, if p is the case, then q is not, and the other way around. The model contains two possible worlds w and v, connected by the indistinguishability relation R_K (which is reflexive, transitive and symmetric, as required), and the valuation function assigning the corresponding truth-values to the propositional variables, according to possible worlds. For the sake of clarity we also write, in the possible worlds represented in Figure 4.5, the negations of the atomic formulas being false in there.

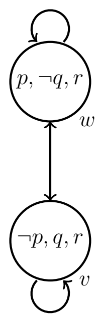


Figure 4.5: The initial model, before performing any act of imagination.

Now, our agent decides to perform an act of imagination with an initial premise $\delta = (\neg p \vee \neg r) \rightarrow ((q \wedge \neg r) \vee \neg q)$. Note that, although one of the existing possible worlds already satisfies this formula (specifically world v, by making p false in it), performing such act of imagination will create possibly many different possible worlds in which the formula $(\neg p \vee \neg r) \rightarrow ((q \wedge \neg r) \vee \neg q)$ holds, even if they are not already considered as epistemically possible by the agent. Let's see how the `ImgAlg` works by following each one of the steps defined in Section 4.4, and responding to the following "evaluation" of the dynamic formula of our language:

$$\mathcal{M}, w \vDash \mathrm{Img}((\neg p \vee \neg r) \rightarrow ((q \wedge \neg r) \vee \neg q))$$

1. The `ImgAlg` starts being called with an initial premise $\delta = (\neg p \vee \neg r) \rightarrow ((q \wedge \neg r) \vee \neg q)$ and a world of reference $w^R = w$. As $(\neg p \vee \neg r) \rightarrow ((q \wedge \neg r) \vee \neg q)$ is not a contradictory formula, the `ImgAlg` continues with its execution.

2. The DNF of $(\neg p \vee \neg r) \rightarrow ((q \wedge \neg r) \vee \neg q)$ is computed as follows:

Initial formula:	$\Rightarrow$	$(\neg p \vee \neg r) \rightarrow ((q \wedge \neg r) \vee \neg q)$
Eliminate implications:	$\Rightarrow$	$\neg(\neg p \vee \neg r) \vee ((q \wedge \neg r) \vee \neg q)$
Move negations inwards:	$\Rightarrow$	$(\neg\neg p \wedge \neg\neg r) \vee ((q \wedge \neg r) \vee \neg q)$
Replace double negations:	$\Rightarrow$	$(p \wedge r) \vee ((q \wedge \neg r) \vee \neg q)$
Formula is in DNF:	$\Rightarrow$	$(p \wedge r) \vee (q \wedge \neg r) \vee \neg q$

3. Now, for each clause in $(p \wedge r) \vee (q \wedge \neg r) \vee \neg q$, the algorithm must create one new possible imaginary world: in particular, as there are three different clauses in here (namely: $(p \wedge r), (q \wedge \neg r)$ and $\neg q$), the algorithm must create three new possible worlds, to which will refer as w_1, w_2 and w_3. This defines the expanded set of possible worlds in the following way:

$$W^+ = W \cup \{w_1, w_2, w_3\}$$

After performing this step, the model (which is still being expanded, so it cannot be considered $\mathcal{M}^+$ yet) looks like the one in Figure 4.6:

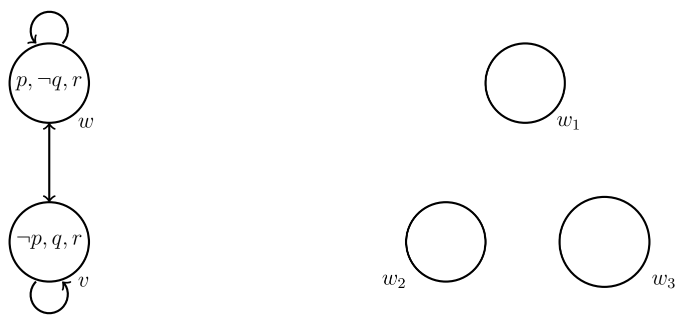


Figure 4.6: Creating new possible worlds for each clause in DNF(δ).

4. The algorithm must now create the new imaginary relations between the world of reference (in this case, w) and the new imaginary worlds (w_1, w_2 and w_3), which will represent different possible states of affairs accounting for the initial premise

$(\neg p \vee \neg r) \rightarrow ((q \wedge \neg r) \vee \neg q)$. This defines the expanded set of imaginary relations as follows:

$$\begin{aligned} R^{+}_{Img} = R_{Img} \cup \{ \quad & (w, w_1, (\neg p \vee \neg r) \rightarrow ((q \wedge \neg r) \vee \neg q)), \\ & (w, w_2, (\neg p \vee \neg r) \rightarrow ((q \wedge \neg r) \vee \neg q)), \\ & (w, w_3, (\neg p \vee \neg r) \rightarrow ((q \wedge \neg r) \vee \neg q)) \ \} \end{aligned}$$

At this point, the model looks as in Figure 4.7 (for the sake of clarity, we use δ to label the imaginary relations, instead of the whole formula).

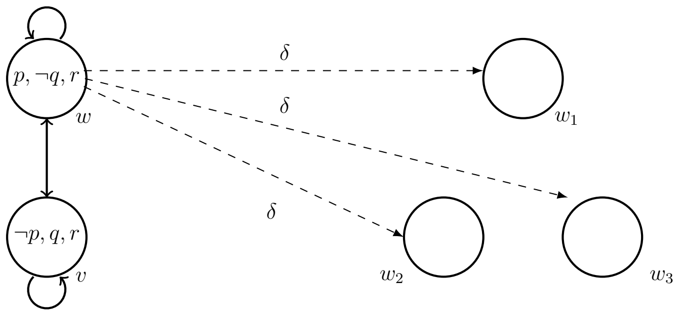


Figure 4.7: The imagination relation shows which worlds result from imagining δ.

5. When imagining the scenario initiated by the premise δ, we have already seen how it can be accounted by three different imaginary worlds. Therefore, there are three different states of affairs which are epistemically indistinguishable by the agent, when she imagines δ. In order to represent this, the `ImgAlg` must relate every new imaginary possible world through the epistemic indistinguishability relation, thus defining its expanded version:

$$\begin{aligned} R^{+}_{K} = R_{K} \cup \{ \quad & (w_1, w_1), (w_1, w_2), (w_1, w_3), (w_2, w_1), (w_2, w_2), \\ & (w_2, w_3), (w_3, w_1), (w_3, w_2), (w_3, w_3) \ \} \end{aligned}$$

The model, at this point, corresponds to the one in Figure 4.8.

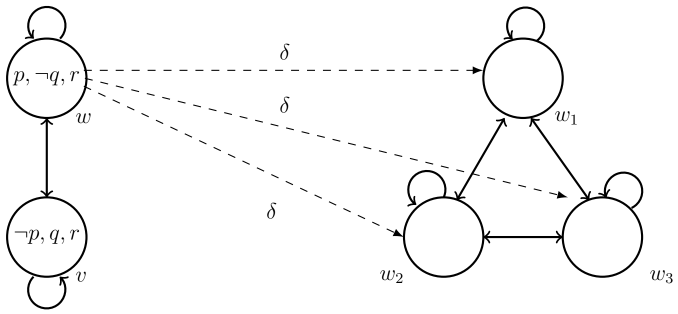


Figure 4.8: Any of the new imaginary worlds represents a world satisfying δ.

6. Once the structure regarding the new imaginary possible worlds has been computed, the `ImgAlg` must assign a name to those new possible imaginary worlds by adding their corresponding new nominals (in this case, we add nominals k_1, k_2 and k_3), and expanding the nominal function:

$$\mathrm{NOM}^+ = \mathrm{NOM} \cup \{k_1, k_2, k_3\}$$

$$N^+ = N \cup \{(k_1, w_1), (k_2, w_2), (k_3, w_3)\}$$

6. Now, the `ImgAlg` must determine which atomic propositions hold at each new imaginary possible worlds. This must be done both according to the initial premise δ, which has preference when creating the new imaginary worlds, and the atomic propositions that can be imported from the world of reference w^R.

   (a) By computing DNF(δ), we already have three possible alternatives that make δ true; namely, $(p \wedge r), (q \wedge \neg r)$ and $\neg q$. Each one of these formulas will be used to determine which atoms are true and which atoms are false in one of the new imaginary possible worlds. This, in turn, determines the expanded valuation function as follows (note that we only have to add the positive atoms to the valuation function; therefore, nothing needs to be expanded regarding world w_3, to which we assign formula $\neg q$):

$$V_1^+(p) = V(p) \cup \{w_1\}$$

$$V_1^+(r) = V(r) \cup \{w_1\}$$

$$V_1^+(q) = V(q) \cup \{w_2\}$$

The model, at this point, looks as in Figure 4.9; note that, for the sake of clarity, we also indicate which negated atomic propositions are related to each new imaginary possible world, although they do not strictly appear in the expanded valuation function V^+.

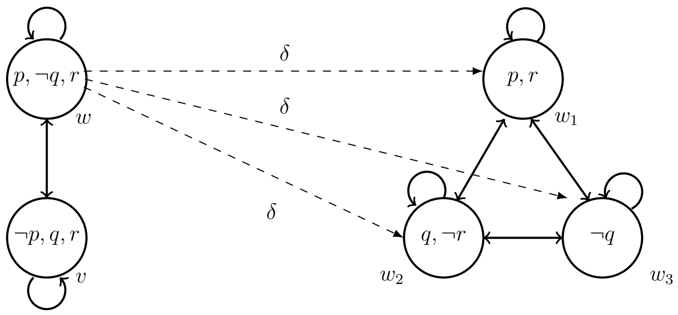


Figure 4.9: Each new imaginary world represents one of the alternatives satisfying δ.

(b) Once the corresponding atomic propositions have been clamped to each new possible world, the `ImgAlg` must import any atomic proposition holding at the world of reference w^R (which is w in this example), as long as they do *not* appear in the corresponding clause from the DNF(δ) that was used to determine the valuation of each new possible world. We require this because the atomic propositions occurring within the initial premise δ have priority over the atomic propositions that the agent imports from the world of reference. This, therefore, defines the last step in the expansion of the valuation function as follows:

$$V^+(p) = V_1^+(p) \cup \{w_2, w_3\}$$

$$V^+(r) = V_1^+(r) \cup \{w_3\}$$

Note that the `ImgAlg` only considers, for each new imaginary possible world w_1, w_2 or w_3, the atomic propositions that do not appear in the clause that is used in the previous step. Specifically, as world w_1 is created using the clause $(p \wedge r)$, we only have to consider q; however, q does not hold in w, so it should not be imported; world w_2 is created using the clause $(q \wedge \neg r)$, so only p should be considered; as p is true in w, it is imported into w_2; similarly, as world w_3 is created using the clause $\neg q$, both p and r should be considered; as they are both true in w, both are imported into w_3. Figure 4.10 represents the model after this step, which corresponds to the already finished expanded model $\mathcal{M}^+$;

in this figure, though, we include the negated atoms for the sake of clarity.

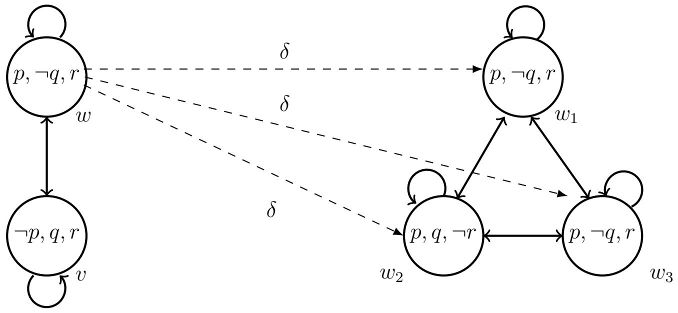


Figure 4.10: The `ImgAlg` finishes after importing atoms from the world of reference.

7. The `ImgAlg` has finished its execution. Figure 4.10 already represents the expanded model $\mathcal{M}^+$, which results from executing an act of imagination with an initial premise $\delta = (\neg p \vee \neg r) \rightarrow ((q \wedge \neg r) \vee \neg q)$ at a world of reference $w_R = w$ in a Model for Imaginary Scenarios $\mathcal{M}$. Recalling what we introduced in Section 4.3.1, this execution of the `ImgAlg` has taken the real world w as the world of reference, and, through imagining δ, has defined a new imaginary scenario containing three new imaginary worlds w_1, w_2 and w_3.

4.6.1 About Equivalent Imaginary Worlds

In the previous example it can be seen how, in this particular case, the act of imagination performed by the agent leads to the creation of three new imaginary worlds, two of which end up being equivalent: w_1 and w_3. Due to the fact that formal logics usually try to be as optimal as possible, one might argue that, being equivalent, one of such worlds could be removed. It would be possible to add one further step to the `ImgAlg` to check, once all worlds have been created, whether there are any of the new imaginary worlds that are equivalent, and remove one of them if that was the case.

Nevertheless, we do not want to do such thing. Even though certain imaginary worlds may end up being equivalent (such as w_1 and w_3 in the example), those worlds will have been defined by following different conditions. In particular, each clause of the DNF(δ) would lead to a different set of atomic propositions being true at each new imaginary world, and thus they would represent, before importing any atomic proposition from the world of reference, a different state of affairs. Take, for instance, the three different clauses in the previous example: $(p \wedge r), (q \wedge \neg r)$ and $\neg q$. Each one of these clauses represents a different state of affairs satisfying δ, and each one is characterized by requiring a different set (or conjunction) of propositional atoms to hold in there. What happens, then, with the rest of atomic propositions not appearing in each one of the previous clauses? Simply, that they do *not* matter, regarding the satisfaction of δ. We decide to import any other unspecified atomic proposition because we are defining our logic as being classical and, in this setting, truth-value gaps are not allowed.

However, if two or more new imaginary worlds end up being equivalent, this would be because the atomic propositions set while satisfying δ and the atomic propositions imported from the world of reference are "accidentally" equivalent. Imaginary worlds that end up being equivalent are not *meant* to be equivalent, but they are just as a coincidence following the particular case being modeled; imagining δ by taking a particular w_1^R may lead to the appearance of equivalent imaginary worlds, but imagining δ again using a different w_2^R could lead to no worlds being equivalent at all.

Precisely because we are focusing, in our proposal, in the *process* of creating those imaginary worlds, we want to "interfere" in the whole process as little as possible, in terms of optimizing its results. When people engage in certain imagination acts, they may end up realizing that some imaginary worlds do actually depict the same state of affairs; moreover, it may be even arguable that, in those case, most people do actually collapse both worlds into a single one. However, this *act of realization* of the agent becoming aware that two imaginary worlds are actually equivalent involve some sort of internal reasoning or comparison between the two worlds, prior to acknowledging that they are equivalent, and prior to collapsing them both into a single one. As our system cannot account for this internal reasoning step in an explicit way, we do *not* want it to happen, as it may end up hiding information to us, as modelers; if, while considering a specific imagination act, it is the case that we end up having equivalent worlds, we as modelers want to know so, and thus we do not want our `ImgAlg` to behave as a sort of "black box" that deletes those worlds as part of its execution without letting us know.

4.7 Exploring Imagination Validities

Up until now, we have defined the language, semantics and the algorithm responsible for handling the way voluntary acts of imagination create new imaginary worlds. In this section, we show how our Logic of Imaginary Scenarios can express some interesting features of imagination using formulas. We discuss what the interpretation of such formulas is, and we prove that they are semantically valid, with respect to the semantics of our language. Let's start by trying to capture our desired features of voluntary imagination acts by using formulas of the Logic of Imaginary Scenarios. What do we want to capture, using our logic? Which properties do we want to represent?

For instance, we want our logic to guarantee that what the agent is explicitly imagining will always be the case in the imaginary worlds that result from such imagination act; in other words, we want to make sure that the initial premise is "clamped". This means that, if the agent imagines δ, then δ must hold at the new imaginary worlds. We refer to this property as *Clamped Premise*, and we can express it through the following formula, which is valid in our system:

$$\vDash @_i\langle I(\delta)\rangle j \rightarrow @_j\delta \tag{4.1}$$

Proof. In order to prove the validity of $\vDash @_i\langle I(\delta)\rangle j \rightarrow @_j\delta$, take an arbitrary model $\mathcal{M}$, an arbitrary world $w \in W$, and suppose $\mathcal{M}, w \vDash @_i\langle I(\delta)\rangle j$. We will prove that the formula is valid by contradiction, by supposing also that $\mathcal{M}, w \vDash \neg @_j\delta$. By the definition of $N(i)$, we know that there is a unique world v such that $N(i) = v$; by the definition of @, we know then that $\mathcal{M}, v \vDash \langle I(\delta)\rangle j$. By the definition of operator $\langle I(\delta)\rangle$, we know there must be a world u such that both $(v, u, \delta) \in R_{Img}$ (recall that relation R_{Img} has been expanded into R^+_{Img}, at some point, through executing the `ImgAlg`; however, and as we explain in Section 4.5, once the expansion is completed we often refer to the expanded elements by their usual names R_{Img} and such; in the following proofs, we follow the same practice) and also $\mathcal{M}, u \vDash j$, according to formula $\langle I(\delta)\rangle j$ and by the definition of satisfiability for nominals. Again, by the definition of operator $\langle I(\delta)\rangle$, we know that $\mathcal{M}, u \vDash \delta$. As $\mathcal{M}, u \vDash j$, we know, by the definition of operator @, that $\mathcal{M}, u \vDash @_j\delta$. Thus, we also get that $\mathcal{M}, w \vDash @_j\delta$, which contradicts our supposition that $\mathcal{M}, w \vDash \neg @_j\delta$, thus proving the validity $\vDash @_i\langle I(\delta)\rangle j \rightarrow @_j\delta$. □

Regarding our static imagination operator $\langle I(\delta)\rangle$, we want to capture the fact that, if the agent performs an act of imagination with an initial premise δ, and this act of imagination leads to the creation of a certain imaginary world, then whatever holds at that imaginary world "follows" from this specific act of imagination. This roughly corresponds to the definition of the static operator itself (as seen in Section 4.5), and it can be captured by the following valid formula, which we call *Imagination Bridge*: intuitively, it expresses that, if we can access a world j via a specific act of imagination, then anything that holds in j can be said to follow from that act of imagination itself:

$$\vDash @_i\langle I(\delta)\rangle j \wedge @_j\varphi \rightarrow @_i\langle I(\delta)\rangle\varphi \tag{4.2}$$

Proof. In order to prove $\vDash @_i\langle I(\delta)\rangle j \wedge @_j\varphi \rightarrow @_i\langle I(\delta)\rangle\varphi$, take an arbitrary model $\mathcal{M}$, an arbitrary world $w \in W$, and suppose $\mathcal{M}, w \vDash @_i\langle I(\delta)\rangle j \wedge @_j\varphi$. We will prove the validity by contradiction, by supposing also that $\mathcal{M}, w \vDash \neg(@_i\langle I(\delta)\rangle\varphi)$. First, by the definition of $\wedge$, we know that we have both $\mathcal{M}, w \vDash @_i\langle I(\delta)\rangle j$ and $\mathcal{M}, w \vDash @_j\varphi$. Now, by the definition of $N(i)$, we know that there is a world v such that $N(i) = v$; furthemore, by the definition of @, we know that $\mathcal{M}, v \vDash \langle I(\delta)\rangle j$. By the definition of operator $\langle I(\delta)\rangle$, we know there is a world u such that $(v, u, \delta) \in R_{Img}$ and also such that $\mathcal{M}, u \vDash j$, according to formula $\langle I(\delta)\rangle j$. Now, since $N(j) = u$, by the definition of @, and recalling that we have $\mathcal{M}, w \vDash @_j\varphi$, we get that $\mathcal{M}, u \vDash \varphi$. Now, as we know that $(v, u, \delta) \in R_{Img}$ we have, by the definition of $\langle I(\delta)\rangle$, that $\mathcal{M}, v \vDash \langle I(\delta)\rangle\varphi$, and considering that we also have $\mathcal{M}, v \vDash i$, and by the definition of operator @, we have that $\mathcal{M}, v \vDash @_i\langle I(\delta)\rangle\varphi$. Therefore, we also get $\mathcal{M}, w \vDash @_i\langle I(\delta)\rangle\varphi$, which contradicts our supposition that $\mathcal{M}, w \vDash \neg(@_i\langle I(\delta)\rangle\varphi)$, therefore proving the validity we wanted. $\square$

There is another rather interesting property, regarding what happens in the imaginary worlds created as a result of an act of imagination, that our logic can capture. By the way our `ImgAlg` works when creating new imaginary scenarios, we know that an imaginary world created as a result of a certain imagination act can only be the epistemic alternative of another imaginary world created by the *same* imagination act; moreover, these possible worlds should have been created by taking the same possible world as its world of reference. This property captures the fact that the agent can tell whether two imaginary worlds have been created by the same imagination act, and thus belong to the same imaginary scenario. Besides being able to tell the difference between what is real and what is imaginary, the agent should also be able to tell the difference between distinct acts of imagination: we do not want our agent to mix the imaginary worlds concerning an imaginary scenario about a tea party, and the ones about a dragon-hunting fantasy story! Therefore, we want to express the fact that, if a certain possible world results from a specific act of imagination, and there is another possible world which is epistemically accessible from the former, then the latter results from the same act of imagination as well. This corresponds to the following validity, which we call *Imaginary Possibilities*:

$$\vDash @_i\langle I(\delta)\rangle j \wedge @_jMk \rightarrow @_i\langle I(\delta)\rangle k \tag{4.3}$$

Proof. In order to prove the validity of $\vDash @_i\langle I(\delta)\rangle j \wedge @_j Mk \to @_i\langle I(\delta)\rangle k$, take an arbitrary model $\mathcal{M}$, an arbitrary world $w \in W$, and suppose $\mathcal{M}, w \vDash @_i\langle I(\delta)\rangle j \wedge @_j Mk$. We will prove the validity by contradiction, by supposing also that $\mathcal{M}, w \vDash \neg(@_i\langle I(\delta)\rangle k)$. By the definition of $\wedge$, we know that we have $\mathcal{M}, w \vDash @_i\langle I(\delta)\rangle j$ and also $\mathcal{M}, w \vDash @_j Mk$. Now, by the definition of $N(i)$, we know that there is a world v such that $N(i) = v$; furthemore, by the definition of @, we know that $\mathcal{M}, v \vDash \langle I(\delta)\rangle j$. By the definition of $\langle I(\delta)\rangle$, this means that there exists a world u such that $(v, u, \delta) \in R_{Img}$ and $\mathcal{M}, u \vDash j$. By the definition of @, we get $\mathcal{M}, u \vDash Mk$, which means that there exists a world z such that $(u, z) \in R_K$ and $\mathcal{M}, z \vDash k$. By the way our model works, we know that an element in R_{Img} must have been created by an execution of the `ImgAlg`; in particular, and as we know that $(v, u, \delta) \in R_{Img}$, we know that there has been a call `ImgAlg`(δ, v) that has created, possibly among others, a new R_{Img}-relation from v to this world u. Now, by the way indistinguishability relations are created between the new imaginary worlds in the `ImgAlg`, we know that if $\mathcal{M}, u \vDash Mk$, then the world identified by k (to which we already assigned the world z) must have been created by this execution of `ImgAlg` (otherwise, there is no way by which a world created in an execution of the `ImgAlg` can be related via R_K to any other world in the model). Then, as world z has been created by `ImgAlg`, then we know that this world must be also related to the world of reference v via the imagination relation: that is, we know that $(v, z, \delta) \in R_{Img}$. By the definition of $\langle I(\delta)\rangle$, we get that $\mathcal{M}, v \vDash \langle I(\delta)\rangle k$, as $\mathcal{M}, z \vDash k$. As we know that $\mathcal{M}, v \vDash j$, and by the definition of @, we get $\mathcal{M}, w \vDash @_j\langle I(\delta)\rangle k$, which contradicts our assumption, and proves the previous validity. □

Last, but not least, there is one of the main points of our desiderata which we want our system to account for: namely, we want acts of imagination to be both dynamic and voluntary. The initial model is formed by a single-agent epistemic model, and it is only through executing the `ImgAlg` that new imaginary scenarios are created as a result of acts of imagination. Therefore, if there *exists* an imagination relation within the model, then it *must* be because a dynamic act of imagination has taken place at some point within that model. In other words: behind every static imagination formula of the form $\langle I(\delta)\rangle\varphi$, there must be an explicit execution of an act of imagination, which has been called by the dynamic imagination formula $\mathrm{Img}(\delta)$. The following validity accounts for this property, which we call *Voluntary Imagination*:

$$\vDash @_i\langle I(\delta)\rangle\varphi \to @_i\mathrm{Img}(\delta) \tag{4.4}$$

Proof. In order to prove the validity of $\vDash @_i\langle I(\delta)\rangle\varphi \to @_i\text{Img}(\delta)$, take an arbitrary model $\mathcal{M}$, an arbitrary world $w \in W$, and suppose $\mathcal{M}, w \vDash @_i\langle I(\delta)\rangle\varphi$. We will prove this validity by contradiction, by supposing also that $\mathcal{M}, w \vDash \neg @_i\text{Img}(\delta)$. Now, by the definition of $N(i)$, we know that there is a world v such that $N(i) = v$; furthemore, by the definition of @, we know that $\mathcal{M}, v \vDash \langle I(\delta)\rangle\varphi$. By the definition of $\langle I(\delta)\rangle$, we know that there exists a world u such that $(v, u, \delta) \in R_{Img}$ and, furthermore, that $\mathcal{M}, u \vDash \delta$. By the definition of R_{Img} and $\text{Img}(\delta)$, we know that any imaginary world (that is, any world being accessible from another one through the R_{Img} relation) has been created by an execution of the `ImgAlg`, which can be only called, in our logic, by formula $\text{Img}(\delta)$. As we know that such relation $(v, u, \delta) \in R_{Img}$ exists, then there must have been a call to `ImgAlg`, made from world v and using a premise δ; thus, there must have been a call `ImgAlg`(δ, v). By the definition of $\text{Img}(\delta)$, this means that $\mathcal{M}, v \vDash \text{Img}(\delta)$ holds, which, as we also know that $\mathcal{M}, v \vDash i$, and by the definition of operator @, we get $\mathcal{M}, v \vDash @_i\text{Img}(\delta)$; as the world of evaluation for @ can be swapped to w, this contradicts our assumption and therefore proves the validity we wanted to. □

Aside from these main features we wanted to represent with our logic, there are other formal properties we want to consider, regarding the way our logic represents that acts of imagination behave. For instance, we want to study how the static operator of imagination $\langle I(\delta)\rangle\varphi$ works, with respect to other logical connectives. In particular, we want to see how this operator behaves regarding conjunction and disjunction, both when dealing with the "result" of the act of imagination (that is, the φ of the previous formula) and its initial premise (that is, the δ itself).

Let's see how the logic behaves regarding conjunction. With respect to the evaluation of formulas following an act of imagination, if the agent creates an imaginary scenario with premise δ, and this leads to an imaginary world satisfying $\varphi \wedge \psi$, then we also want to ensure that such imagination act leads to φ, and that it also leads to ψ (separately). Our logic does account for this, as the following validity shows[7]:

$$\vDash \langle I(\delta)\rangle(\varphi \wedge \psi) \to (\langle I(\delta)\rangle\varphi) \wedge (\langle I(\delta)\rangle\psi) \tag{4.5}$$

Proof. By following the definition of satisfiability of operator $\langle I(\delta)\rangle$, we know that there exists an R_{Img}-accessible world (call it v) in which formula $\varphi \wedge \psi$ hold. By the definition

[7]The converse, nevertheless, is not valid in our system. There could be a δ-accessible world v satisfying φ and another δ-accessible world u satisfying ψ, but this would not amount to their conjunction being simultaneously satisfiable.

of operator $\wedge$, this means that both φ and ψ are true at v. Due to this, and again by the definition of $\langle I(\delta)\rangle$, we know that both $\langle I(\delta)\rangle\varphi$ and $\langle I(\delta)\rangle\psi$ must be true; then, and by the definition of $\wedge$, we can say that $\langle I(\delta)\rangle\varphi) \wedge (\langle I(\delta)\rangle\psi$ is also true. $\square$

Regarding the content of the initial premise, things get more delicate. Do we want our logic to be monotonic[8], with respect to the initial premise of an act of imagination? In other words: if I imagine δ and this leads to a world satisfying φ, should it also be the case that, by imagining $\delta \wedge \gamma$, I can also reach a world satisfying φ?

Intuitively, we should not want this to happen. Take, for instance, an example in which I imagine that I am at a tea-party; when elaborating the scenario based on reality-oriented facts, I would probably import the way I am dressed at that moment (say, black trousers and an informal T-shirt), and so it will be true, in that imaginary scenario, that I am dressed that way. However, if now I toss out this first imagining and I imagine that I am at a tea-party *and* that I am wearing a tuxedo, my knowledge about me being dressed with trousers and a T-shirt will be blocked, and therefore not imported. Thus, in this second imaginary scenario, it will be false that I am wearing trousers and a T-shirt. According to this, we do *not* want our logic to be monotonic with respect to the content of the initial premise determining an imagining.

First and foremost, it is worth noting that, in our logic, two premises δ_1 and δ_2 will result, by the way our `ImgAlg` behaves, in two different acts of imagination. The scenario in which I just imagine being at a tea-party, and the one in which I imagine being in a tea-party and wearing a tuxedo are different imaginings[9]. Therefore, formulas $\langle I(\delta_1)\rangle\varphi$ and $\langle I(\delta_2)\rangle\varphi$ would be about two different imagination relations, unless δ_1 and δ_2 are exactly the *same* formula[10].

Second, and by sticking to the informal definition of the Imagination Algorithm dis-

[8]In a nutshell, a logic is considered *monotonic* whenever the conclusions derived from a certain inference cannot be altered by adding new information, whereas it is considered *non-monotonic* whenever, by adding new premises into an inference, new consequences could be withdrawn. Non-monotonic logic refers to a family of formal frameworks devised to represent *defeasible inference*; that is, the kind of inference that can vary its conclusions, whenever new information is added in their premises. For a comprehensive introduction to this kind of logics, see [48].

[9]Note that imagining just that I am at a tea-party and *then* adding a new premise saying that I wear a tuxedo is not the same act of imagination as imagining both simultaneously. Although the resulting imaginary scenarios could end up being equivalent, the first case requires two different acts of imagination, whereas the second one only requires one. Besides, the former has an additional, intermediate imaginary scenario describing a certain state of affairs, which the second does not.

[10]Logically equivalent formulas are *not* the same formulas in terms of the `ImgAlg`. This, however, is not to be seen as a drawback of our system; quite the contrary, it captures the fact that the agent cannot check whether two initial premises lead to the same imaginary scenario, until she actually creates both.

cussed in Section 4.1, we want to import as much information about the world of reference as possible: therefore, if by initiating an act of imagination with content $\delta \wedge \gamma$, it turns out that this γ blocks a certain atomic formula that would not have been blocked by δ alone, then the resulting imaginary scenario would be different from the one that would result by imagining only δ.

Our system, then, is non-monotonic with respect to the static imagination, and this is shown through the following invalidity:

$$\not\models \langle I(\delta)\rangle\varphi \rightarrow \langle I(\delta \wedge \gamma)\rangle\varphi \tag{4.6}$$

Proof. Consider the Model for Imaginary Scenarios represented in Figure 4.11. The model represents a real possible world w in which p and q are the case, an imaginary world w_1 created as a result of an imagination act with initial premise p, and which fulfills p and q, and another world w_2, created as a result of an act of imagination with content $p \wedge \neg q$, and which fulfills p and $\neg q$ (we write $\neg q$ inside w_2 within the picture for reading purposes). This model represents two acts of imagination: one with content p, and one with content $p \wedge \neg q$. If we interpret, in our previous invalidity, $\delta = p$, $\gamma = \neg q$ and $\varphi = q$, we see how $\models \langle I(p)\rangle q \rightarrow \langle I(p \wedge \neg q)\rangle q$ is not valid in this model. Particularly, this is due to the fact that formulas already clamped within the imaginary world may block atomic propositions that would be otherwise imported from the world of reference. In this case, clamping $\neg q$ in world w_2 prevents q from being imported. $\square$

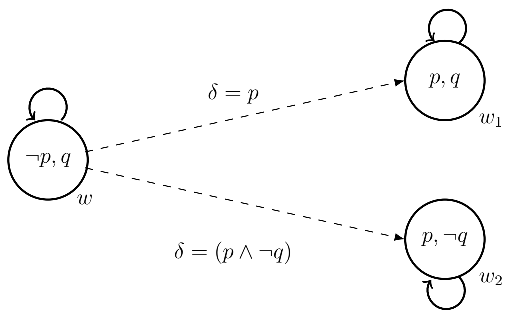


Figure 4.11: A countermodel for formula 4.6.

Regarding disjunction, our logic behaves classically with respect to the formulas that follow from an act of imagination; that is, if the agent imagines δ, and this leads to a possible world satisfying φ, then it is also true that imagining δ leads to a possible world satisfying $\varphi \vee \psi$, as our operator $\vee$ is defined in a classical way. Therefore, the following validity holds in our proposal:

$$\models \langle I(\delta)\rangle\varphi \rightarrow \langle I(\delta)\rangle(\varphi \vee \psi) \tag{4.7}$$

Proof. By the definition of $\langle I(\delta)\rangle$, we know that there exists an R_{Img}-accessible world (call it v) at which φ is true. By the definition of operator $\vee$, a formula $\varphi \vee \psi$ is true at a possible world whenever at least one of its subformulas is true in there, which is the case. As $\varphi \vee \psi$ holds at v, then, again by the definition of $\langle I(\delta)\rangle$, formula $\langle I(\delta)\rangle(\varphi \vee \psi)$ is true at the world of evaluation. □

Again, and similarly as in the conjunction case, things get more complicated when we focus on the content of the initial premise. Imagining δ does *not* lead to the same imaginary scenario that would be described by imagining $\delta \vee \gamma$: the former clearly requires δ to hold (as a result of being clamped) in the resulting possible worlds, whereas the later describes different possible scenarios in which *either* δ or γ may hold, but not necessarily any of these two, and so not necessarily δ (by the definition of operator $\vee$). We want our system to capture this fact, and so we want to represent that, if after imagining δ, a certain formula φ holds, then it need *not* be the case that, after imagining $\delta \vee \gamma$, the same formula φ holds, as the resulting imaginary worlds created by both acts of imagination will clearly be different. Our logic does account for that, as the following invalidity shows:

$$\not\models \langle I(\delta)\rangle\varphi \rightarrow \langle I(\delta \vee \gamma)\rangle\varphi \tag{4.8}$$

Proof. Consider the Model for Imaginary Scenarios depicted in Figure 4.12, which represents a model containing a real world w, an act of imagination with initial premise p, which results in world v (note that the relation from w to v is an R_{Img} relation), and a different act of imagination with initial premise $p \vee q$, which results in three different worlds: u_1, u_2 and u_3. Now, if we interpret $\delta = p$, $\gamma = q$ and $\varphi = Kp$, we can see how $\models \langle I(p)\rangle Kp \rightarrow \langle I(p \vee q)\rangle Kp$ is not valid in this model. Whereas in the imagination act leading to world v we indeed have Kp at any resulting imaginary world, we do not have Kp at the worlds resulting from the act of imagination with content $p \vee q$, as world u_2 particularly does not satisfy p, and so it prevents Kp from being true there. □

Looking back at the theories of imagination we reviewed in Chapter 2, there is also one very important property imagination must account for, and which we also want our system to represent: we want to account for the *quarantine* effect. Briefly recalling, the

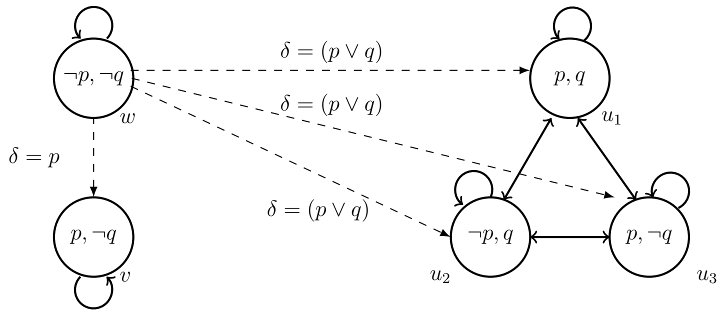


Figure 4.12: A countermodel for formula 4.8.

quarantine effect says that people do not (usually) mix their imaginings with their beliefs, and so that they are able to tell the difference between an imaginary world and the real one. In our proposal, this can be expressed by saying that if a possible world is imaginary (and we can express this, thanks to hybrid logic, by saying that such world needs an act of imagination in order to be reached), then the agent does not consider it to be an epistemic alternative to the actual possible world. By using nominals, this can be nicely expressed in our logic with the following validity:

$$\models \langle I(\delta) \rangle i \to \neg Mi \tag{4.9}$$

Proof. By the definition of satisfiability of operator $\langle I(\delta) \rangle$, we know that we have an R_{Img}-accessible world identified by nominal i (and, moreover, by the definition of satisfiability for nominals, we know that there is only one world identified by i). By the way the `ImgAlg` works, we know that the only R_K relations that are added to a model are between those worlds that are created during the execution of the algorithm; similarly, those worlds are accessible through R_{Img} only from the world of reference (which, obviously, is not one of the worlds created during the execution of the algorithm). Therefore, we know that if i is a world accessible, from the current world, through the R_{Img} relation, then it must be a world that was created by performing an act of imagination, and taking the current world as the world of reference; moreover, we know that i cannot be related by the epistemic indistinguihsability relation, as this would mean that the `ImgAlg` has drawn an epistemic indistinguishability relation between the world of reference and one of the new imaginary worlds, which is not possible. If this was the case, it would amount to the agent believing in something that is imaginary; the way our `ImgAlg` works prevents that from happening. $\square$

As it has just been shown, our formal system can express some of our desired features. The motivation to integrate hybrid logic into our setting was to increase the expressive power in order to be able to capture complex properties of imagination, and, as we have just seen in the previous paragraphs, nominals have been of great help in doing so. By being able to explicitly refer to the possible worlds, we can more clearly draw the distinction between the imagination relation, aimed to capture acts of imagination carried out by the agent, and the epistemic indistinguishability relation, aimed to represent the uncertainty of the agent about the state of affairs of some scenario, be it real or imaginary.

4.8 Thoughts on the Logic of Imaginary Scenarios

So far, we have introduced the motivation behind our approach, we defined the language, models and semantics of our system, and discussed some interesting properties that the logical language allow us to capture. However, and because we wanted to start building our system step by step, there are some features that our system can indeed account for, while there are some others that, although also in our desiderata, it cannot yet capture. During the remaining of this chapter, we discuss the relation of our system with the philosophical properties we want to represent in it.

4.8.1 The Good

We wanted to define a logic able to capture the dynamics involved in acts of imagination, while clearly distinguishing between the *voluntary* and the *involuntary* role of the agent in this process. The Logic of Imaginary Scenarios captures the former mode through its dynamic operator $\mathrm{Img}(\delta)$, representing the action of the agent engaging in an act of imagination with content δ, while it captures the latter through the way the `ImgAlg` handles the creation of the imaginary worlds. In this sense, the Logic of Imaginary Scenarios is on the right track.

Furthermore, the relation between the agent's imaginings and the real world accounts for the *quarantine* effect: namely, the agent does not believe that imaginary worlds can *actually* be the case. Imaginary possible worlds and real possible worlds are only related via the imagination relation, and there are no epistemic indistinguishability relations linking a real world with an imaginary one: this is ensured by the way the `ImgAlg` works. When considering epistemic alternatives to real possible worlds, the agent is only concerned with other *real* possible worlds; when entertaining an imaginary scenario, the agent is only concerned with *imaginary* possible worlds. Our proposal, therefore, does account for this quarantine effect.

However, this distinction between the real and the imaginary may sometimes fail, thus resulting in *contagion*: namely, the agent's imaginings affect the agent's beliefs (see Section 2.2.2 within Chapter 2 for a more detailed explanation). Light cases of contagion, such as being afraid after seeing a horror movie and imagining a monster crawling from under our bed, is something we experience quite frequently. However, the contagion effect is sometimes responsible for situations that are far more serious. For instance, some of the most severe cases of contagion take place in psychic disorders involving hallucination episodes, in which the agent is no longer able to tell the difference between what is real and what is not. A particular instance of these psychic disorders is the *Capgras delusion* (see [34], page 118), in which the agent believes that someone close to her (say a partner, a friend or even a pet) has been replaced by a clone (roughly speaking and without going into the details about this disorder). Now, although our system respects the quarantine effect by default, it *can* also represent this disorder with the kind of models it already has available.

For example, consider an agent who knows, in the real world, that "my brother is my actual brother" (represented by the atomic formula b) and who, as a result of the aforementioned condition, suddenly comes out with the idea that her brother has been replaced by a clone. This is, of course, and imaginary world, as it cannot be the actual case, but nevertheless the agent starts considering that it *could* actually be the case, and starts being unable to distinguish whether the real world is such that her brother is himself, or not: namely, she starts being unable to tell the difference between a real world, and an imaginary one. We can represent this effect in our system by performing an act of imagination in which the agent imagines $\neg b$ (representing that "it is not the case that my brother is my actual brother"), and then adding an epistemic indistinguishability relation from the real world to the imaginary one: in this model, the agent cannot longer tell the difference between the real world, in which her brother is her actual brother, and the imaginary world, in which that is not the case, as Figure 4.13 represents[11].

Note how, as this kind of situations are rare cases, the way accessibility relations behave differ from the restrictions we imposed when defining the `ImgAlg`. Nevertheless, and although we will pursue this issue no further in this work, we think that the Logic of Imaginary Scenarios could be a promising way to represent psychic disorders involving

[11]Epistemic indistinguishability relations are represented by solid lines, whereas the imagination relation is represented by a dashed line.

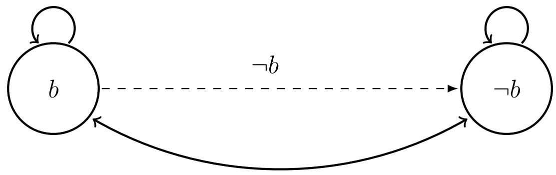


Figure 4.13: Modeling the Capgras delusion.

contagion or hallucination between real worlds and imaginary worlds. In order to do so, it could be possible to define different versions of the `ImgAlg` which, representing different psychic disorders, behaved differently in the way acts of imagination are performed[12].

4.8.2 The Bad (and the Ugly)

One of the main shortcomings of our system is related to the way the agent imports information not directly specified by the initial premise. As seen in the theories reviewed in Chapter 2, an important feature of imagination is that most of the information developed in an imaginary scenario is based on reality-oriented rules and facts which, in turn, are believed or known to be true by the agent; this corresponds, precisely, to the *mirroring* effect.

The first shortcoming, with respect to the mirroring effect, is that our formal system misses a way to account for the notion of "reality-oriented rules". We may argue that, once the initial premise is clamped into the new imaginary worlds, the `ImgAlg` looks into the world of reference and imports "reality-oriented facts", being the atomic formulas that describe the state of affairs represented by the world of reference. Nevertheless, importing specific *facts* into the imaginary scenario is not the same as using *rules* to infer what else would be the case in there.

In particular, take, as an example, an imaginary scenario initiated by using the premise "I have wings". In this scenario, we could obtain new information such as "Bird have wings as well", or "Paris is the capital of France", by importing facts from the world of reference; nevertheless, if we wanted to obtain new information like "I can fly", we would need to import not just a specific *fact* detailing how things actually are in the real world, but rather we would need to use a sort of *rule* stating, for instance, that "If I had wings, then I would be able to fly", and then add this new information by inferring it from the

[12]In particular, note how $\langle I(\delta)\rangle i \vDash \neg Mi$, which is a validity of the Logic of Imaginary Scenarios, would *not* be valid in a model representing the Capgras delusion.

previous rule and the fact that I am imagining that I have wings. As we can see, our Logic of Imaginary Scenarios can account for importing reality-oriented facts, but it cannot yet capture the way we develop imaginary scenario by using reality-oriented rules.

Still related to the mirroring effect, the second shortcoming of our formal system is directly inherited from the fact that the single-agent epistemic logic we took to build our logic upon cannot, at least for now, represent beliefs in an explicit way. Even though we can interpret operator M as a weak form of belief in which, by being uncertain about p, the agent "believes that p could be the case", this interpretation lacks the kind of "preference ordering" which is often attributed to our beliefs, making some of them more plausible than others. Nevertheless, this is the stance we have taken in our proposal regarding the way we use the single-agent epistemic logic, and in order to bypass the lack of an explicit belief operator. By performing an act of imagination at a possible world w, the agent considers, in the resulting imaginary worlds, that her beliefs are those facts represented in world w. This, however, usually forces our agent to imagine more than she is supposed to; needless to say, this may do the trick as a first approach, but it is not as accurate as we would like it to be.

Consider a possible world w satisfying $\{p, q\}$, and an epistemically accessible world v satisfying $\{p, \neg q\}$. Now, suppose the agent performs an imagination act with an initial premise $\neg p$ by taking world w as the world of reference. In the resulting imaginary world, the agent would clamp $\neg p$, as expected, and then she would import her "belief" of q (which is true at world w), thus resulting in an imaginary world satisfying $\{\neg p, q\}$. Regrettably, this represents more than what we wanted: in this case, the agent is not just imagining $\neg p$, but she is also imagining q (or, more precisely, she is imagining to believe q, and then importing it). The whole example is represented in Figure 4.14.

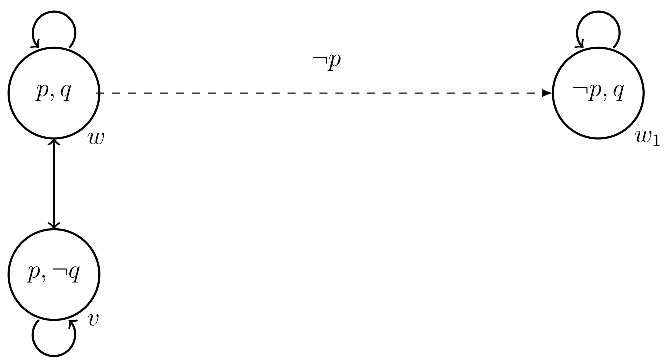


Figure 4.14: The agent imagines more than what she should.

Aside from adding an explicit representation for beliefs in our logic, we would need to tune up the `ImgAlg` in order to handle beliefs appropriately (once the underlying system was able to represent them). Up until now, and given the way we use the expression $M\varphi \equiv \neg K \neg \varphi$ in our logic (as a weak form of belief), it is enough to import the propositions holding at the world of reference, as they reflect one of the possible scenarios the agent believes (considers possible) that could be the case. However, we are still missing the "ordered" flavor that beliefs have: after all, we do not usually believe that everything we consider possible is *equally* plausible. As we briefly introduce in the following lines, there are different approaches in formal logic that allow to represent this notion of "ordered belief". Therefore, if we update our logic with this notion of ordered belief, then importing the atomic propositions being true at the world of reference will not be enough: we would need to tune up our `ImgAlg` so that only those atomic formulas believed by the agent are imported, regardless of whether they are true at the world of reference or not. As "belief" is, in those approaches, a modal attitude just as knowledge, then the `ImgAlg` should take into account not only the world of reference, but also those worlds that are belief-accessible from it.

Therefore, the solution needed to overcome the issue regarding the mirroring effect needs two changes in our logic:

1. Expand the system with a new operator B able to explicitly account for "believe" as a modal attitude.

2. Update the third step of the `ImgAlg`: instead of going over the atomic propositions that are true at the world of reference, the algorithm should check, for each atomic proposition in the model, which ones are actually *believed* by the agent, and import only those.

There are many alternative ways to represent an agent's beliefs (see [50]), but there are two approaches that are specially interesting for us: either to add a new binary relation to our current setting (as suggested in Chapter 7 of [50]), or to use the so-called *plausibility models* (also in Chapter 7 of [50], or in [1]).

The first alternative would be more similar, technically speaking, to the logic we have already defined: to add a new binary relation R_B, standing for a doxastic relation, and defining a new modal operator $B\varphi$ standing for "the agent believes φ". There is a vast amount of literature concerning which formal properties the belief relation should have, and so we should need to study in detail which alternative seems more promising. Maybe the most straightforward approach is to define a binary relation with almost no restrictions at all (regarding its relational properties), and read it as a pointed relation (w, v) going from world w to v: then, the agent's beliefs would be whatever is the case at the world or set of worlds where the binary relation leads to[13].

The other alternative, using plausibility models, allows to represent different worlds over which the agent has a preference relation, leading to those ones she considers more likely to be the case. Even though this approach differs more from our initial setting that the previous alternative, it has shown very good results in the logics of belief revision literature (see Chapter 7 of [50] for a comprehensive introduction). As this approach can also represent the agent knowing (or not knowing) something, it looks like a promising way of solving our belief-import issues as well.

Peeping Into Plausibility Models

Things are seldom as simple as they sound, however, and adding an explicit representation for beliefs in our system is not an exception. The reader may be asking himself "well, if the solution is so straightforward, why not simply add it now to the Logic of Imaginary Scenarios?". After considering how we could add beliefs into our system, and when giving a closer look on how our logic would behave in such kind of models, we already foresee certain issues we would have, regarding new limitations that such models would impose in the way our agent imagines.

In order to illustrate our concerns, let's follow the example depicted in Figure 4.15, which corresponds to a plausibility model[14] in which there are no acts of imagination yet.

Now, let us suppose that we have already tuned up our Logic of Imaginary Scenarios in order to be used with plausibility models. Without getting into technicalities, we would need, as specified in the previous paragraphs, a version of the `ImgAlg` that would import *only* those atomic formulas holding at the most plausible world, which is the one that is actually believed by the agent. In the example depicted in Figure 4.15, those believed atomic formulas would be p, $\neg q$ and r.

[13]This solution, nevertheless, is not as straightforward as it seems and has its own paradoxes: see page 322 of [50] for an example showing certain issues derived from this approach

[14]The accessibility relation in this kind of models is called the *plausibility relation*, and it determines the order in which the agent thinks that a certain possible world is more plausible than another one. In the example of Figure 4.15, the agent considers that both worlds w and v could be the actual state of affairs, but she beliefs that w is more likely to be the case (or more plausible) than v. Even though we could say that the agent considers possible any of the two existing possible worlds, her *beliefs* are only represented by the state of affairs corresponding to world w.

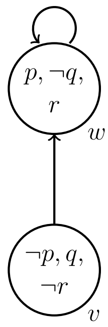


Figure 4.15: An initial plausibility model, with no acts of imagination.

Suppose that we now execute an act of imagination Img(q) while taking, as the world of reference, w, which is the possible world the agent actually believes to be the case. In this case, the `ImgAlg` (which is initiated with parameters `ImgAlg`(q, w)) needs to clamp q at the new imaginary world, and then import those atomic formulas that are *believed* by the agent at the real world. This time, we do not need to import those atomic formulas $s \in$ ATOM such that $\mathcal{M}, w \vDash s$, but rather those such that $\mathcal{M}, w \vDash Bs$. Although we do not provide the formal definition of the semantics for operator B in plausibility models (again, we refer to Chapter 7 of [50]), it suffices to say that the agent believes φ (that is, $B\varphi$) if and only if φ is true at one of the so-called *top-worlds*, or worlds that are at the top of the pointed plausibility relation. In our example, as we are looking for atomic formulas, we would need to import p and r, as $\mathcal{M}, w \vDash Bp$ and $\mathcal{M}, w \vDash Br$. In this particular act of imagination `ImgAlg`(q, w), the world of reference is taken to be, precisely, the same possible world which also contains the agent's beliefs. The result of such act of imagination is depicted in Figure 4.16.

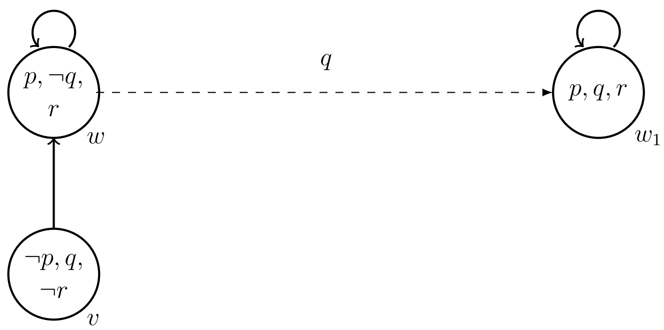


Figure 4.16: The agent imagines by taking her believed world as world of reference.

The previous act of imagination behaves as we expected. The agent imagines something and, conversely to what happens in our actual definition of the Logic of Imaginary Scenarios, which cannot explicitly account for beliefs, in this setting our logic would be able to import not any atomic formula being true at the world of reference, but only those atomic formulas that are actually believed by the agent. Although this feature may seem, at first sight, a valuable addition, it soon becomes vacuous, if we keep considering how our system would behave: in order to see why, let's keep moving forward in our example.

Consider, now, that we execute the same act of imagination with content q, but by taking world v as the world of reference this time. As expected, the `ImgAlg`, called as `ImgAlg`(q, v), would create a new imaginary world, clamp q into it, and then it would need to import the atomic formulas believed by the agent. It turns out that, conversely to what happened before, v is *not* one of the possible worlds believed by the agent, and so the atomic formulas at v do not represent what the agent believes. In fact, we have that, as it happened before, $\mathcal{M}, v \vDash Bp$ and $\mathcal{M}, v \vDash Br$, which means that the `ImgAlg` would need to import, again, atomic formulas p and r, as it did in the previous act of imagination. The imaginary world resulting from this new act of imagination, which took world v as the world of reference, would be equivalent to the one we created in the previous act of imagination by taking w as the world of reference, as it can be seen in Figure 4.17.

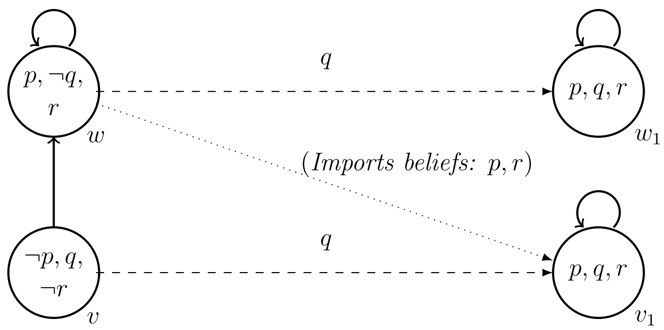


Figure 4.17: The imported atoms come from the same world as before.

Therefore, we can see how, when using plausibility models or, in fact, any explicit representation for beliefs in our logical system, the world of reference would no longer matter, as the atomic formulas believed by the agent would always be those atoms holding at the same possible world, no matter what[15].

[15]It can also be the case, in plausibility models, that we have not just a single top-world, but rather a set of various top-worlds, which the agent believes to be the case, but over which she has no preference.

After considering this simple example, we ask ourselves: what would be the role of the world of reference, then, if our Logic of Imaginary Scenarios could account for an explicit representation of beliefs? Moreover, and if the agent's beliefs were always taken from the same real possible world, would the belief structure represented by the real possible worlds even matter, in our system? If the imported atomic formulas were always taken from the same real possible world, why would we even consider any real possible world different from the top-world?

Taking this into account, we argue that, although adding an explicit representation for beliefs is worth considering, our current approach to the Logic of Imaginary Scenarios, in which we have chosen to account for a weak form of beliefs, is, after all, an approach that gives us more freedom, in terms of exploring the different acts of imagination the agent could perform. This is not to say that we could not include explicit beliefs as part of our system, but we have just shown that, in such cases, we would also need to come up with a way of avoiding a kind of "non-top-world-trivialization" or, in other words, of avoiding that the presence of possible worlds that are not among the agent's explicit beliefs becomes trivial.

With respect to the shortcoming we identified in our current approach to the Logic of Imaginary Scenarios, in which the agent, in some cases, imagines more than what she is supposed to, we can see how this shortcoming may not be that bad, after all —specially if we were to reinterpret our reading of imagination acts as "the agent imagines δ while considering that the actual state of affairs is represented by the world of reference". In the end, it all comes up to a matter of balancing: either we allow to take a different "weakly-believed" possible world each time the agent imagines, and so we allow for the same imagination act to be unfolded in different ways, depending on the world of reference, or we clearly identify a possible world believed by the agent, and so we trivialize any real possible world other than that, when it comes to importing the agent's beliefs. As our main goal is to represent and understand the dynamics involved in the creation of imaginary worlds, we claim that the approach we have taken in the Logic of Imaginary Scenarios provides a more suitable and flexible setting for doing so, and that, at least for now, the way it account for beliefs is enough to allow our system to capture some interesting properties of voluntary imagination acts.

We will not unfold the technical details of how this case should work, but the `ImgAlg` would probably need to duplicate the same structure of top-worlds in the imagining to represent the fact that the agent's plausibility order over certain facts is not determined.

4.9 Foreseeing an Epiphany

In this chapter, we have defined a logic aimed to capture dynamic acts of imagination, as presented in the theories reviewed in Chapter 2.

While discussing what the system can account for, however, we have ended up with a kind of sweet-and-sour feeling. Even though it is true that the Logic of Imaginary Scenarios can account for some relevant features of our desiderata, such as the quarantine effect and some interesting properties captured by certain validities, there are still important shortcomings. The lack of an appropriate way of capturing the reality-oriented development of an imaginary scenario, which concerns both the representation of reality-oriented rules and the explicit account for beliefs, is what concerns us the most. By the end of this last section, we have seen how, by improving our basic logical system, we could accommodate a (seemingly) more satisfying version of the mirroring effect, regarding an explicit representation of beliefs —although, when doing so, we would also need to consider how to account for the relevance of the world of reference, as we have shown through an example. However, while considering different solutions and looking for alternative approaches, we realize that there is still something we feel is missing.

When revisiting the theories of imagination reviewed in Chapter 2, we feel that our logic ends up being a bit straightforward, and it is still a bit far away from our intuitive understanding of imagination. In particular, the way we elaborate imaginary scenarios by using reality-oriented rules and facts feels a bit distant, even if we could try to find a solution to account for the mirroring effect by integrating beliefs. As we have already noticed, we miss the notion of "rule", as part of the development of an imaginary scenario that is based on reality-oriented rules and facts. Our system now aims to import just facts, but it cannot account for rules, and adding an explicit representation for beliefs would not change that.

The next chapter represents a spin in our work. In there, we briefly discuss whether the direction we are in is the right one, according to a philosophical point of view on our goal. The main aim of this work is to understand, aided by formal tools such as logical systems, how the dynamics of voluntary acts of imagination work in our minds, and how we create and develop imaginary worlds. Are we closer to our initial goal? Is the approach we are following the right one? Is formal logic helping us reach our goal and, if so, up to which amount? The answers to these an other questions motivate, in the next chapter, a change of approach in our work. As we explain in there, this change will require us to dive deeper into the philosophical theories of imagination, up to the point of questioning whether the current theories of imagination are detailed enough for our purpose. Then, we will choose again logic as our companion, but from a different perspective: instead of in its formal language and semantics, our new approach to a logic of imagination will focus deeper on its dynamics and the algorithms that capture it.

Chapter 5

Conclusions on Part I

Throughout the first part of the present work, we have introduced, in Chapter 2, the notion of imagination from a philosophical perspective, and reviewed three influential theories about it. Then, in Chapter 3, we have briefly presented the basics of formal logic and reviewed different logics aimed to represent imagination. After arguing that such logics do not account for acts of imagination in a dynamic way, we have presented, in Chapter 4, the Logic of Imaginary Scenarios, which captures the dynamic side of imagination acts through the Imagination Algorithm; as we have discussed by the end of that chapter, our proposal has some strong points, but also some caveats that concerns us.

The present chapter represents a brief stop in our work, in which we come back to the surface, take a deep breath, and take a look at the holistic picture of how our work is going so far. In particular, we ask ourselves the following kind of questions: is our logic in the right track in order to capture acts of imagination? And, if it is not, where does the problem lie, and why is it a problem? Is the problem mainly formal, or does it also have a theoretical dimension?

5.1 About the Formal Approach

Although in sections 4.8.1 and 4.8.2 we already discuss and analyze in detail the Logic of Imaginary Scenarios, that discussion is directed towards its formal behavior and its relation to the theoretical background. Looking at a bigger picture, we wonder whether, regardless of possible improvements that could be made in our proposal, the general approach is the right one.

In particular, we feel that the way our logic captures the dynamics of imagination acts by using a single algorithm is not modular enough. Our Imagination Algorithm captures an act of imagination as a single process (see sections 4.1 and 4.4 for details): the agent sets an initial premise δ, creates a set of new possible imaginary worlds, imports atomic formulas from the world of reference, and the act of imagination is considered finished. Then, if the agent wants to imagine something else regarding one of the recently created imaginary worlds, she must perform a new act of imagination with a different premise γ; this, by the way we understand our algorithm, amounts to intentionally adding a new premise to the imaginary scenario, and thus corresponds to the *atypical* or *non-usual development* mechanism identified by the theories of imagination (see Section 2.4).

However, this seems a rather simplified account of all the things that happen *within* an act of imagination. Our agent sets the initial premise of a new imaginary scenario, unfolds every possible imaginary world represented by it, and imports every remaining detail; then, if she wants to keep elaborating on the scenario, she must replicate the same kind of process. In other words: as soon as the agent has specified the initial premise describing the new imaginary scenario, she loses control over the rest of the process. Is this the way acts of imagination work? Even though the actions of initiating an imaginary scenario and afterwards adding a new premise into it do behave in that way, the development of imaginary scenarios falls short; surely there must be something more to it than just importing facts from the world of reference. After all, we do have some kind of control over our imaginings and the way or the "direction" towards which we develop them (this is also pointed out by Peter Langland-Hassan, whose theory we reviewed in Section 2.3.3). In particular, it is almost impossible to tell, just by looking at the initial premise of an imagining, where would that imagining go, and how would the agent develop it. However, the way our proposal handles the reality-development mechanism does not allow us to capture and account for this different possible ways in which an imaginary scenario could be elaborated.

The key to solve this issue, then, lies in digging deeper into how imaginary worlds are developed, by following reality-oriented rules and facts. Importing atomic propositions the way our logic does is not enough: we also want to account for this kind of agentiveness hidden behind the way an imaginary scenario develops, and thus we need to understand, in more detail, how this mechanism works. However, this pops out a new question into our minds: is this issue something purely formal, or are its roots deeply buried within the theoretical background as well?

5.2 About the Theoretical Approach

We want to improve the way a logic of imagination acts captures the *reality-oriented development* mechanism, as the mechanism responsible for determining, once an imaginary scenario has been created, how it would usually unfold, if it was real.

This intuition, however, goes beyond the way our logic captures this mechanism. In the Logic of Imaginary Scenarios, the *reality-oriented development* of a scenario corresponds to importing those facts (represented by atomic propositions) that are true at the world of reference, and which do not contradict the clamped premise in the imaginary world. So, for instance, if the agent imagines that "human beings have wings", the way our logic develops the resulting imaginary scenario would import things such as "the capital of France is Paris", or "I am wearing jeans and a T-shirt", or "the Moon is not made of cheese"; namely, it will import atomic propositions describing different *facts* about the actual state of affairs.

However, and although those facts may still be useful when filling up the details of an imaginary scenario, they do not capture another kind of details that one would likely expect to pop up in such scenario; namely, something like "if human beings had wings, then they would be able to fly". This kind of reasoning is something that, most probably, everyone imagining such situation would reach to, but it is not captured by what our logic imports. Moreover, there are also a different kind of rules one could use to elaborate on that imaginary scenario; for instance, something like "if human beings had wings, then those wings would be made of feathers". In this case, one could add information regarding certain facts of the imaginary scenario that were not even there before; namely, *how* do human wings look like, or what are they made of. The reason behind this lies in the fact that our Imagination Algorithm imports *facts*, but does not account for *rules*. Our Logic of Imaginary scenarios elaborates the new imaginary worlds by filling up the details about what else would be the case in there, regarding the actual state of affairs in which the imaginary worlds are based, but they are not elaborated using these kind of *hypothetical rules* aimed to capture the consequences of certain non-actual facts: our main problem, then, lies in there.

Once the problem has been identified, the solution may seem at first pretty straightforward: we should modify our algorithm in order to account for those "rule-like" kind of formulas to elaborate our imaginary scenarios. There is still, however, a kind of security-check we have to make before going back to the formal setting: is the *reality-oriented development* mechanism detailed enough to allow a new version of the Imagination Algorithm to handle all this? In other words; is just this notion of "rule" all we were missing when defining the Logic of Imaginary Scenarios, or do we lack something more, even before diving into the depths of formal languages?

5.3 A Turn of Events

The considerations we present in this chapter of how our work is going so far fuels a change in our approach. By analyzing how well our formal logic succeeded in capturing acts of imagination, we have detected certain shortcomings of our formal system which, in the end, suggest that there is something fishy about the theories of imagination we used as our theoretical background. Therefore, we need to go back to the theories of imagination, as our suspicion is that they are not fine-grained enough to be represented in a formal way through a precise algorithm.

Due to this, our work takes an important spin: instead of elaborating upon the Logic of Imaginary Scenarios, we go back to the theoretical background to perform a critical reanalysis of the theories of imagination. In Chapter 2 we saw what they had to say, and we took them for granted as our theoretical framework; now, we go back to them with a critical eye in order to asses whether they are exhaustive enough to define all the pieces of the puzzle we want to solve, and we end up proposing our own theory for imagination acts. This turn of events, then, represents the end of the first part of our journey, and the beginning of its second part.

We have been down the rabbit-hole; now, let's squeeze ourselves even deeper into it.

Part II:

Deeper Down the Rabbit-Hole

"But I don't want to go among mad people," Alice remarked. "Oh, you can't help that", said the Cat: "we're all mad here. I'm mad. You're mad." "How do you know I'm mad?" said Alice. "You must be," said the Cat, "or you wouldn't have come here."

—Lewis Carroll
Alice's Adventures in Wonderland

Chapter 6

A New Theory for Imagination Acts

In the previous part of this work, we have reviewed, in Chapter 2, three relevant theories of imagination, and taken them as the background upon which to build a logic for imagination acts. Although those theories do provide a detailed account for how imagination acts work, plugging in a formal system has revealed certain shortcomings.

In this chapter, we pick up those theories of imagination and dig deeper into their mechanics. After doing so, we argue why the three mechanisms we previously identified are actually not fine-grained enough. Then, we propose our own theory of imagination acts, which we claim identifies the least number of different mechanisms that any theory of imagination should account for: the Common Frame for Imagination Acts.

Once the Common Frame for Imagination Acts has been presented, and keeping in mind that the main goal of our work is to provide a systematic analysis of imagination acts, we define a measurement tool that can be used to provide a visual representation of their dynamics: the Rhombus of Imagination. After defining it, we show how it can be used to create the "blue-print" of different kinds of imagination acts, and how it allows us to characterize certain properties that every imagination act accounts for.

6.1 The Common Frame for Acts of Imagination

Briefly recalling what we saw when comparing the theories of imagination in Section 2.4, we identified three different mechanisms involved in an act of imagination:

1. An *initiating* of the imagining, representing the voluntary action of the agent to imagine a scenario characterized by a certain initial premise.

2. A *reality-oriented development* of the imaginary scenario, which follows rules and facts that are believed or known by the agent, and which mirror the way the scenario would likely be, if it was real.

3. An *atypical* or *non-default addition*, which represents another voluntary action of the agent to add a new premise that does *not* follow from reality-oriented rules and facts regarding the way the imaginary scenario is.

Although these mechanisms seemed precise enough at first glance, after trying to capture them in a formal logic, we saw how there is something still missing.

In particular, our concerns lie within the *reality-oriented development* mechanism. An imaginary scenario can mirror the real world in different ways, depending on why we are entertaining such scenario. Consider, for instance, an agent who decides to imagine a chessboard by creating a new imaginary world initiated by the premise "there is a chessboard". Among other reasons, she could be imagining a chessboard because:

1. She wants to paint one upon a canvas.

2. She wants to decide which next move is the best, given a certain setting of the pieces in the board.

Now, according to each of these possibilities, the agent could develop her initial imagining by:

1. Filling up the details about how the chessboard is; for instance, whether is it made of wood, what color is it, and so on.

2. Considering the possible moves she could make and evaluating their consequences.

Now, a question pops up into our minds: in the first case, should the agent be concerned with *what* happens within the chessboard, meaning how she could move the pieces? If she is just imagining the chessboard because she wants to paint it, then it would be of no use. Regarding the second case, should the agent be concerned with *how* the chessboard is, like whether it is made of wood or not? If she just wants to asses which is the best movement to make, it would be completely useless.

Having reached this point, our concerns have been made clear. When assessing whether the Logic of Imaginary Scenarios was on the right track, in Chapter 5, we identified that we were missing the notion of "rules" upon which the imaginary scenario was developed in a reality-oriented way. When looking closer to how the existing theories of imagination treat the reality-oriented development, we notice how this mechanism actually embeds different ways of developing an imagining. As we see it, one could fill up the *static* details of the (which would correspond to case 1 of the previous example about imagining a chessboard) without involving any kind of *dynamic* action that would typically take place in it (which would correspond to case 2 of the chessboard example), and the other way around. Although they both use reality-oriented rules as their input, we think it is worth considering them apart.

In the following subsections, we identify and define the processes that form our new theory for imagination acts, which we call the *Common Frame for Imagination Acts*, and which we claim that is better suited to study the dynamics of imagination acts, compared to the theories previously reviewed in Chapter 2.

Note that, from now on, we use the word "process", instead of "mechanism", to refer to the distinct parts forming the Common Frame. We do this because, unlike previously, where we just wanted to understand what characterized an act of imagination as a whole, we now want to account for a more modular vision of acts of imagination. As we will see later, this modular approach would allow our theory to characterize different imagination acts according to how much they rely on each one of the processes we define. Once our proposed theory has been introduced, we show where the previously-reviewed theories match with it, and where they are actually collapsing distinct processes into a single mechanism.

6.1.1 The Initialization

We call the first process of our frame the *Initialization*, and we associate to it the question "what describes the initial scenario?". It is characterized by the agent performing a mental action initiating an episode of imagination with a certain initial premise.

This premise may be more or less specified; for instance, it can either be something like "I am the singer of a jazz band", or "I am the singer of a jazz band formed by a piano, a double-bass and drums, where we are all wearing classy tuxedos". Either way, a new imaginary scenario will be created: the difference is that the scenario resulting from the former premise will be more less detailed, with less things being characterized in there, whereas the scenario resulting from the latter premise will already contain more details right up from its creation.

Figure 6.1 depicts the *Initialization* process. It is important to stress the fact that this figure is *not* a Kripke model, like the ones we introduced in Section 3.1.2, or the ones we defined and used throughout Chapter 4; rather, it aims to schematically represent how the process takes places, and what it involves. In the next subsections, we also use similar figures which, like in this case, are not to be understood as being formal models, but rather as schematic representations.

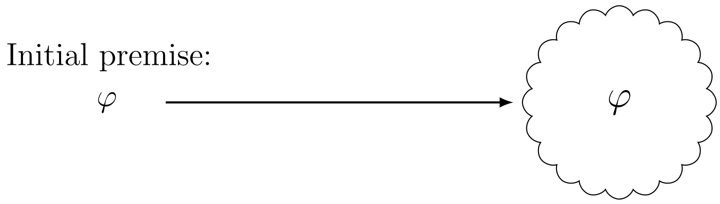


Figure 6.1: The *Initialization* creates a new imaginary world, given an initial premise.

6.1.2 The Description

This second process, which we call the *Description*, and which answers the question "how does the scenario look like?", corresponds to the "static" part of the elaboration of an imaginary scenario. This process captures the action of filling up the details of how the initial imaginary scenario could be like, regarding what other facts could also be the case in there, but without evolving it into a new, different scenario. We call the kind of rules governing the static description of an imaginary scenario the *factual rules*.

For example, an imaginary scenario initiated by the premise "we are having a tea-party" may be enriched with details about the shape of the tea-pot, the presence or absence of cookies on the table, the room we are in, etc. Note how these details refer only to the *current* scenario we are describing: we are not yet putting our imagining in motion, but just forming a more detailed static version of the initial scenario.

We could draw an analogy with logical languages, in which this kind of enrichment would correspond to "factual formulas" of the sort $t \to c \vee k$: "if there is a tea-party, then there are also cookies or cake". Note how atomic formulas c and k say something about the *same* scenario that satisfies t: we are just adding details to it.

At this point is where our proposal starts departing from the previously reviewed theories, and starts being also more fine-grained than them. Although, when considering reality-oriented development, Nichols and Stich do mention that both "facts" and "scripts" (or "paradigms") are involved in this process, they collapse them both into the UpDater mechanism[1]. Langland-Hassan's theory behaves similarly to Nichols and Stich's, and he attributes the description process to the lateral constrains which, again, are the same ones that govern how the scenario would typically unfold. The same goes for Williamson's theory: involuntary imagination handles both the description of the imagined scenario, and unfolding the way it would typically evolve.

Figure 6.2 represents how the *Description* process works. In particular, the process involves using factual rules that provide details about how the current imaginary scenario could be, regarding its state of affairs, and with respect to what already is the case in there.

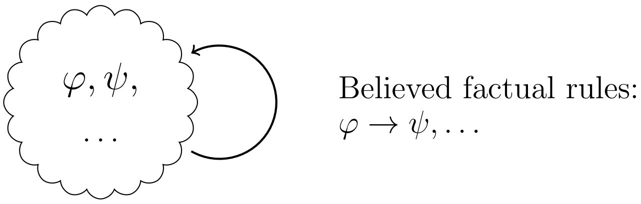


Figure 6.2: The *Description* process elaborates on the static details of the scenario.

6.1.3 The Default Evolution

Following the considerations made in the previous paragraph, we call the third process the *Default Evolution*, and we associate to it the question "what would typically happen in the scenario?". Thus, the main point of this process is to determine, still by using reality-oriented rules, but of a different kind, how the scenario would typically evolve. We call the kind of rules describing the possible ways a scenario could move forward the *scripts*.

For instance, in the example from the previous section where an imaginary scenario was initiated with the premise "we are having a tea-party", our scripts about tea-parties could tell us something about how the scenario may typically move forward, or what actions could take place in there: "if I am in a tea-party, I could pour tea into my cup", or "if there is a tea-party, everyone is going to sip tea slowly, in a well-educated manner". Note how, with respect to our previous step, these rules do *not* provide information regarding how the scenario could look like, but they tell us what could happen (or what could I do, or how should I behave) in that scenario; they are not rules that elaborate on

[1] Particularly, on page 118 of [43], they identify both processes as the "inferential elaboration", without distinguishing between their static or dynamic character.

the current scenario, but rather rules that describe some “accessible”, or “possible”, or “future” scenarios based on the current one.

Therefore, and going back to the logical analogy, by using modal logic (see Section 3.1.2 for details) we can represent these scripts as something like $t \rightarrow \Diamond p$; meaning that “if there is a tea-party, then I could pour tea into my cup”. The main difference between the scripts used in this process, and rules used in the *Description* process, is that these scripts do *not* tell us anything new about the same world in which t is the case, whereas the rules of the *Description* do. Roughly speaking, whereas the rules used in the previous process tell us something like “if it is the case that t, then it is *also* the case that p”, the scripts we are currently introducing say that “if it is the case that t, then it could *become* the case that p, given a certain change of state triggered by an event”. The scripts, therefore, encode how an imaginary world could change, by encoding what could happen in there as a result of certain events or actions.

Again, and as a consequence of not distinguishing between the static and the dynamic elaboration of an imaginary scenario, this process is also embedded into Nichols and Stich’s UpDater mechanism, into Williamson’s involuntary imagination, and into Langland’s lateral constrains.

Figure 6.3 represent the *Default Evolution* process. Note how, unlike the *Description* process, in this case the scripts are *not* used to elaborate the details of the current imaginary scenario in a static way, but rather to evaluate the possible ways in which such scenario could “evolve”, or move forward in time. Obviously, a specific scenario could usually move forward in many different ways; it is up to the agent, then, to decide which course of events she wants to choose in her imagining.

6.1.4 The Unscripted Additions

We call the fourth and last process distinguished in our frame the *Unscripted Addition*, and we associate to it the question “how does the agent voluntarily change the scenario?”. This process corresponds to those ways of developing an imaginary scenario that are not typical, nor that can be inferred from reality-oriented rules or scripts.

Note how the question with which we characterize this process strongly emphasizes the voluntary intention of the agent to go “off-script” with additions that deviate from what one would expect to happen. As the agent chooses again a new premise to be put into the imaginary scenario, this process behaves like the *Initialization* process of our frame: the agent clamps a new premise into an (already existing) imaginary scenario, and

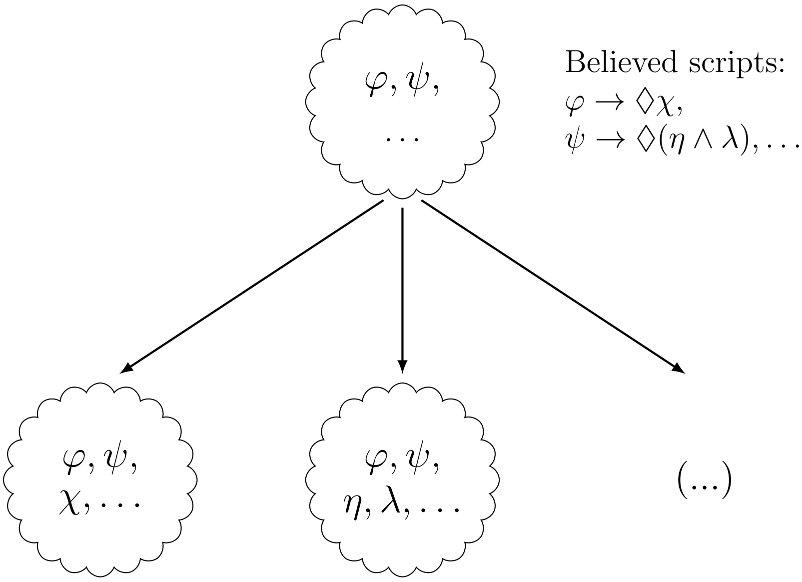


Figure 6.3: The *Default Evolution* process opens possible outcomes of the scenario.

then the cycle of filling up the details, imagining how the scenario moves forward and, if it comes to that, coming up with yet another off-script premise, begins again.

Unlike the *Initialization* process, which is responsible for creating a brand-new imaginary scenario, these additions are built upon an already existing imaginary scenario that is already specified up to a certain point. The new premises added by this process may be completely new, but nevertheless it may happen that they override something already set in a previous imaginary scenario: if the new premise can be held together with what is already the case in the imaginary scenario, so be it; otherwise, it should override whatever was already the case in the scenario that conflicts with it. In other words: these new premises have priority over what else is the case in the imaginary scenario.

For instance, the agent can decide to add, in the tea-party scenario, a new premise w stating that the White Rabbit from *Alice's Adventures in Wonderland* knocks at the door and joins the tea-party: regarding what has already been detailed about how the tea-party looks like, or how should one behave in there, this new premise should not be in conflict with anything. However, if the agent later wants to add a premise a stating that the tea cups are actually cameras, and that someone is spying from them as part of an ambush, then a lot of things concerning the static details of the scenario, and also concerning what could happen now in there, would be in conflict with that, and so they should be either withdrawn or updated to accommodate the new premise.

This process can be mapped into Nichols and Stich's Script Elaborator, and into Langland's cyclical involvement of top-down intentions (driven by a decision of the agent); as we have already argued, Williamson's work does not consider this mechanism of imagination acts, as he is only interested in those imaginings that are reality-oriented and suitable for guiding our actions and improving our epistemic state.

Figure 6.4 represents the *Unscripted Addition* process. Note how this process is, in fact, pretty similar to the *Initialization*: these kind of additions are not implied by any kind of factual rule nor script, but they are instead voluntarily added by the agent. As we have already said, the reason for adding a new premise into the imaginary scenario follows a desire for elaborating the scenario in a certain way.

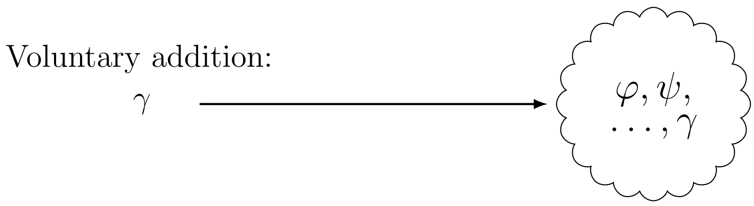


Figure 6.4: The *Unscripted Addition* clamps a new premise into the imaginary world.

6.1.5 Binding Everything Together

Up to this point, we have already identified what we claim are the four processes involved in every act of imagination. Note, however, that most of these processes are not required to be performed in a sequential way; in fact, the only one that is indeed required to happen in a specific order is the first one, the *Initialization*. After all, without a voluntary mental action initiating an imagining, it would not be even possible to entertain such imagining. The other three processes, nonetheless, could be (and usually are) performed in a mixed way, without necessarily following any order.

The Common Frame for Imagination Acts can be summarized in comparison to the previously reviewed theories in the Table 6.1. As we can see, the theories of imagination we reviewed earlier fail to distinguish between some of the processes we have identified in acts of imagination; specifically, the *Description* and the *Default Evolution* processes collapse into a single mechanism in the previous theories. We believe that these processes are different enough and involve using mechanisms that are different enough to be properly distinguished in any theory of imagination.

Casas-Roma et al.	*Nichols / Stich*	*Williamson*	*Langland-Hassan*
Initialization	Premises into PWB	Volunt. imag.	Top-down intention
Description	UpDater	Involunt. imag.	Lateral constrains
Default Evo.			
Unscripted Add.	Script Elaborator	-	Cyclical top-down int.

Table 6.1: The Common Frame with respect to other theories of imagination.

We believe that our analysis of acts of imagination, and the Common Frame for Imagination Acts we propose, can be taken into account as a useful guide for anyone working in the analysis and understanding of the dynamics of imagination. Following our own advice, we take the Common Frame as our underlying theory throughout the remaining of this work, and we refer to it when defining, in Chapter 7, a formal system able to represent the dynamics of imagination.

6.1.6 A Remark About Agentiveness

While giving a closer look to the reality-oriented development of an imaginary scenario, we have come to realize different facts. The most important of them, and the one we already presented, is that reality-oriented development must be split into two distinct processes that use different mechanisms and a different kind of rules in order to be unfolded. Nevertheless, there is something else we noted in our analysis.

The previous theories of imagination (that is, Nichols and Stich's, Williamson's, and Langland-Hassan's theories) recognize the reality-oriented development of an imaginary scenario as something *involuntary*, conversely to initiating a new scenario and adding more premises into it, which are *voluntary* processes. Nevertheless, we believe that this consideration needs an important remark: namely, that the reality-oriented development of a scenario is *not* completely involuntary, but just constrained by certain rules and scripts. This, in turn, affects the way imaginary scenarios are elaborated in two different ways: namely, it affects the "inputs" that will be used in such elaboration, but also the "outputs" that will be followed after that.

Choosing the Inputs: Rules and Scripts

Consider the brief example we propose at the beginning of Section 6.1 about the agent initiating a new imaginary scenario with the premise “there is a chessboard”. Now, as we argue in there, once the initial scenario has been created, it can be elaborated in many different ways; that is, either by filling up the static details of the scenario, by evaluating which possible moves to make, or even to imagine that the pieces of the chessboard are alive in a story-telling session.

The following question is what concerns us: if, once the initial scenario has been set, it is supposed to be elaborated in an involuntary way, then how it can be possible that the same initial scenario can be developed in completely different ways?

Our answer lies in adding a nuance to the meaning of “involuntary”, and bringing into the picture one further voluntary decision hidden just before the reality-oriented development. Whenever an agent imagines something, she does so because she has a certain goal in mind: for instance, one imagines a chessboard either because she wants to form a mental image of it and paint it into a canvas, or because she wants to decide which next move is the best to make within a chess game. Imaginary scenarios, then, are *not* initiated completely independent from a certain goal that the agent has in mind when deciding to entertain them.

Even though we agree with the fact that, once the initial scenario has been set, it is developed “involuntarily” by following certain rules and facts, there is still an important detail that has been left outside of the loop: why does the agent develop the scenario using certain rules or scripts, and not different ones? Why does the agent focuses on how the chessboard looks like in one scenario, but on how could she move the pieces into another? Because, before elaborating the scenario, the goal the agent has in mind (that is, the reason why she is entertaining such scenario) is used to “select” a certain set of rules and scripts she will use, and leaving other rules and scripts out. It may be possible that this choosing of rules and facts is, in fact, kind of automatic, with respect to the goal the agent has in mind; she needs not to “think” about the set of all rules and build her own bag of “rules-to-be-used” before engaging in the imagining, but, nevertheless, they are still chosen in that way, and not in a different one, because the agent has a certain goal, about which she is actually aware.

But there is still more regarding the way the agent uses these rules and scripts. Once there is a certain set of rules and scripts that can be used in a specific imaginary scenario, using one over the other is also something that entails an agentiveness as well. For instance, in the imagining where the agent evaluates which possible move to make in a chess match, why does she evaluates certain possible moves, and leave others unevaluated? Most probably, because she already thinks that only a certain set of moves could be really useful for her, but still, she *decides* to prioritize the evaluation of certain moves over some others. This, in turn, involves an active role of the agent as well. Not only the agent chooses which set of rules and scripts she will use during the imagining, but she also chooses which rule or script she wants to evaluate at any given step.

A proper theory of imagination acts, therefore, should account for this phenomenon. The agent, when elaborating on the details of an imaginary scenario, does so by taking into account not *every* rule and script possibly available for that particular scenario, but just a certain subset of them. Similarly, determining which particular rule or script is used to elaborate on the scenario, also involves a decision of the agent.

Choosing the Outputs: Possible Outcomes

Furthermore, we can find the same kind of "hidden agentiveness" not only when considering the set of rules and scripts involved in the elaboration of an imaginary scenario, but also when considering how, those rules and scripts, determine the way the imagining is elaborated.

For instance, when evaluating the next possible move within an imagined chessboard, it is *not* predetermined which move the agent should imagine herself performing (even if there is a better, more rational move to make): if the agent considers she could either move a rook or the queen, the particular option she decides to consider is *not* involuntary. It may be biased by a certain evaluation that points out to a move that seems better in the long run, but, in the end, the agent is the one who *decides* how the imaginary scenario moves forward. Therefore, reality-oriented development should be considered *involuntary*, but only in the sense that it develops the imaginary scenario by following a certain set of rules and scripts which, according to certain inputs or antecedents, unfold a certain set of possible outputs or consequences; nevertheless, the ultimate decision regarding which possible outcome, among all the possible ones, the agent will end up choosing, is something that comes down to the decision of the agent[2].

These considerations, therefore, call for an important clarification regarding the way the *Description*, the *Default Evolution* and the *Unscripted Addition* processes work. The

[2] Peter Langland-Hassan, whose theory we review in Section 2.3.3, also points out that the agent chooses, among all the available options, which one she wants to follow in her imagining.

concerned reader may think "what, then, makes the *Unscripted Addition* more voluntary than the previous processes? If the agent *chooses* the outcomes of the so-called reality-oriented rules and scripts, then she is in fact adding the premise characterizing such outcome into the imagining, after all". The answer to this question lies in the set of different ways the agent may choose the scenario to move forward, or to increase its detail.

In particular, what the agent chooses after considering the possible outcomes following a rule or script is constrained by such rules and scripts. So, for instance, the agent may choose to imagine that she moves the rook or the queen, or even a knight, according to the way those pieces are allowed to move in a chessboard; however, what she cannot choose to imagine, when evaluating her possible next move in a reality-oriented fashion, is that she takes a pawn at the other side of the board and moves it 8 positions across until it threatens the opponent's king: it does not follow from any reality-oriented script that she could do so, as it violates the rules of chess. In other words, the kind of voluntary decision of the agent that is involved in both the *Description* and the *Default Evolution* processes is not "totally free", but it rather comes constrained by the outcomes of some particular rules and scripts, whereas the new premises added by the *Unscripted Additions* do not. The agent may choose to follow any possible consequence of a rule or a script (either being about the static scenario, or about a possible course of actions), but she cannot deviate from such set of consequences: the agentiveness related to reality-oriented development, in this sense, comes only to choosing among certain possibilities.

Conversely, what we identify as the *Unscripted Addition* process is intended to add new premises that do *not* follow from any rule or fact the agent has available. They are not simply a matter of the agent *voluntarily choosing* something to happen within the imaginary scenario, but it is about her introducing something that is *not* the outcome of any rule or script involved in neither the *Description*, nor the *Default Evolution* processes.

We believe that this remark is important not only as a contribution when refining the notion of "involuntary", as it is used by the previous theories of imagination, but also to strengthen the way we understand these processes within the Common Frame for Imagination Acts, and so how our use of "involuntary" does not exactly match with the way it is used in the previous theories of imagination. As we are looking for a fine-grained, specific analysis of the way imagination acts work, our theory needs to account the distinction between this kind of *constrained agentiveness*, that is hidden into the reality-oriented processes within imagination acts, and another kind of *free agentiveness*, that comes into play in the *Unscripted Addition* process, and which deviate from any possible outcomes foreseen by reality-oriented rules or scripts.

During the remained of our work, this distinction not only plays an important role when analyzing specific instances of imagination acts in the next section, but it will also be one of the features captured by the formal system we propose in Chapter 7.

6.2 The Rhombus of Imagination

After having introduced the Common Frame, and keeping in mind that our main goal is to systematically analyze imagination acts, the following question pops out into our minds: would it be possible to use these distinct processes within the Common Frame to compare different kinds of acts of imagination? This comparison would not be about the *content* of such acts (say, whether one imagining is about a chessboard and another one about a tea-party), but rather about their *structure*; namely, how similar and different they are, with respect to the processes they use. Moreover, would it be possible to use these four processes in order to analyze, distinguish and classify different acts of imagination according to how much they rely on each process?

This is, precisely, the aim of this last section: to propose a tool for representing up to which point different kinds of imagination acts use each of the previously identified processes, and therefore to identify a sort of "blue-print" characterizing them. We call this measurement tool for imagination acts the *Rhombus of Imagination*.

In order to provide a visual and intuitive way of representing this, our tool is formed by a rhombus shape, which we call the *outer rhombus*, and in which each vertex corresponds to one of the processes of the Common Frame for Imagination Acts; then, each vertex is connected to the center of the rhombus by a line, which is used to measure, comparatively to the other lines in the rhombus, how much does a certain act of imagination rely on using that particular process. Specifically, the farther away from the center of the rhombus, the more that imagination act uses a certain process. In particular, there are three distinguished measures in each line, corresponding to a certain *degree of relevance*:

1. The first measure, coinciding with the center of the Rhombus, corresponds to a *marginal use* of a particular process; this represents that a certain kind of imagination acts does not rely at all on that specific process, and that its uses are almost negligible in the overall evaluation.

2. The second measure, placed at the center of the line, corresponds to a *standard use* of the process; it is used during the act of imagination and it adds useful information

during the act of imagination.

3. The third measure, placed at the vertex of the Rhombus, corresponds to a *high use* of the process; it is considered to be crucial in order to reach the goal the agent had in mind when engaging in such acts of imagination.

Therefore, when considering a specific kind of acts of imagination using this tool, we end up having a rhomboidal shape, called the *inner rhombus* and embedded within the outer rhombus, and which corresponds to a sort of blue-print representing the way this kind of imagination acts use each one of the processes (or, in other words, how much they rely on each process).

The empty Rhombus of Imagination, with no act of imagination inscribed in it yet, is shown in Figure 6.5.

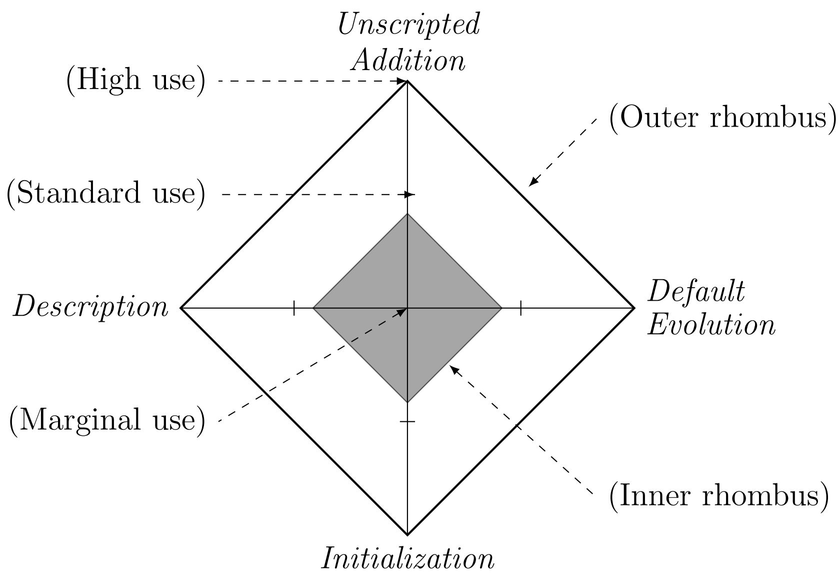


Figure 6.5: The (empty) Rhombus of Imagination.

Note that the way we have distributed each process among the rhombus is not fully arbitrary. We start by relating the *Initialization* process to the lower vertex of the rhombus; as the *Initialization* is the basic process needed to create an imaginary scenario, we associate it to the "base" or the "standing point" of the rhombus. Then, the *Description* and the *Default Evolution* processes are responsible for developing the imaginary scenario, and so we place them forming the "body" or the "core" of the rhombus, at its horizontal axis. Finally, the *Unscripted Addition* process represent additions that depart from what would be typical in the imagining, and so they kind of “go beyond” or “fly away” from what one would usually expect; therefore, we associate it to the upper vertex of the rhombus, also inspired by the expression “let your imagination fly”. Now, how should we use the Rhombus of Imagination?

6.2.1 Classifying Acts of Imagination

In order to classify acts of imagination into different kinds, the Rhombus of Imagination should be understood as a tool for the *qualitative analysis* of the processes involved in such imagination acts. In this sense, a qualitative analysis corresponds to considering, for a particular kind of imagination acts, the degree of relevance that each process has within this kind of imagination acts. Namely, how much do imagination acts of this kind rely on a particular process, in order to reach their particular goal?

Obviously, different instances of specific acts of imagination will use each one of the different processes in different ways, but what the Rhombus of Imagination aims to capture is how relevant they typically are, for this kind of imagination acts. Therefore, conceiving the Rhombus of Imagination for a specific kind of imagination acts requires a sort of general analysis over how such imagination acts typically work, while trying to avoid focusing on particular details of specific instances of such acts. The analysis we have to make in this case, thus, requires an abstraction over the way multiple instances of a particular kind usually work.

In order to show how the qualitative version of the Rhombus of Imagination works, let’s consider three different kinds of imagination acts: hypothetical reasoning, story-telling, and engaging in a preexisting fiction.

Hypothetical Reasoning

When engaging in hypothetical reasoning, one uses imagination as a tool for assessing what would likely happen, or how things would likely be, if things were different as they are in the current scenario. Therefore, imagination acts in hypothetical reasoning are used to guide our actions and knowledge in the real world; thus, in this setting, imagination is highly constrained by reality-oriented rules and facts. Coming up with unexpected plot twists, or evolving the scenario in ways it would seldom do, falls outside the point of these kind of imagination acts.

When defining the Rhombus of Imagination for hypothetical reasoning, we identify the following patterns:

1. Hypothetical reasoning requires the hypothetical scenario to be clearly set before entertaining it. If one wants to evaluate what next move to make in, say, a chess match, or if one wants to perform a specific thought experiment, then the initial premise defining such scenario is highly relevant (be them the distribution of the pieces over the chess board, or the premises defining the thought experiment). Someone may argue that this premise should be equally relevant for any imagination act, as this is precisely what creates the imaginary scenario as it is. Nevertheless, hypothetical reasoning highly relies on the very specific set of conditions defining the scenario to be evaluated, as the consequences of its evaluation are expected to affect the agent's behavior in the real world; a different or less carefully stated set of premises would define a different scenario, in which case the hypothetical reasoning may result in misguiding what the outcome would be if it was real. We consider that hypothetical reasoning makes a *high use* of the *Initialization* process.

2. When entertaining an hypothetical scenario, one is usually concerned with what would or could happen in that scenario: counterfactual reasoning is typically used to guide our own knowledge and actions in the real world. However, not all uses of hypothetical reasoning concern what would happen in a specific situation. For instance, when imagining how the staging of a theater play could be, or how one could decorate the room of a friend for a surprise birthday party, hypothetical reasoning allows us to foresee *how* a scenario would look like, rather than what would happen in there. In this sense, hypothetical reasoning often makes a *high use* of the *Description* process as well[3].

3. One of the most common uses of hypothetical reasoning usually involves predicting or foreseeing which actions would someone take in an alternative scenario. Both when wondering how one would react in a zombie apocalypse, or when playing a chess match and trying to foresee our opponent's next moves, hypothetical reasoning is used to predict how a situation could typically evolve; in other words, it is used to predict what would usually happen in the scenario we are entertaining, if it was real. Thus, hypothetical reasoning also makes a *high use* of the *Scripted Evolution* process.

4. Since now, hypothetical reasoning has been using each of the previous step at the

[3] As we already pointed out, this fact depends on the specific hypothetical reasoning being carried out. We are identifying classes or kinds of acts of imagination, and therefore collapsing some differences between particular executions into a more general group.

highest grade; this, at least, means that hypothetical reasoning is a quite demanding imaginative exercise. However, things are about to change. When engaging in hypothetical reasoning, we are not (or should be not) interested in voluntarily adding facts to the hypothetical scenario that would hardly follow, if it was real. We may be interested in following the scenario in its scripted evolution, or maybe even in resetting the whole hypothetical scenario and entertaining a different one (described by a different set of initial premises), but we should not allow our imagination to add things to the imaginary scenario that would never occur in it, regarding our set of reality-oriented facts and scripts. Say, if one is imagining how the setting for the next theater play could be, it would be useless to imagine that one has available a set of hoovering platforms sustained by their own gravitational field: it may be fun to imagine it, but it would surely be a waste of time when trying to figure out how the setting could actually be. Similarly, when evaluating my opponent's next moves in a chess match, it would be totally useless if I imagine my opponent taking my king, swallowing it and claiming that she is now the winner of the match (in virtue of the "Royal Banquet" rule). Therefore, we consider that hypothetical reasoning make a *marginal use* of the *Unscripted Addition* step[4].

Following the previous considerations, the qualitative Rhombus of Hypothetical Reasoning is as shown in Figure 6.6. As it can be seen, this sort of imagination acts make extensive use of three of the processes identified in the proposed Common Frame for Imagination Acts, but it completely neglects the fourth one, which is consistent with the fact that their goal is to provide trustworthy information based on reality-oriented rules.

Thought Experiments: A Remark

In the previous section, we have discussed and analyzed the way hypothetical reasoning uses each different process of the Common Frame, but without saying anything about the overall goal of the specific hypothetical reasoning imagining. As a result, we have considered that, in hypothetical reasoning have, the use of the *Description* and the *Default Evolution* processes are equally important, depending on our particular goal. However, the reader may be concerned with the fact that, in this general analysis, the processes we have willingly distinguished from the reality-oriented development do, in fact, coincide when analyzing their relevance.

[4] Note how changing the hypothetical scenario in order to evaluate a different one would not be an Unscripted Addition, but rather to simply toss away the current one and initiate a new, different scenario with a new initial premise (that may be similar to the one used, but which is different nonetheless).

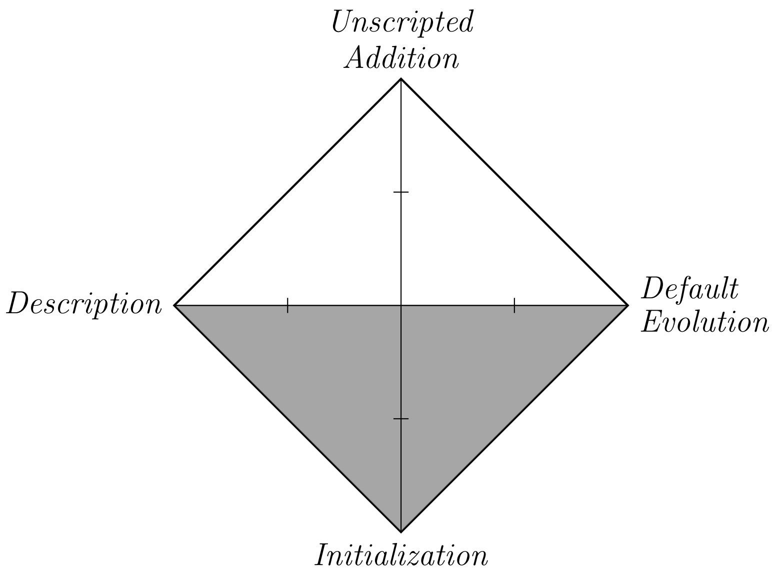


Figure 6.6: The Rhombus of Hypothetical Reasoning.

In order to show how this must not be necessarily the case, and thus in order to show how the distinction we made is indeed legit, let's consider the Rhombus of Imagination for a particular sub-kind of hypothetical reasoning imaginings: thought experiments.

In a nutshell, thought experiments are particular cases of hypothetical reasoning in which one sets an initial scenario described by certain initial premises, and then derives the consequences of such premises in order to foresee how the resulting scenario would be, if it was real (see [9] for an extensive introduction).

One of the earliest thought experiments in the literature (see [9] for more details) was proposed by Lucretius in his *De Rerum Natura* in order to prove that the universe was infinite. The thought experiment goes like this:

1. The universe is either infinite, or it is not (i.e.; it is finite)
2. Let's assume the universe is finite.
3. If the universe was finite, then there would be an edge of the universe.
4. If someone went to the edge of the universe and threw a spear at it, two different things could happen:
   (a) The spear would bounce back.
   (b) The spear would go through the edge of the universe.

5. From each possible course of action, we can derive:

   (a) If the spear bounced back, then it would mean that the edge of the universe is the limit beyond the universe and something else being behind it.

   (b) If the spear went through, then it would mean that there is something behind the edge of the universe.

6. In both cases, we can infer that, no matter what happened with the spear, there would be something beyond the edge of the universe, contradicting the assumption that the universe is, in fact, finite.

7. Therefore, we can conclude that the universe is *not* finite, and so that it must be infinite.

Note how, in this case, we make a negligible use of the *Description* process (we infer that there is an edge of the universe), whereas we make a high use of the *Default Evolution* process (regarding how the spear would behave, and what it would imply); regarding the *Unscripted Additions*, we also can consider that we make a negligible use (besides putting ourselves at the edge of a universe with a spear in our hands). As thought experiments are often concerned with the consequences of certain scenarios, the process they rely on the most is the *Default Evolution*.

Therefore, we can represent the Rhombus of Imagination for thought experiments as in Figure 6.7; note how, being thought experiments a particular sub-kind of hypothetical reasoning, their rhombus is different to the one of hypothetical reasoning (in Figure 6.6), but it is nevertheless *contained* in it.

Note how, if we had not distinguished between the *Description* and the *Default Evolution* processes, we would be unable to distinguish the way thought experiments, in particular, and hypothetical reasoning, in general, differ. Moreover, we claim that, if we stick with the mechanisms identified by the theories of imagination reviewed in Chapter 2, distinguishing these different kinds of imagination acts is not possible, at least regarding their structural composition of processes involved in the act of imagination itself.

This example, therefore, reinforces the fact that the distinction we make between the processes involved in any imagination act is, with respect to the other theories of imagination, more precise, and provides a better setting for a suitable analysis of imagination acts from a dynamic perspective.

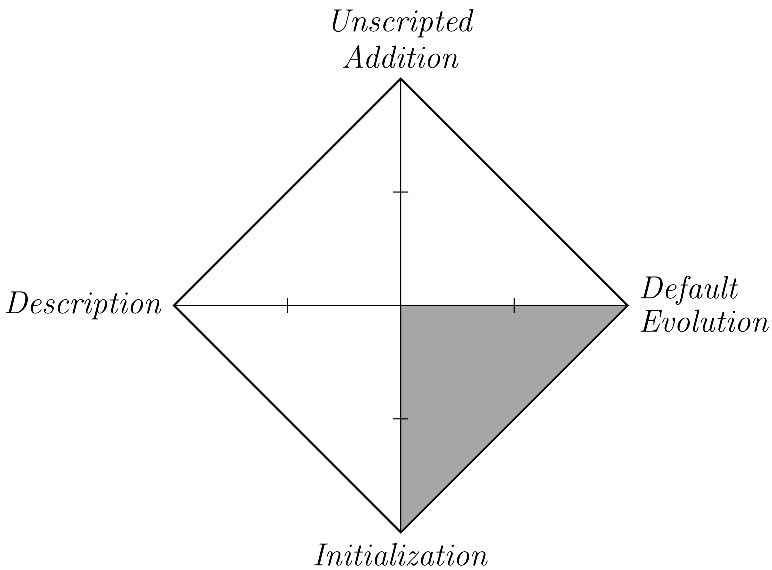


Figure 6.7: The Rhombus of Thought Experiments.

Story-Telling and Theater Improvisation

Things change dramatically, though, when considering other kinds of imagination acts. When engaging in, say, story-telling, or theater improvisation, or even in pretense play, the aim of such imagination acts often lies beyond entertaining scenarios that develop as they would, if they were real, but rather in creating fictional scenarios where unexpected and amusing things happen.

Reality-oriented rules and facts should be also present in order to build a common grounds in which the imagining takes place; specifically, in virtually every fictional story it is assumed that most physical laws remain intact, that people can still speak and communicate, that water can be boiled, and so on. However, once the scenario has been set, detailed and put in motion, we need something else for these imaginative episodes to be genuinely interesting. A movie in which the main character is a banker, and which just shows what usually happens in her daily routine, would hardly make for a very interesting one.

Therefore, we need imagination to come up with unexpected twists, and funny or tragic situations that does not naturally follow from the imagined scenario. The Rhombus of Imagination for story-telling can be defined by the following patterns:

1. As argued in the hypothetical reasoning case, every act of imagination makes use of this step: after all, it is through it that the imaginary scenario is created. However,

the initial conditions describing an imaginary scenario are not usually as important when engaging in story-telling as they are in hypothetical reasoning; at least, the consequences of initiating the scenario with a less strict set of initial premises are not as severe (for the result of the imagination episode in the real world) in the former case than they are in the latter. Furthermore, most episodes of story-telling, theater improvisation or pretense play are typically initiated by a very small set of initial conditions, such as "we are having a tea-party", or "I am the waiter of a fancy restaurant", or "it was a bright cold day of April, and the clocks were striking thirteen". Thus, we consider that story-telling makes a *standard use* of the *Initialization* step.

2. Embellishing the initial scenario with details regarding how the scenario looks like is usual in story-telling or theater improvisation, for instance. This process is responsible of describing what elements there are in our tea-party, of describing how am I dressed as a waiter, or what else in there, in that cold day in April. Therefore, this kind of imagination acts makes a *standard use* of the *Description* step as well.

3. Similarly, allowing the imagined scenario to evolve in ways it would typically do is something usual in story-telling and theater improvisation, even if at least at the beginning of the story. One would not typically start enacting the scenario about the restaurant waiter by pretending to shoot a laser-ray gun to one of the diners (at least not before asking whether they would like to eat or drink something first: it would be rude). As we have already said, the *Default Evolution* is not what characterizes this kind of imagination acts, but nevertheless it is present, and we typically make use of it throughout the imaginative scenario (at least regarding how physical laws would behave in our story, for instance). Thus, we also consider that this kind of imagination acts make a *standard use* of the *Default Evolution* step.

4. Good stories usually tell us something unexpected, funny, interesting, mysterious or tragic. As we have previously argued, the daily life of many of us is not usually very exciting (or, at least, not exciting enough to be told by a story-teller, or enacted by theater actors): we need something more to happen in order to catch our interest. These unexpected twists (say, an everyday person working in the IT department of a company and accidentally discovering, in an online forum, an undercover plot to assassinate her best friend) are seldom considered "usual ways" a scenario would typically evolve, and so they fall under the scope of the *Unscripted Additions.* As

we have already argued, this is probably the most characteristic part of this kind of imagination acts, and thus we assign a *high use* to the *Unscripted Additions* step.

The rhombus that represents imagination episodes involving story-telling, theater improvisation or pretense play is as shown in Figure 6.8. As we can see, and when comparing it to the rhombus in Figure 6.6, both rhombus are dramatically different. The rhombus of story-telling has a highly-pointed vertex towards the *Unscripted Additions*, whereas the rhombus of hypothetical reasoning is completely flat on that side. Conversely, the rhombus of hypothetical reasoning makes a high use of the processes that rely on reality-oriented facts and scripts, whereas their use is lower in the rhombus of story-telling.

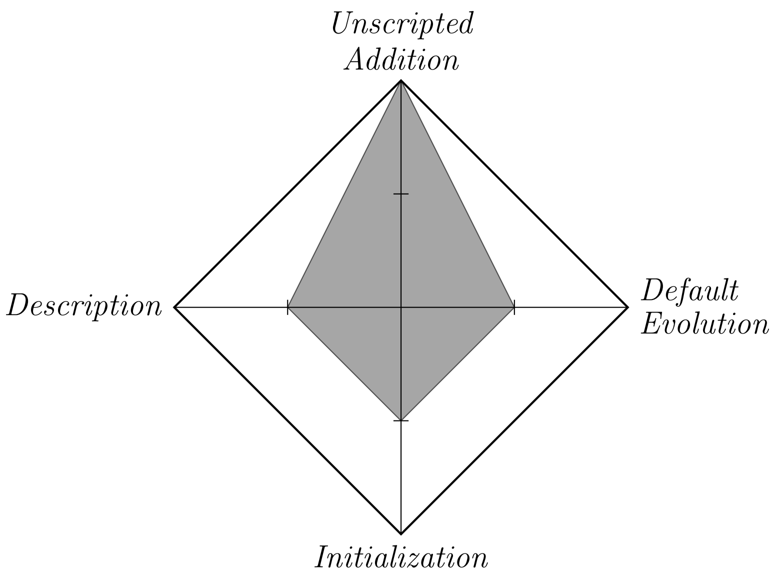


Figure 6.8: The Rhombus of Story-Telling.

Engaging in a Preexisting Fiction

There is yet another kind of imagination acts that we want to analyze and represent by using the Rhombus of Imagination: engaging in an already existing fiction, say, by reading a novel, or by listening to someone telling a story[5].

[5] We omit watching a movie in this group for various reasons. As the steps we identified in the Common Frame refer to the processes aimed to embellish the imaginary scenario, or to imagine how it would normally evolve, movies already take care of most of these processes. It is true that a spectator can (and most of the time do) imagine things beyond what is shown in the screen, but we think that the cases of the agent reading a novel or listening to a story are more relevant for the point we want to make regarding this kind of imagination acts.

The interest beyond this kind of imagination act lies in the fact that, while being related to story-telling as well, the role of the agent changes dramatically, depending on whether the agent is the one telling the story (or writing the book), or listening to it (or reading it). Whereas in story-telling the agent is the one who creates the story, and thus has to use plenty of her own imaginative resources to tailor it, when the agent becomes the reader of an existing fiction (or the hearer of someone in the role of the story-teller) her imaginative resources are highly guided by the story being told. The initial conditions of the imaginary scenario are being given, as well as its description (or part of it) and most of the things that happen within the story.

The patterns followed when engaging in a fiction, and that determine the shape of the Rhombus of Imagination for this kind of imagination acts, are as follows:

1. The initial conditions of the imaginary scenario are described by the story-teller the agent is listening to (or the book she is reading). We argue that, as it happens when considering hypothetical reasoning, this step is crucial when engaging in a preexisting fiction in order to properly follow the story afterwards. How many times have we heard (or asked), while watching a film, "why does the character do that?", and have been answered with a "because such and such happened earlier, don't get distracted or you'll miss the plot!". The premises describing the scenarios involved in the story are being fed into our imagination in a way that allows us to imagine and follow what happens in them. Therefore, we consider that this step is of great importance in this sort of imagination acts, and we consider that engaging in preexisting fictions involves a *high use* of the *Initialization* step.

2. Although many details as to how the scenario looks like can be fed to by the story-teller or the book, many details are often left unspecified. The story-teller may tell us that the hunter stops by a huge tree, higher than any other, and dense with foliage: however, the way we picture the tree in our mind typically has a lot more detail than that. If we form the image of the tree (or describe it in terms of propositions), we may assign to it a certain height, a certain foliage density, the leaves' shape and color, etc. All these unspecified details, although being prompted by a certain set of initial details, fall within the second process of our Common Frame. Thus, we assign a *standard use* of the *Description* step to this kind of imagination acts.

3. A similar thing happens with the way an imaginary scenario would usually evolve. While listening to a story-teller, we do not imagine the characters doing what we

think they would usually do, but rather doing what the story tells us they do. We may be surprised (and we may even think "I would never do that", specially when a character steps outside a forest cabin in the middle of the night after hearing a snarl by the window), but we do not "force" the characters in our imagining to do something different as we are being told, even if it is unexpected. We do, however, assume certain "off-story" actions: for instance, when the story tells us that a co-worker of the main character dies, we might automatically imagine that the other co-workers are devastated, even if the story does not explicitly tell us so. Similarly, we imagine that the main character is in pain when she is shot in the leg, even if the story does not explicitly states that "the main character is in pain". Due to this, we consider that this kind of imagination acts do make a *standard use* of the *Default Evolution* step as well.

4. Regarding the last process, its use is practically negligible in this kind of imagination acts. As we already mentioned in the previous step, when listening to or reading a story we do not imagine that the characters do something different as what we are being told, nor we imagine that the story goes in a different way, or that it takes place in some other scenario that the one we are being told. We may imagine what would have happened *if* the story had gone different, but this would initiate another different imagination act in which we are no longer the spectators of a story, but rather its tailors, thus leading to one of the two previous kinds of imagination acts. Therefore, we consider that engaging in a preexisting fiction makes a *marginal use* of the *Unscripted Additions*.

After the analysis of this kind of imagination acts, the resulting rhombus is the one represented in Figure 6.9.

There are some interesting things we can say when comparing this rhombus to the previous ones; in particular, neither the rhombus of hypothetical reasoning (in Figure 6.6) nor this one make any use of the *Unscripted Additions*. The reasons for this are that both uses of imagination are constrained by a very specific set of rules and facts, and the aim of both imagination acts lies on developing the imaginary scenario (be it reality-oriented or not) by following these rules and facts. When engaging in hypothetical reasoning, we draw from our own knowledge and beliefs about reality in order to unfold the imaginary scenario; when engaging in a fiction, we are given the way an imaginary scenario advances, and we only make a light use of our knowledge and belief to embellish it (this is why the rhombus of engaging in a preexisting fiction makes a lighter use of the

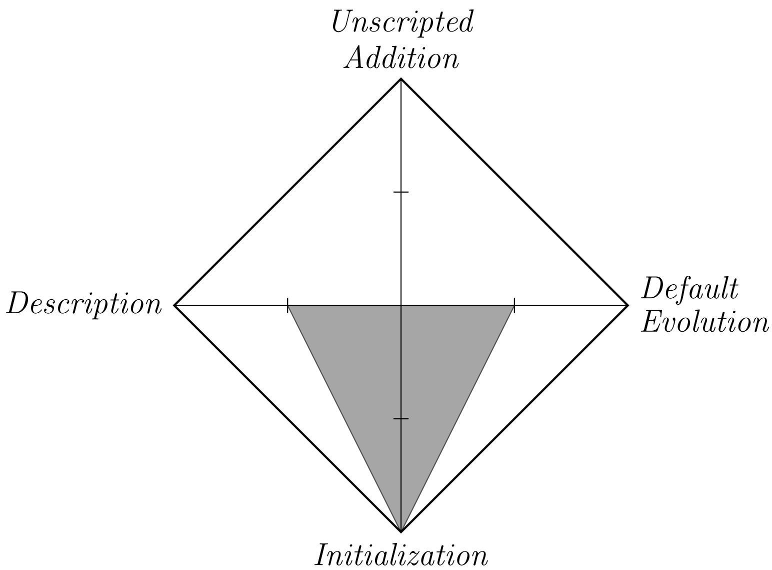


Figure 6.9: The Rhombus of Engaging in a Preexisting Fiction.

Description and the *Default Evolution*); however, in neither case we intentionally deviate the scenario from what these rules and facts tell us.

Regarding the rhombus of story-telling or theater improvisation (in Figure 6.8), we have already mentioned that their differences result from a "role-shifting": being the one who comes up with the story requires a quite different use of one's imaginative resources than being the one who listens to it. This gets represented by the fact that both rhombuses have little in common, specially regarding the *Initialization* and the *Unscripted Additions*.

As we have just seen, the Rhombus of Imagination can indeed be used as a valuable tool for analyzing and classifying imagination acts into different kinds, according to how much they rely on each process identified by the Common Frame for Imagination Acts. Nevertheless, the Rhombus of Imagination has still more to offer. The level of detail in which we can split the processes occurring in acts of imagination allows us to identify what is happening, at any point, when an agent is elaborating on an imagining. Now, if we can identify what particular process is taking place at any point, we should also be able to count how many occurrences of each process have taken place in that act of imagination. Could we use the Rhombus of Imagination, in that case, to generate not a general blue-print for a kind of imagination acts, but rather a unique blue-print for *that* particular instance of imagination act? It turns out that, indeed, we can.

In the next section, we tune up the Rhombus of Imagination a little bit, and we show how its modified version can be used to compute the blue-print of any instance of imagination act, as long as we can access to the details of the processes that took place within it.

6.2.2 Analyzing Specific Acts of Imagination

In order to use the Rhombus of Imagination for this purpose, we need to consider a particular instance of an act of imagination, about which we know all the processes that have been carried out in it; that is, how the imagining has been initiated, how it has been further developed (both statically and dynamically), and whether further premises have been added later on. Once we have all this information, we can use the Rhombus of Imagination to perform a *quantitative analysis* of the weight each process has had into that particular imagination act.

In order to do this, we need to slightly modify the Rhombus of Imagination by getting rid of the measures we added to each line. In this version of the Rhombus, we no longer want to attribute a qualitative label to certain sections of each line, but we rather want to use them as a sort of "scale" for measuring the weight of each process within the imagination act. The quantitative version of the Rhombus of Imagination, therefore, is as shown in Figure 6.10; note that the only difference with respect to the qualitative version of the Rhombus, in Figure 6.5, is that we no longer split the lines in the present case.

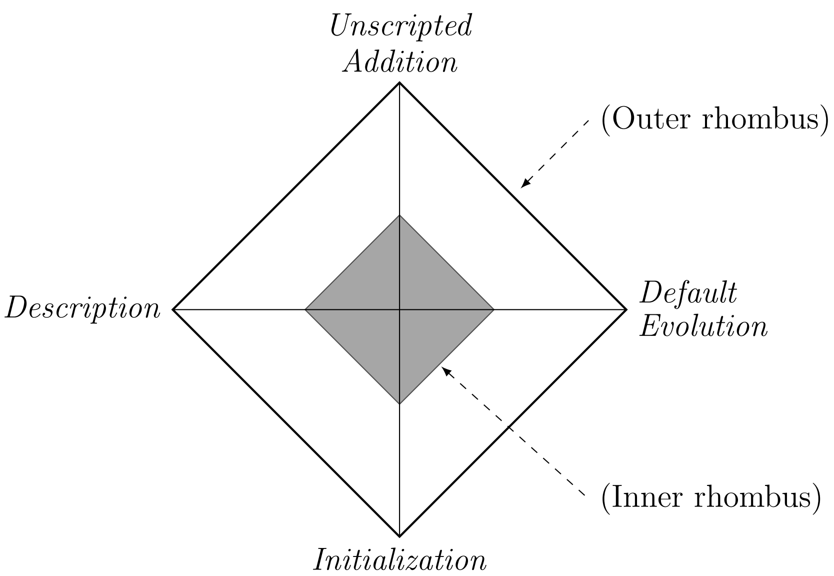


Figure 6.10: The (empty) Rhombus of Imagination.

Determining the weight of each process within the whole act can be done using Formula 6.1. We use $p \in \{\text{init}, \text{descr}, \text{evo}, \text{add}\}$ to refer to one of the processes of the Common Frame; then, we use $w(p)$ to denote the *weight* of process p, which measures the relevance of this process within the whole imagination act. Similarly, we use $\#(p)$ to denote the *number of occurrences* of process p, corresponding to the number of times process p has been called within the imagination act:

$$w(p) = \frac{\#(p)}{\#(\text{init}) + \#(\text{descr}) + \#(\text{evo}) + \#(\text{add})} \tag{6.1}$$

Once we have calculated the particular weight of each process $w(p) \in [0, 1]$ (that is: $w(p)$ is a real number between 0 and 1), we can draw each process' weight into its corresponding line, and then draw the *inner rhombus* that results from drawing the shape that follows from connecting each weight, starting at the *Initialization* line and following the rhombus in a clock-wise fashion until reaching the starting point again.

Now, in order to show how the quantitative version of the Rhombus of Imagination works, let's follow an example of one specific imagination act described in detail, about which we afterwards provide its blue-print.

In the following lines we provide a detailed description of an example of an act of imagination, in which the processes that take place are clearly identified. After the agent finishes entertaining this particular imaginary scenario, we count how many calls to each process the agent has done, compute their weight into the overall scenario, and then draw the rhombus of that particular act of imagination. Obviously, the example we provide in here could have gone in a different way, and could be much longer: this is only for illustrative purposes.

Imagining a Dystopian Future

In this imagining, the agent fantasizes about how the world would be[6], if it was a kind of dystopian future setting governed by despotic machines.

1. *Initialization*: "the world is in chaos and machines have taken over the control"; the agent initiates an imagining with the initial premise.

[6] Note that, following the considerations we make in Section 6.1.6, this example depicts the agentiveness of both the *Description* and the *Default Evolution* processes. Not only the agent chooses to use a particular rule or script, which develops the scenario in certain ways and opens certain possible outcomes, but she also chooses, when there is more than one possible outcomes available (as in step 8), which one she wants to follow in her imagining. We willingly want to account for imagination acts that work in this way, and so we want our theory of imagination acts to identify, account for and capture this feature.

2. *Description*: "if the world is in chaos, then everything is in ruins"; the agent elaborates on the static details of the scenario by taking the antecedent from the premise clamped in step 1.

3. *Description*: "if everything is in ruins, people live hiding in abandoned buildings"; the agent elaborates the scenario by drawing on the consequences of step 2.

4. *Description*: "if machines have taken over, they must be hostile towards human beings"; the agent elaborates by following, this time, another one of the premises clamped in step 1.

5. *Description*: "if machines were threatening humanity, there would be some kind of resistance group"; the agent elaborates on the consequences derived by step 4.

6. *Unscripted Addition*: "I am a scout of the resistance group"; the agent decides to add a new premise into the imagining, which does not follow any specific rule or script.

7. *Default Evolution*: "if I was a scout of the resistance, I would infiltrate the headquarters of the machines": the agent determines a course of event that would take place in the imagining, based on the antecedent of the premise she added in step 6 and using a script detailing what would happen, given that antecedent. Note how this script needs not be related to the actual world, but rather to a paradigmatic cliche from science-fiction works.

8. *Default Evolution*: "if I was to infiltrate the headquarters, I could be caught or I could succeed": the agent imagines two possible courses of action, according to the action she imagined in step 7.

9. *Description*: "if I was infiltrating the headquarters of the machines, I would have some sort of weapon to fight them"; the agent elaborates on more static details of the scenario by following what results from step 7.

10. *Default Evolution*: "if I succeeded in my infiltration, I would use my weapon to destroy the main computer"; the agent takes one of the two courses of action described in step 8 (note that this implies some kind of decision made by the agent to choose one instead of another) and determines what would happen after it.

11. *Description*: "once the main computer had been destroyed, the world would no longer be in chaos"; by following the action described in step 10, the agent derives a static detail about the scenario, which overrides a part of the premise clamped in step 1.

12. *Unscripted Addition*: "instead of celebrating, I would disappear in the shadows like a legendary *vigilante*"; the agent adds a new premise determining a new course of actions she would follow, regarding the description of the current imaginary scenario given by step 11.

At this point, the agent finishes her imagining. The imaginary scenario has been created, developed in different ways[7], and finally it has come to an end. It is undeniable that the previous example could have been expanded in many different ways, enriched with much more details, given many more possible courses of action, and so on; needless to say, it is pretty simplified, when compared to the kind of imagining we usually engage in. Nevertheless, it is still enough to illustrate our purpose.

Now, let us compute the weight of each process within this imagination act, and draw its own rhombus using the Rhombus of Imagination; the number of occurrences of each process is equal to the sum of the corresponding steps in the previous example:

- Weight of the *Initialization* process: 0.083.

$$w(\text{init}) = \frac{1}{1+6+3+2} = 0.083$$

- Weight of the *Description* process: 0.5.

$$w(\text{init}) = \frac{6}{1+6+3+2} = 0.5$$

- Weight of the *Default Evolution* process: 0.25.

$$w(\text{init}) = \frac{3}{1+6+3+2} = 0.25$$

[7]Note that we are not concerned, at any moment, with the origin of the rules used by the agent. As noted, for instance, by Langland-Hassan in [35] (page 72), the rules used by the agent to develop an imagining may either be about the real world itself, or about certain paradigms and fictions detailing how they are usually represented. For example, in the step 2 of this example, the agent associates chaos with ruins, maybe because of what she knows about wars within the history of the world, or maybe because of certain kind of movies or fictional works; the same goes for other rules elaborating on the static description of the scenario, and also about the default scripts, such as the one in step 7.

- Weight of the *Unscripted Addition* process: 0.17.

$$w(\text{init}) = \frac{2}{1+6+3+2} = 0.17$$

Now, the blue-print for the previous act of imagination, which we call the "Dystopian Future" is as shown in Figure 6.11. Note how, in this case, the agent has focused mostly on the static description of the scenario (either by imagining how the world would look like, or by imagining what else would be the case in there).

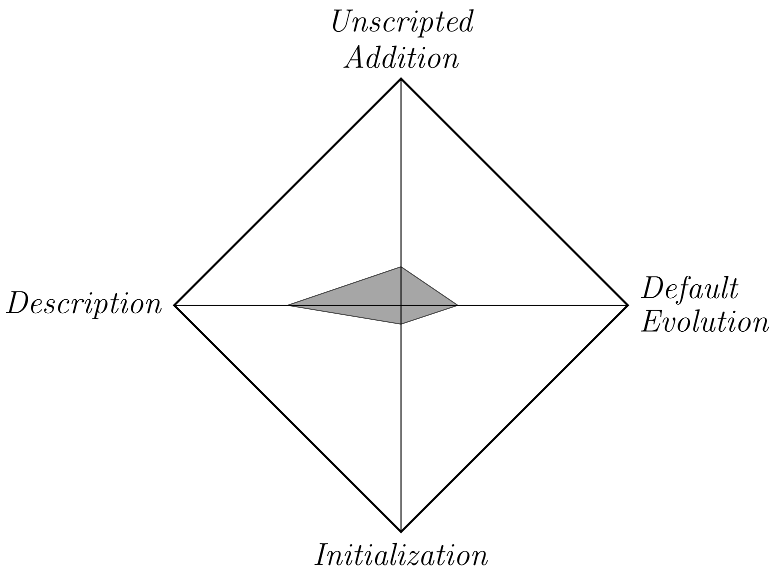


Figure 6.11: The Rhombus of the Dystopian Future.

Even though Figure 6.11 shows the rhombus of that specific imagination act, as we wanted, the overall surface of the inner rhombus, with respect to the outer rhombus, is rather small. In fact, and taking into account that each weight represents its relation among the total sum of all weights, we know that the inner rhombus will always be sort of small, in comparison to the outer one. Nevertheless, and as we are interested in the shape of the blue-print, rather than in its size, we can scale up the results in order to account for a bigger, more readable version of the inner rhombus.

What we do is simply to take the highest of all weights, and associate it to 1; this defines a constant representing the ratio between the highest weight and 1, which we call the *proportion constant*, and we refer to it by k. Then, the rest of the weights should be computed following the same proportion rule. Therefore, in this case we get:

- Proportion constant $c = 2$.

$$0.5x = 1 \Rightarrow x = 2$$

- Expanded weight of the *Initialization* process: $0.083 \cdot 2 = 0.16$.
- Expanded weight of the *Description* process: $0.5 \cdot 2 = 1$.
- Expanded weight of the *Default Evolution* process: $0.25 \cdot 2 = 0.5$.
- Expanded weight of the *Unscripted Addition* process: $0.17 \cdot 2 = 0.34$.

Now, the Rhombus of the Dystopian Future, when using the expanded weights for each process, is as shown in Figure 6.12. Note how, by using the expanded weights, it becomes much more easy to spot, at a single glance, how strongly this act of imagination relies on each of the processes involved in it.

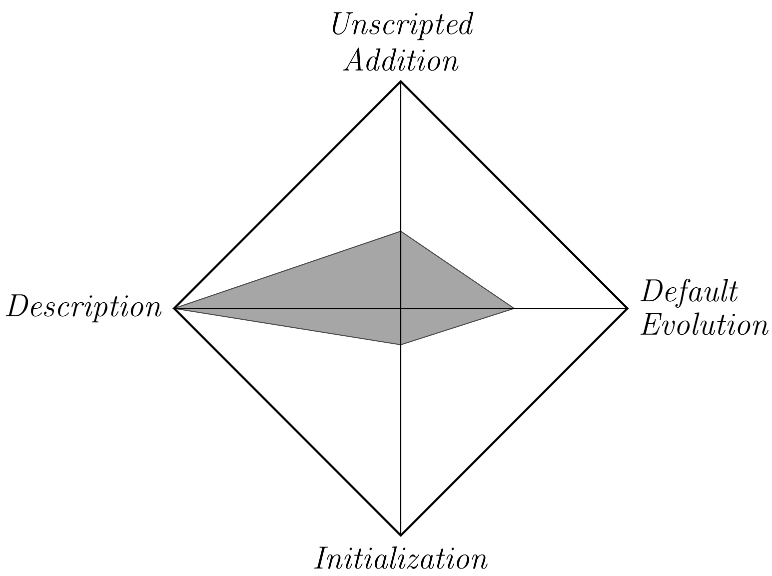


Figure 6.12: The Rhombus of the Dystopian Future using expanded weights.

The previous example shows that the Rhombus of Imagination is a suitable measuring tool to be used in a systematic analysis of specific imagination acts. Moreover, when combining the qualitative and quantitative version of the Rhombus of Imagination, we end up with an analysis tool that can be used both when we have access to all the structural details of a specific imagination act, and also when we consider kinds or classes of imagination acts in a more general way. Having a way of characterizing both general and particular imagination acts by a certain shape, which provides useful information at a single glance, is a valuable contribution towards acquiring a more precise and more analytic understanding of the dynamics of imagination.

6.2.3 The Rhombus as an Analysis Tool for Imagination Acts

After having introduced the Rhombus of Imagination, and the different ways in which it can be used as a tool for systematically analyzing imagination acts, we want to show how studying the dynamics of imagination in this way provides valuable insights into the way imagination acts work. In particular, we want to point out to certain regularities we can identify in imagination acts, according to the results provided by the examples considered in the previous sections.

1. As we have seen when using the Rhombus of Imagination, *every* imagination act, be it a particular one or a general kind, makes use of the *Initialization* process. No matter how the imagining develops afterwards: each and every act of imagination is initiated by a mental action of the agent that creates an imaginary scenario described by a certain initial premise[8]. Therefore, we say that imagination acts are bound under a *necessity of initialization*.

2. Moreover, there is still something else we can say about the *Initialization* process. Not only is it *necessary* to happen in order to initiate an imagining, but it is also required to happen exactly *once*. Every imagining needs to be initiated, but it can only be initiated by a single execution of the *Initialization* process. Note how adding further premises to an already existing scenario is *not* handled by the *Initialization*, but rather by the *Unscripted Additions*. Even if we "modify" the initial premise of an imagining, we either do it through adding new premises (which take an already existing imaginary world as the world of reference, and thus is considered an *Unscripted Addition*), or by tossing out the first imagining and creating a new one from scratch; this uses the *Initialization* process indeed, but it uses it to create a *different* imagining. Therefore, we say that imagination acts are bound under a *singularity of initialization*.

[8] Although we focus on voluntary acts of imagination in this work, it can also be argued that involuntary acts of imagination (that is: those imaginings that pop out into our minds all of a sudden) are also initiated by a mental action of the agent, although not voluntary in those cases.

3. Imagination is never fully separated from reality: we do, after all, guide our imaginings (or parts of them) through reality-oriented rules and scripts. Therefore, we can say that every act of imagination makes a certain use of the *Description* and the *Default Evolution* process in one way or another. It is worth noting that, although we have distinguished both processes in the Common Frame, the *Description* and the *Default Evolution* processes are strongly interrelated, because they are both intended to reflect what would be the case in the real world. Even if we are just interested in what would typically happen in an imaginary scenario, the way the scenario evolves can affect how the scenario is, or what things are true in that scenario (meaning that the *Default Evolution* process may also put in motion the *Description* one); for instance, if an actor involved in a restaurant scenario pretends to play the piano, then the description of the scenario would be updated to include such instrument. Similarly, by elaborating on the details regarding how a imaginary scenario is, one can then realize that, if such and such things are in the imaginary scenario, then certain actions are more likely to happen; say, if I imagine that there is a cake in the tea-party scenario, then the action "someone eats the cake" would become part of the default set of actions that could likely happen in there. Therefore, we say that imagination is always characterized by a *closeness to reality.*

4. We have just stated that imagination never completely departs from reality, and so our knowledge or beliefs about the real world inevitably affect our imaginings. What happens with *Unscripted Addition* process is quite the contrary. As we have seen when characterizing, for instance, the Rhombus of Imagination for hypothetical reasoning, or for engagement in preexisting fictions, it is possible to carry out an act of imagination without coming up with any kind of unscripted twists or original ideas. We can say, in this sense, that imagination is *always* grounded on reality, but, conversely, that it only departs from it on occasions (usually in many occasions, but still not always). Regarding this, we say that imagination is *occasionally fictional* (keeping in mind that we mean not always).

Table 6.2 summarizes the previous properties of imagination acts.

By taking the Common Frame for Imagination Acts as our underlying theory, and by using the Rhombus of Imagination as a support tool for analyzing both general and specific imagination acts, we have been able to provide valuable insight regarding the dynamics of imagination, both in the form of blue-prints for imagination acts, and in the form of four regularities we identify in them. This, therefore, proves that the Common Frame for

Property	*Expresses that*
Necessity of initialization	Every imagining must be initiated by a voluntary action
Singularity of initialization	Every imagining can only be initiated by a single voluntary action
Closeness to reality	Reality-oriented rules and scripts always determine some part of the imagining
Occasionally fictional	Imaginings do not necessarily require new voluntary additions of premises

Table 6.2: Properties of imagination acts, as recognized by the Rhombus of Imagination.

Imagination Acts and the Rhombus of Imagination are, indeed, a valuable contribution towards a better and deeper understanding of the dynamics involved in imagination.

Throughout the rest of this work, therefore, we take the Common Frame for Imagination Acts as our underlying theory, and we use it to define, in the next chapter, a dynamic logic system aiming to represent how an agent creates and elaborates on an imaginary scenario.

Chapter 7

The Logic of Imagination Acts

In the present chapter, we come back to the topic of formal logic, but with a fresh idea in mind. After defining the Logic for Imaginary Scenarios in Chapter 4, and arguing how the approach we followed in there was not precise enough to reach the level of detail with which we want to represent voluntary imagination acts, we revisited the theoretical foundations of our work. Then, we have defined the Common Frame for Imagination Acts, which captures the dynamics of imagination in more detail than the already existing theories, and we have argued how this theory should be the underlying frame for a new logic aimed to capture imagination acts.

In the next pages, we define the Logic of Imagination Acts, a new formal system that allows us to capture the dynamics of imagination acts in a way which is much more precise and modular than our previous formal system. This new logic takes the Common Frame for Imagination Acts as its underlying theory and is defined in such a way that captures and distinguishes the different processes identified by it. After defining this new approach, we discuss its strong and weak points.

7.1 Towards a New Perspective

As we have already seen, our first attempt to represent acts of imagination through a single algorithm (as part of the logic defined in Chapter 4) ends up being too simplified.

Handling an imagination act through a single execution of a single algorithm reduces imagining to a matter of computing every possible combination of truth-values of a set of atomic formulas involved in an initial premise; then, the development of the imaginary world is just based on importing any atomic formula that is true at the world of reference, and that does not contradict what has already been set in the new imaginary world. At this point, the act of imagination is considered finished: if the agent wants to further elaborate on the imagining, she must do so by voluntarily adding new premises, and thus by beginning the same procedure again. As we have already explained in Chapter 5, we miss both the notion of rule-based elaboration, in which the imaginary worlds are shaped according to certain rules and scripts, and the implicit agentiveness that hides behind choosing which rules will be used to elaborate the scenario.

The way we elaborate our imaginings is way more complex than just setting the truth-conditions of the initial premise and importing any other factual detail from reality. We want our logic to account for that: to account for our agent following certain factual rules and scripts, and for the resulting imaginary worlds to get more and more elaborated without the need for voluntarily clamping new premises into them.

This change in the underlying analysis of imagination acts, steamed by the Common Frame for Imagination Acts (introduced in Section 6.1), requires a drastic change in our formal approach. We are no longer interested in defining *one* algorithm capable of handling a whole imagination act; rather, we now want to capture *each* of the four processes identified by the Common Frame in distinct algorithms.

Before giving a formal definition of the logic, we present the intuitions that should be captured by each one of the algorithms that will now take part in an act of imagination.

7.1.1 Initialization: a New Imaginary World

This process is the responsible for *creating* new imaginary worlds from scratch, based on a certain world of reference representing the state of affairs of the real world, as believed by the agent, and by using a certain initial premise determining the content of the imagining to be initiated.

This process, therefore, is closely related to the way our `ImgAlg` works in the Logic of Imaginary Scenarios, defined in Chapter 4; beyond importing atomic propositions from the world of reference, there is no further "elaboration" of the imaginary worlds, and so the process we are currently defining still conforms to the one we followed before. Following the same intuitions, this process requires an initial premise δ that must be clamped into a new (or various) imaginary worlds; similarly, the state of affairs described by δ must be prioritized, as it is what the agent is voluntarily imagining. Once the new imaginary worlds satisfying δ have been created, the remaining details must be imported from the world of reference, as long as they do *not* contradict what has been explicitly set by evaluating premise δ; Figure 7.1 depicts the whole process. Once the new imaginary worlds are created and stabilized, this process is considered finished; further elaboration of the imaginary worlds requires, as pointed out by the Common Frame, using other processes. Note that importing details from the world of reference does not need to be seen explicitly as an elaboration of the imaginary world, but rather as a sort of "default way" of filling up the details that δ leave unspecified.

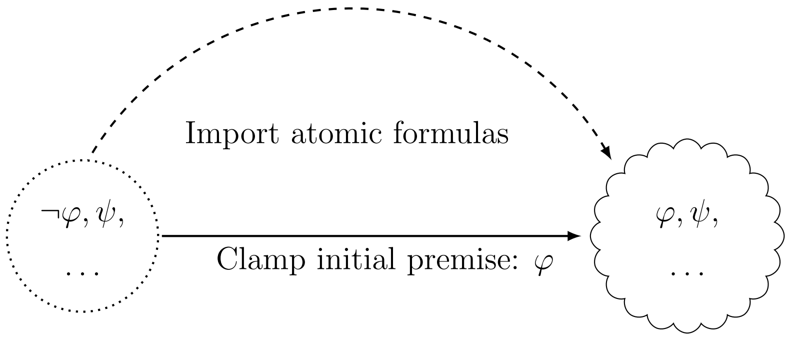


Figure 7.1: Initiating the imagining using an initial premise.

7.1.2 Description: Adding More Details

In order to capture the Description process of the Common Frame, our formal setting must be able to account for the notion of *factual rules*. Briefly recalling, factual rules are a kind of implication-like formula capturing a sort of hypothetical conditional of the form "if φ was the case, then ψ would also be the case". We want to use this kind of rules in the Description step to say, precisely, that if the imaginary world fulfills a certain condition, then it could also fulfill a certain outcome as well.

An example of this could be a very simple imaginary scenario initiated by the agent imagining that "it is raining", through an execution of the Initialization step. Now, the agent could browse through her beliefs on factual rules, and find a rule stating that "if it was raining, that I would be carrying an umbrella": as the current imaginary scenario fulfills the conditions that "it is raining", then the agent can further elaborate it by adding the fact that "I am carrying an umbrella".

This elaboration of the imaginary scenario, nevertheless, must be handled in a very specific way. We want our formal models to be able to keep track of the different imaginative processes the agent uses, and to show how they affect the elaboration of the imaginary worlds. Therefore, if we were to represent the application of the rule "if it was raining, I would be carrying and umbrella" by adding the fact "I am carrying an umbrella" to the world already satisfying the fact "it is raining", we would lose track, in the formal model, of *how* that world became a world in which I am carrying an umbrella. One might argue that we can add a sort of reflexive relation capturing the fact that I use this factual rule on the world itself, and so we can deduce that "it is raining" was previously the case at that world, and so by applying the factual rule we have come to add "I am carrying an umbrella". Nevertheless, if we keep applying more and more factual rules to the same world, we may soon lose sight of what was the case there in the first place, and what has been added later. Furthermore, we may even elaborate the world in ways that override something that we previously added: if we were to capture this by using reflexive relations and updating always the *same* world, we would soon be unable to tell what was initially the case in there, what we added later by using what rule, and what information may have been overridden by successive applications of different factual rules.

Taking these considerations, we want the Description process to be represented, in our formal models, as an accessibility relation linking two *different* worlds: the world of reference, where the antecedent φ of a factual rule $\varphi \to \psi$ holds, and a different imaginary world, similar to the world of reference in everything, except by the fact that ψ also holds at that new world as a result of applying that specific factual rule. In other words, this kind of formulas (and the way they are processed by our logical system) represents a kind of *Modus Ponens* rule in which the antecedent is evaluated at the world of reference, but the consequent is evaluated at a different accessible world[1].

In order to capture all these considerations in our formal setting, we will need a set of specific formulas aimed to represent those factual rules. Conversely to the initial premises we use in the Initialization step, however, this set of factual rules should *not* be formed by any arbitrary formula: after all, the agent can choose to imagine whatever she wants, but she does not choose to believe the rules that describe how an imaginary situation

[1]It is worth noting how this kind of hypothetical conditionals are similar to the kind of formulas in which David Lewis is interested in his work *Counterfactuals* (see Section 3.2.1). Nevertheless, the way Lewis uses them is different from the way we do: in his work, Lewis evaluates a formula of the kind $\varphi \mathrel{\Box\!\!\rightarrow} \psi$ at a world aimed to represent the real world, and operator $\Box\!\!\rightarrow$ moves the whole evaluation to an accessible counterfactual world in order to assess whether the conditional $\varphi \to \psi$ holds in there. His way of evaluating hypothetical conditionals, therefore, is by moving the whole conditional to an alternative world. Our way of understanding them, however, will be to asses whether the antecedent φ holds in the current world of evaluation, and then to determine that a *new* world, fulfilling ψ and defined by taking the current one as the reference, must be created. Our understanding of this kind of conditionals, therefore, will be used by our Description algorithm to determine the way a certain world could change, given the information provided by the specific formula being evaluated.

would be by following reality-oriented rules —as this rules are, precisely, determined by her beliefs about reality. This set of factual rules, to which we will formally refer as FACT, is therefore limited to contain a determined set of formulas representing the factual rules the agent believes in. Similarly, the form of such formulas should not be arbitrary; a formula like $\varphi \wedge \psi$ does not provide new information on *how* an imaginary scenario would be, but it rather states a fact which do not follow any kind of conditional rule (even if we could translate this formula into an equivalent conditional-like formula). Thus, we require the formulas in FACT to be conditional formulas *similar* to $\varphi \rightarrow \psi$. Why are we saying "similar", though? Because, although being a conditional operator, $\rightarrow$ does not really convey the sense of "modal conditional" we want to express with such formulas. Therefore, we define a new derived operator $\langle\rightarrow\rangle$ standing for this kind of modal conditional, and which we use to represent the formulas standing for factual rules in FACT. The formulas standing for the set of factual rules believed by the agent have the following form:

$$\text{FACT} = \{\varphi_1\langle\rightarrow\rangle\psi_1, \varphi_2\langle\rightarrow\rangle\psi_2, \ldots\}$$

The formal definition of the $\langle\rightarrow\rangle$ operator, together with certain restrictions we impose at both the antecedent φ_i and the consequent ψ_j of any formula in FACT, are detailed in Section 7.2.

The execution of the algorithm representing the Description process, therefore, will need a world of reference w^R and a certain formula $\varphi\langle\rightarrow\rangle\psi$ within FACT. Then, the algorithm must check whether the antecedent φ is true at w^R and, if it is, it must create a new imaginary world v, accessible from w^R, and in which ψ holds; the rest of atomic formulas determining the state of affairs of the new world v will be taken from w^R. In other words, the only changes that v will have with respect to w^R are those changes needed to make ψ true at v. Figure 7.2 represents how the Description process works in a formal setting.

7.1.3 Default Evolution: Moving Forward

The algorithm aimed to capture the Default Evolution process should behave similarly to the one capturing the Description process: after all, and although they are indeed different, both processes use certain kinds of rules to unfold the consequences that one could derive from an imaginary scenario.

The main difference between both processes, and so between both algorithms, is that

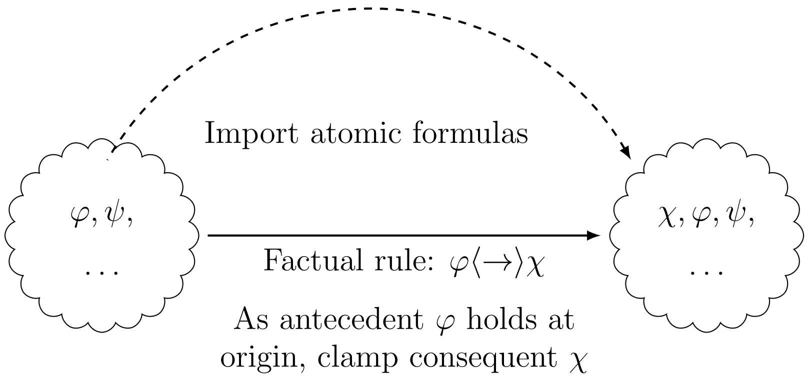


Figure 7.2: Follow a factual rule to elaborate on the static details.

the Default Evolution involves a certain *action* or *event* to happen, whereas the Description focuses on the relation between the static facts of the situation. In this case, therefore, the scripts that detail how an imaginary scenario could typically evolve or move forward must capture something like "if φ was the case, and action α happened, then ψ would also be the case". Therefore, and similarly to what happened in the case of the Description process, we want to capture the scripts the agent believes in by using a certain set of specific formulas, which we call SCRIPT, and which represent different ways an imaginary scenario could evolve, given a certain action or event taking place, and conditioned by the fulfillment of certain antecedents.

The kind of behavior we want to capture, with this notion of moving forward an imaginary scenario by using these scripts, is similar to the way the Description algorithm behaved. In this case, we also want to capture the sense of "modal conditional" by evaluating whether the antecedent of a script holds, at the world of reference, and, if it does so, then clamp the antecedent of such script into a new accessible imaginary world. Nevertheless, and as we already pointed out, we need these scripts to take also into account what particular action or event occurs that make the consequence appear. Thus, we do not only need a "simple modal conditional", as we did in the previous case, but instead a "signed modal conditional", which would depend on a certain action α. In this sense, we take the same operator we used to represent this modal conditional, but we now add a superscript to sign it with a particular action: $\langle\rightarrow\rangle^{\alpha}$. Therefore, a formula such as $\varphi\langle\rightarrow\rangle^{\alpha}\psi$ stands for "if φ was the case and event α happened, then ψ could also be the case". We also define the corresponding box-operator of the signed modal conditional as $[\rightarrow]^{\alpha}$, and thus interpret a formula $\varphi[\rightarrow]^{\alpha}\psi$ as "if φ was the case and event α happened, then ψ would surely be the case as well". Considering this, we represent the set of scripts the agent believes in as follows:

$$\text{SCRIPT} = \{\varphi_1\langle\rightarrow\rangle^\alpha\psi_1, \varphi_2\langle\rightarrow\rangle^\alpha\psi_2, \varphi_3[\rightarrow]^\alpha\psi_3, \ldots\}$$

We provide the formal definition of the the two new operators $\langle\rightarrow\rangle^\alpha$ and $[\rightarrow]^\alpha$, and detail certain restrictions applied to both the antecedent φ_i and the consequent ψ_j of any formula in SCRIPT, Section 7.2.

The way of executing the Default Evolution process, though, is different from the Description process. Whereas in the Description our agent picked up a specific factual rule and used it to elaborate on the scenario, we do not want our agent, in this case, to imagine that a single script affects the scenario, but rather to imagine that a certain action α takes place, and then infer *every* consequence that α would carry with it: this represents a major change in the way the algorithms for the Description and the Default Evolution processes will work.

Why do we do this, though? Why not allow the agent to select which consequences of α she wants to focus on, or why not force the agent to unfold every possible static consequence of an imaginary scenario in the Description, according to the antecedents holding at the world of reference?

Our decision is mainly motivational. By considering and analyzing different particular cases of imagination acts, we note that the way we elaborate on the details of a static scenario is more selective than the way we unfold the consequences of a dynamic action. Take, for instance, an imaginary scenario about a tea-party in which we just stated that there is a table in front of us. Our beliefs about factual rules concerning tea-parties could tell us a lot of things that could also be the case in the scenario where there is a table; like, for instance, that "if there is a table and we are in a tea-party, then there are a kettle and a tea cup on the table as well", or that "if there is a kettle, then it is surely filled up with tea", or that "if I have a tea cup in front of me, it could either be empty or full", and so on. Nevertheless, we do not imagine, simultaneously, *all* the different static scenarios that could follow from the fact that there is a table in our tea-party, but rather we start by unfolding certain details and we "follow" them in a step-by-step way, while leaving other details that could possibly be inferred aside, as long as we do not focus on them (for instance, we may not be interested, or at least not yet, in whether there are cookies in the tea-party).

Conversely, we claim that, when elaborating on the consequences of certain events or actions, we are not as selective as with the static details. Take the same imaginary scenario about a tea-party, and suppose that we have specified, in addition, that the kettle is burning hot (and so is the tea inside it), that we are really thirsty, and suppose we imagine to pour tea into a cup. Now, say that we imagine to drink the whole tea cup in a single swallow. Would we typically imagine that, by drinking the whole cup of burning-hot tea, we *just* become satiated, without taking into account which other consequences would follow from drinking the burning-hot tea? Would not we imagine as well that we would burn our mouths, by drinking the burning-hot tea in a single gulp?

Taking this into account, we claim that, while the static elaboration of an imaginary scenario is usually more selective (regarding how we want to elaborate it, and choosing to follow a certain path or focusing on certain details), the dynamic elaboration of it is way more exhaustive (in terms of considering all the consequences that a certain action would carry with it). Due to this, we decide to define the algorithm for the Description process as being guided by a step-by-step process in which the agent chooses which details to focus on, and which factual rules to use on the elaboration of the scenario, whereas the algorithm for the Default Evolution process is defined by taking into account all the consequences that an action would carry with it in that imaginary scenario.

The algorithm for the Default Evolution, therefore, takes a world of reference w^R and a certain action or event α; then, by looking over all the scripts believed by the agent that concern action α, the algorithm determines what consequences would follow in the resulting imaginary worlds, as a result of α taking place at the world of reference. Figure 7.3 represents how this process work. Note how, conversely to the case of the Description process, in this case we need to take into account a whole set of formulas (the scripts concerning α), instead of focusing on a single one; as we will see when we formally define this algorithm, this fact increases the algorithm's complexity.

7.1.4 Unscripted Addition: Intervening in the Imagining

The process used for voluntarily adding new premises into the imaginary scenario, the Unscripted Addition process, is closely related to the Initialization process. Whereas in the Initialization the agent creates a new imaginary scenario by taking reality as the reference, in this case the agent decides to clamp a new premise into an already existing imaginary scenario.

In this case, we go back to the agent choosing to clamp an arbitrary formula into the imagining, and so we do not need, in this case and conversely to what we did in

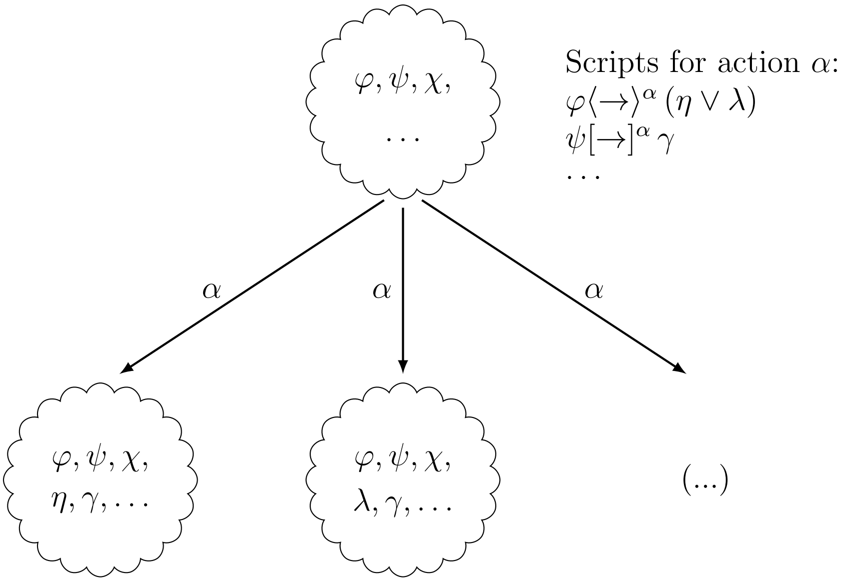


Figure 7.3: The scripts determine all the possible consequences of an action.

the previous two algorithms, any limited set of formulas that could be used; instead, we now allow our agent to imagine any formula δ, as we did in the Initialization process. The only difference that this algorithm will have, with respect to the one defined by the Initialization process, is that the current one will require the world of reference to be an imaginary world, instead of a real world: even if voluntary, the Unscripted Addition is a way of elaborating on an imaginary world, and the agent cannot elaborate on an imaginary world if it has not even been created.

Keeping this in mind, then, the algorithm for the Unscripted Addition takes an imaginary world w^R as the world of reference, creates the new imaginary worlds that could satisfy the added premise δ, and imports the rest of the atomic formulas from the world of reference to the newly created worlds. Figure 7.4 depicts how this process works. Note how the only different with respect to Figure 7.1, corresponding to the algorithm for the Initialization process, is that the current figure takes an imaginary world as the world of reference, instead of a real world.

7.2 Syntax

Conversely to what we did when defining the Logic of Imaginary Scenarios in Chapter 4, we define our Logic of Imagination Acts without taking the single-agent epistemic logic

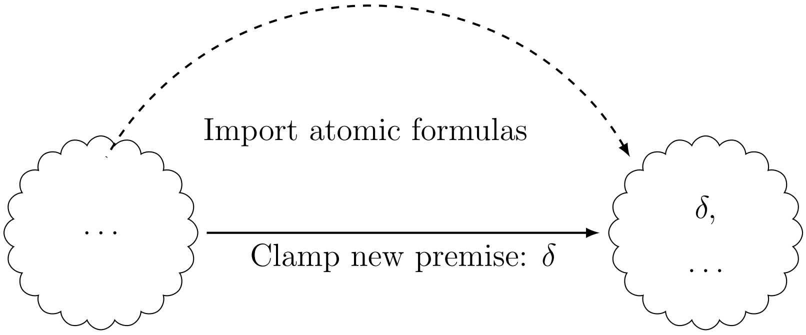


Figure 7.4: Adding new premises into the imagining.

as our base logic. In this case, the beliefs or knowledge of our agent will be represented implicitly by the way we define the whole system. We take this stance in order to be able to focus on what we want to capture: the dynamics of imagination acts. Mounting our proposal upon a single-agent epistemic logic, as we did before, involves a series of constrains inherited from it that diverts us from our actual goal; in this occasion, we want to focus only on the dynamics of imagination.

As it happened in Chapter 4, we introduce the definitions of the Logic of Imagination Acts in such a way that we need to refer to certain elements of the system before explicitly introducing them in a formal way. The reason, again, lies in the fact that both the syntax of the logic is closely related to the different algorithms our system uses, and the other way around.

Now, the language of the Logic of Imagination Acts is formed by a countably infinite set of *atomic formulas*, called ATOM, and represented by the lowercase letters p, q, and so on. There is also a countably infinite set of *nominals* (taken from hybrid logic), called NOM and represented by the lowercase letters i, j, and such. Besides, we have a countably infinite set of *atomic actions* (or simply *actions*), called ACT, and represented by the Greek letters α, β, and so on; note that these actions will only be used, in our language, to sign a special modal operator that we introduce in the further lines[2].

[2]Sometimes we also talk about *events*, instead of *actions*. Even though the terminological difference does not affect our formal setting, there is an important nuance in the meaning, when we use it in an informal way. In particular, we associate, to any action, an agent responsible for doing it; conversely, we do not need anyone to "do" an event. Whereas events *happen*, without the need for an intervention of an agent (at least an agent represented in our system), actions are *done* by some agent. For example, we would say that "it rains" is an event, whereas "open the umbrella" is an action. Our set of actions ACT may include any of the two: the only thing that matters, in our formal system, is that anything belonging to ACT represents something that occurs over a certain period of time (even if just an instant),

We use the standard propositional operators $\neg, \wedge, \vee, \rightarrow$ (standing for "negation", "conjunction", "disjunction" and "material implication", respectively); besides, we include the hybrid operator @. We use bracket symbols (, [,),] as usual (and we omit them when the context is clear).

We introduce four new dynamic operators, and four new static ones. Intuitively, each dynamic operator is responsible for calling one of the four algorithms (which will be defined in brief), and each dynamic operator has a corresponding static operator, used to evaluate the transitions created by the related execution of the algorithm. The new dynamic operators are $\mathrm{Init}(\delta)$, $\mathrm{Descr}(\zeta)$, $\mathrm{Evo}(\alpha)$ and $\mathrm{Add}(\delta)$, which are related to the static operators $\langle\delta\rangle^I$, $\langle\zeta\rangle^D$, $\langle\alpha\rangle^E$ and $\langle\delta\rangle^A$, respectively. Aside form this, we also introduce a new static modal operator $\langle *\rangle$.

The well-formed formulas of the language are defined by induction as follows:

$$
\begin{gathered}
i \mid p \mid \neg\varphi \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid \varphi \rightarrow \psi \mid @_i\varphi \mid \langle *\rangle\varphi \mid \\
\mathrm{Init}(\delta) \mid \mathrm{Descr}(\zeta) \mid \mathrm{Evo}(\alpha) \mid \mathrm{Add}(\delta) \mid \\
\langle\delta\rangle^I\varphi \mid \langle\zeta\rangle^D\varphi \mid \langle\alpha\rangle^E\varphi \mid \langle\delta\rangle^A\varphi
\end{gathered}
$$

where $i \in \mathrm{NOM}$, $p \in \mathrm{ATOM}$, $\{\varphi, \psi\} \subseteq \mathrm{FORM}$, $\delta \in \mathrm{FORM}^*$, $\zeta \in \mathrm{FACT}$ and $\alpha \in \mathrm{ACT}$; we explain in the following lines what these sets of formulas are[3]. Notice that, although in Section 7.1.3 we introduce also a set of distinguished formulas SCRIPT, we do mention it here; as we will see when detailing the formal details of the algorithm for the Evolution process, in Section 7.4.3, those formulas are only used internally by the algorithm, but play no role in the language.

We introduce two symbols $\top, \bot$ to refer to *truth* and *falsity*, respectively, and we define them as follows (for $p \in \mathrm{ATOM}$):

$$
\begin{aligned}
\top &\equiv p \vee \neg p \\
\bot &\equiv p \wedge \neg p
\end{aligned}
$$

Furthermore, and in order to distinguish certain particular formulas belonging to the

dynamically, and that potentially changes the current state of affairs by affecting it somehow.

[3]The way we define the set of actions ACT is inspired by the way Propositional Dynamic Logic, or PDL, defines a set of *atomic programs* Π_0. In PDL (see [27]), these programs are also used to sign a modal operator, just as we do in our case; nevertheless, PDL also defines a set of operators over programs, which can be used to combine them in different ways, but we are not interested in doing such thing with our actions. It could indeed be interesting to study the possibility in a future expansion of the logic, as it would amount to represent how certain combination of actions affect the way our agent develops the details of the imagining.

sets FACT and SCRIPT, we introduce two new symbols that will be used to encode a combination of two different operators as follows (for some $\zeta \in$ FACT):

$$\varphi\langle\rightarrow\rangle\psi \equiv \varphi \rightarrow \langle\zeta\rangle^D\psi$$
$$\varphi\langle\rightarrow\rangle^\alpha\psi \equiv \varphi \rightarrow \langle\alpha\rangle^E\psi$$

We also define, as usual, a derived $\Box$-like operator for each modal operator $\langle *\rangle$, $\langle\delta\rangle^I$, $\langle\zeta\rangle^D$, $\langle\alpha\rangle^E$ and $\langle\delta\rangle^A$ as follows (where # stands for any of the suitable formulas and $*$ for any of the superscripts that could be used in any of these operators):

$$[\#]^*\varphi \equiv \neg\langle\#\rangle^*\neg\varphi$$

The intuitive reading of the dynamic operators is the following: formula Init(δ) is read as "the agent creates a new imaginary scenario using the initial premise δ"; formula Descr(ζ) is read as "the agent elaborates on the static details of an imaginary scenario by using the factual rule ζ"; formula Evo(α) is read as "the agent elaborates on the consequences that action α would have in the imaginary scenario"; lastly, formula Add(δ) is read as "the agent adds a new premise δ into the imaginary scenario".

The associated static operators can be intuitively interpreted as follows: formula $\langle\delta\rangle^I\varphi$ is interpreted as "after an execution of the Initialization process, with initial premise δ, the agent imagines a world in which φ is the case". The rest of the corresponding static formulas are read similarly.

Operator $\langle *\rangle$ is intuitively interpreted as being a kind of "wildcard" imagination operator, and thus a formula $\langle *\rangle\varphi$ can be read as "there is some process of imagination that allows the agent to reach a world where φ holds".

In this proposal, we distinguish certain particular sets of formulas that are aimed to be used only by certain operators. In particular, we have:

- FORM corresponds to the set of all well-formed formulas of the language; we typically refer to elements of FORM by using φ, ψ and so on.

- FORM* corresponds, as it happened in the logic defined in Chapter 4, to the propositional fragment of FORM; we typically refer to elements of FORM* by using δ, γ and such. Therefore, a formula $\delta \in$ FORM* can be of the following form:

  - p (for $p \in$ ATOM)
  - $\neg\varphi$ (for $\varphi \in$ FORM*)

  - $\varphi \vee \psi$ (for $\{\varphi, \psi\} \subseteq \text{FORM}^*$)
  - $\varphi \wedge \psi$ (for $\{\varphi, \psi\} \subseteq \text{FORM}^*$)
  - $\varphi \rightarrow \psi$ (for $\{\varphi, \psi\} \subseteq \text{FORM}^*$)

  No other well-formed formula of the language can be a formula of FORM*.

- FACT is a particular subset of FORM, and it corresponds to the set of formulas aimed to represent the *factual rules* in which the agent believes, and which represent how certain conditions could give rise to certain consequences; we typically refer to elements of FACT by using ζ, ζ_1, ζ_2 and so on. We require every formula $\zeta \in \text{FACT}$ to be of the following form[4]:

  - $\varphi\langle\rightarrow\rangle\psi$ (for $\{\varphi, \psi\} \subseteq \text{FORM}^*$)

  Intuitively, and as we will see in brief, this kind of formulas account for factual rules describing that, if φ is the case in an imaginary world, then it could also be the case in the imaginary world that ψ. The reason why we only allow $\langle\rightarrow\rangle$-formulas, and not also $[\rightarrow]$-formulas, is detailed in Section 7.4.2, where we introduce the algorithm responsible for handling these kind of formulas.

- SCRIPT is another particular subset of FORM, and it corresponds to the set of formulas aimed to represent the *scripts* in which the agent believes, and which represent what consequences, given certain conditions, an action or event will trigger in an imaginary world; we typically refer to elements of SCRIPT by using ξ_1, ξ_2 and such. We require every formula $\xi \in \text{SCRIPT}$ to be of one of the following forms (recall that $\langle\rightarrow\rangle^\alpha \equiv\rightarrow \langle\alpha\rangle^E$):

  - $\varphi\langle\rightarrow\rangle^\alpha\psi$ (for $\{\varphi, \psi\} \subseteq \text{FORM}^*$ and $\alpha \in \text{ACT}$)
  - $\varphi[\rightarrow]^\alpha\psi$ (for $\{\varphi, \psi\} \subseteq \text{FORM}^*$ and $\alpha \in \text{ACT}$)

  Intuitively, this kind of formulas account for scripts describing that, if φ is the case in an imaginary world, after action α happens it will be the case that ψ.

[4]The reason why we only allow the antecedent and the consequent to belong to the propositional fragment of the language follows the same motivations we discussed in Section 4.4, while introducing the former `ImgAlg` of the Logic of Imaginary Scenarios. In a nutshell, we are interested in seeing how imaginary worlds are created and developed: modal and hybrid operators, nevertheless, convey information about either *other* worlds, or about the relation between other worlds. Leaving aside the technical complications that this would involve, imagining about other worlds or their relations falls outside the scope of our current goal, while defining this logic.

7.3 The Models for Imagination Acts

We define a *Model for Imagination Acts* as a structure $\mathcal{M} = \langle W, R_I, R_D, R_E, R_A, V, N\rangle$, where:

- W is a non-empty set of elements called *possible-worlds* or *states of affairs*. We use the lowercase letters $w, v, u, \ldots$ to refer to the elements of W.

- $R_I \subseteq W \times W \times \text{FORM}^*$ is a ternary relation called the *initialization relation*. Intuitively, an element (w, v, δ) captures how, through the Initialization process, and by using an initial premise δ and taking w as the *world of reference*, a new imaginary world v is created. We use triplets of the form $(w, v, \delta), (u, z, \gamma), \ldots$ to refer to elements of R_I.

- $R_D \subseteq W \times W \times \text{FACT}$ is a ternary relation called the *description relation*. Intuitively, an element (w, v, ζ) captures how, through the Description process, and by using a factual rule $\zeta \in \text{FACT}$ and taking w as the *world of reference*, an imaginary world v resulting from the application of ζ is created. We use triplets of the form $(w, v, \zeta_1), (u, z, \zeta_2), \ldots$ to refer to elements of R_D.

- $R_E \subseteq W \times W \times \text{ACT}$ is a ternary relation called the *evolution relation*. Intuitively, an element (w, v, α) captures how, through the Default Evolution process, by performing (or imagining to perform) an action $\alpha \in \text{ACT}$, and by taking w as the *world of reference*, an imaginary world v is created as a result of action α taking place. We use triplets of the form $(w, v, \alpha), (u, z, \beta), \ldots$ to refer to elements of R_E.

- $R_A \subseteq W \times W \times \text{FORM}^*$ is a ternary relation called the *addition relation*. Intuitively, an element (w, v, δ) captures how, through the Unscripted Addition process, using a premise δ, and by taking w as the *world of reference*, an imaginary world v is created. We use triplets of the form $(w, v, \delta), (u, z, \gamma), \ldots$ to refer to elements of R_A.

- $V : \text{ATOM} \to \mathcal{P}(W)$ is a function from atomic formulas of the language to subsets of the power set of W, called the *valuation function*. Intuitively, it keeps track of which atomic formulas are true at which subset of possible worlds.

- $N : \text{NOM} \to W$ is an exhaustive function setting, for each element of NOM, a possible world in W, and called the *nominal function*. Intuitively, it specifies which nominal is used to identify each world.

Notice that, with respect to the Models for Imaginary Scenarios previously defined in Section 4.3, we no longer have a single imagination relation R_{Img}, as we did in there, but we rather have four distinct relations now; this is consistent with the fact that, this time, we have four algorithms instead of one.

It is worth noting that both R_I and R_A are closely related to the former relation R_{Img}. In particular, in our former logical system, we were able to account for two different processes within acts of imagination: the creation of a new imaginary scenario, taking a real possible world as the world of reference, and the addition of new premises into an already existing imaginary world. In the former system, both processes were captured by the same relation R_{Img}; in this setting, nevertheless, our account of the dynamics of imagination is much more refined, and this is also reflected in the accessibility relations. We now have two relations R_I and R_A, which share the same structure (w, v, δ) as relation R_{Img}, but which have more restrictions than the former more, and so are more precise than it. The expressive power of our latter approach with regards to the dynamics of imagination acts, thus, can already be seen when comparing the formal models of both approaches and noting that R_I and R_A are, in fact, a refinement of what R_{Img} was.

7.3.1 Accounting for the Agent's Beliefs

Following the theories of imagination that steamed our work, and more particularly following the Common Frame for Imagination Acts, the agent's beliefs still play a very important role in this: after all, both the Description and the Default Evolution processes draw from the agent's beliefs. Nevertheless, and as we have already seen while defining both the syntax of our current proposal in Section 7.2 and the formal models in the present section, we have decided to build the Logic of Imagination Acts without using an epistemic or doxastic logic as its basis; therefore, we neither have any kind of operator, nor any kind of accessibility relation accounting for beliefs.

The reasons for doing so is because our current goal and main concern is to focus on the dynamics described by the four algorithms that will capture the processes of the Common Frame: adding more technical complexity to the setting could easily deviate us from our goal, and so we have decided to omit an explicit formal representation of knowledge and belief for now. Nevertheless, we will keep talking about and referring to the agent's beliefs: they are an unavoidable part of our theoretical setting, and thus we still consider them as being implicit in our formal proposal.

The way we overcome this apparent shortcoming is by taking the agent's beliefs to be implicitly defined in our models. In particular, we say that every factual rule belonging to FACT is believed by our agent; similarly, we say that our agents believes in every script belonging to SCRIPT. This way of understanding the factual rules and the scripts represented in our models as believed by the agent accounts for the fact that both the Description and the Default Evolution processes need to draw, actually, from what the agent believes in.

Similarly, we also need to represent what our agent believes that the real world to be like. We do this implicitly by understanding that each and every real world initially present in a model represents a possible way the agent believes the real world could be like: if there is more than one real possible world, it would mean that the agent is unsure about which one is the actual case.

We do not, however, include any kind of doxastic relation between such worlds: as they are all intended to represent possible worlds the agent considers epistemically plausible, the doxastic relation between them is also implicit –in fact, as our system represents a single agent, any possible world represented in the model can be interpreted as being "within" the mind of such agent. Similarly, we do neither include any kind of preference or plausibility relation between such worlds: the agent considers them to be equally plausible, and then she just chooses one on which she wants to imagine, and takes it as the world of reference for that particular imagination act.

Even though we know that this way of considering beliefs is a bit simplified (specially when there are so many different approaches that explicitly represent beliefs), we claim that it is enough to allow us to develop our proposal on the dynamics of imagination, while keeping the motivational understanding of the relation between imagination and beliefs.

7.3.2 Setting New Terminology

Throughout the rest of this chapter, we will be using specific terminology to refer to different parts of a model. Some of this terminology follows what we already defined in Chapter 4, while some terms are new.

- We use the term *real possible world* or *real world* to refer to a possible world $w \in W$ that is not the destination of any element belonging to the relations R_I, R_D, R_E or R_A; moreover, if w appears in one of these relations, it can only be as the origin of an element $(w, v, \delta) \in R_I$, but never in any of the other three accessibility relations. Intuitively, a real world stands for a state of affairs the agent believes that could represents the actual state of affairs, and which is not part of any imaginary scenario (unless it is taken as the world of reference used to initiate one).

- We use the term *reality* to refer to the set of all real possible worlds. Similarly, we may refer to such set of worlds using expressions like "how things actually are", or "the actual state of affairs", for instance.

- We use the term *imaginary possible world* or *imaginary world* to refer to a possible world $w \in W$ that is the destination of at least one element in R_I, R_D, R_E or R_A. Note that w could also be the origin of a different element in such relations, but w can *never* be the origin of relation R_I, as only real possible worlds can be used as the world of reference when initiating a new imaginary scenario. Intuitively, an imaginary world belongs to an act of imagination, and describes a possible state of affairs that has been obtained by applying one of the four imagination algorithms.

- We use the terms *imaginary scenario* or *imagining* to refer to the whole set of possible worlds $\{v_1, \ldots, v_n\} \in W$ that result from a *single* execution of one of the four algorithms of imagination, and thus that create a set of elements $\{(w, v_1, \#), \ldots, (w, v_n, \#)\}$ that belong to either R_I, R_D, R_E or R_A (for the same world of reference w and corresponding formula $\#$). Therefore, when we talk about an imaginary scenario or an imagining, we are referring to any of the imaginary worlds that have been defined during a particular execution of an algorithm.

- We use the term *imagination act* to refer to a whole "tree" of imaginary scenarios —which are, at the same time, formed by a set of imaginary possible worlds. Each one of those imaginary possible worlds must be accessible by following a chain of accessibility relations in R_I, R_D, R_E and R_A, with the particularity that the elements of R_I that are part of this chain must have the same real world of reference w and the same initial premise δ. Intuitively, an imagination act embeds the whole set of imaginary possible worlds that have been created by a *single* execution of the `InitAlg`, and then by any number of executions of the `DescrAlg`, the `EvoAlg` or the `AddAlg` by using either the imaginary worlds created by the `InitAlg`, or any other imaginary world created as a result of the other algorithms[5].

[5] Note that, in Chapter 4, we used this term to refer to the outcomes of a *single* execution of the Imagination Algorithm; conversely, we use this term here to refer to the outcomes of the different executions of each of our four algorithms (although limited by a single execution of the `InitAlg`). The difference in terminology comes from the fact that, whereas in Chapter 4 we defined a single algorithm for imagination acts, this time we are defining four algorithms that represent different processes embedded

- We use the term *root* to refer to a single pair, formed by a world of reference w^R and an initial premise δ, which appear in an element belonging to R_I. Intuitively, the root of an imagining is a specific execution of the Initialization process that creates the first imaginary worlds, upon which the whole imagining is based.

- We use the term *horizontal* to relate those possible worlds that result from the same execution of an algorithm, and thus that have the same world of reference and the same premise, factual rule or script (depending on the particular case); in other words, two possible words are said to be horizontal if they belong to the same imaginary scenario. Intuitively, two horizontal imaginary worlds represent alternative outcomes of the same execution of a specific algorithm.

- Conversely, we use the term *vertical* to relate those possible worlds that, although belonging to the same imagination act, do not both result from the same execution of an algorithm (that is: they do not belong to the same imaginary scenario). Intuitively, two worlds are said to be vertical between them if, by following one or more accessibility relations, one of them can be reached from the other.

- We use the term *imaginary story* or an *imagination chain* to refer to a particular path within an imagination act (that is: a vertical sequence of accessibility relations between possible worlds). A path is a whole succession of accessibility relations $(w, v, \delta), (v, u, \#), \ldots, (z, x, \#)$ in which w is a real possible world, element (w, v, δ) belongs to R_I, each possible world of origin is also the possible world of destination from the previous element of the chain (except the first real possible world appearing in it), and the last possible world x of the chain is not the world of origin of any other element in any accessibility relation. Intuitively, an imaginary story corresponds to a specific path within an imagination acts that represents one possible way of elaborating an imagining.

- Within an imaginary story, we use the term *leaf-world* to refer to those possible worlds v which, being the destination of some accessibility relation $(w, v, \#)$, are not the origin $(v, u, \#)$ in any relation. Intuitively, a leaf-world is the end of a specific imaginary story: either the agent stopped developing the imagining any further, or

into an imagination act; therefore, our understanding of the term "imagination act" is now much more modular and detailed than it was before, and this is also reflected by how we use this term, regarding our formal system and the algorithms used by it.

she chose to develop her imagining by following another horizontal world, and thus this particular leaf-world marks the end of a path within the imaginary story.

7.4 The Algorithms

Following the Common Frame for Imagination Acts introduced in Chapter 6, we want our Logic of Imagination Acts to represent and capture the dynamics of imagination at a more fine-grained and detailed level than we did in the Logic for Imaginary Scenarios. In particular, we want to account for each and every execution of any of the four processes identified in the Common Frame for Imagination Acts, and we want our logic to recognize and represent them. Therefore, we can no longer think of an act of imagination as a *single* algorithm, but rather as set of different executions of different algorithms, each one accounting for one particular process within a whole imagination act. The syntax of our logic, defined in Section 7.2, already points to this goal by distinguishing four different dynamic operators, associated each one to a different static modal operator, and which are related to four different accessibility relations, and defined in Section 7.3. Therefore, and following this, we now define not *the* algorithm for imagination acts, but rather the *four* distinct algorithms accounting for each one of the four processes involved in an imagination act.

During the execution of any of these four algorithms, a Model for Imagination Acts $\mathcal{M}$ is expanded into its *expanded model* $\mathcal{M}^+$. We refer to any of the elements of $\mathcal{M}^+$ as the *expanded element* (with its corresponding name), and we identify them as $\mathcal{M}^+ = \langle W^+, R_I^+, R_D^+, R_E^+, R_A^+, V^+, N^+\rangle$. Note, however, that not each algorithm will expand every element of $\mathcal{M}$; nevertheless, we will still talk about the expanded version of such elements, when referring to them either during, or just after the execution of one of such algorithms[6].

Even though each algorithm has its own particularities, there are certain processes within them that are very similar, both in the way they work and in their outcome. We provide a complete specification for each algorithm, but we will assign names or labels to some of those internal processes. The reason for doing so is to help the reader to identify, beforehand, what these processes do, and identify that they stand for similar processes appearing in the other algorithms. Note that, even if they intend to do the same thing, two

[6]Each one of the four algorithms we define in the following pages is executed upon a model $\mathcal{M}$. On the first step of their execution, there are certain initial conditions that each algorithm has to check. As it happened in the `ImgAlg` defined for the Logic of Imaginary Scenarios in Chapter 4, in case some of those conditions are not fulfilled, the algorithm does not expand model $\mathcal{M}$ in any way; in that case, we consider that $\mathcal{M}^+ = \mathcal{M}$.

similar processes in two different algorithms might still be technically different: they may refer to different parts of the model, or they may alter them in different ways. However, the goal they seek to accomplish is the same as in the other algorithms.

It is important to keep in mind that, as it happened with the processes defined in the Common Frame for Imagination Acts, the execution of the algorithms is not required to follow any specific order. If there is yet no imaginary world created, we do need to start with an execution of the algorithm responsible for handling the Initialization of the scenario, but, after this, the agent may choose to elaborate on the imaginary worlds by using any of the other three algorithms, without the need to follow any specific order. This fact is one of the core points of our Common Frame, and so it is of our Logic of Imagination Acts.

Let's proceed to the definition of each one of the four algorithms, corresponding to each process identified by the Common Frame for Imagination Acts.

7.4.1 The `Init` Algorithm

The `InitAlg` corresponds to the Initialization process of the Common Frame for Imagination Acts, and it is responsible of creating a brand-new imaginary world by taking a certain real possible world as the world of reference w^R, and by imagining a certain formula $\delta \in \text{FORM}^*$. The `InitAlg` is called by evaluating the dynamic formula $\text{Init}(\delta)$ at a possible world w^R as follows:

$$\texttt{InitAlg}(\delta, w^R)$$

The algorithm follows these steps:

1. *Check initial conditions*: If world w^R is *not* a real possible world (that is, if there exists some v such that $(v, w^R, \#)$ is in either R_I, R_D, R_E, or R_A, being $\#$ one of the corresponding formulas required by each accessibility relation), do nothing[7]. Similarly, if δ is contradictory (that is, if $\delta \equiv \bot$), do nothing[8].

2. *Compute DNF*: In order to handle the formula in an efficient way, we compute the *Disjunctive Normal Form* (DNF) of δ, to which we refer as $\text{DNF}(\delta)$. In the following

[7]The *Initialization* step is used to *create* a new imaginary world from scratch, and this must be done from a real possible world. Initiating an act of imagination while taking an imaginary world to be the "real reference world" would be a misuse of imagination; instead, one could clamp some new conditions to an already existing imaginary scenario via the *Unscripted Addition* process.

[8]Although other authors (such as Priest or Berto) do suggest that we can imagine true contradictions, we do not think that we actually can do such things, and so we do not allow this in our proposal.

steps of the `ImgAlg`, we refer to the clauses that form the DNF(δ) as follows: DNF(δ) $= \delta_1 \vee \ldots \vee \delta_n$.

3. *Create imaginary worlds*: Create new imaginary possible worlds $w_1, \ldots, w_n$ for each clause $\delta_1, \ldots, \delta_n$ in DNF(δ). This defines the *expanded set of possible worlds* as follows:
$$W^+ = W \cup \{w_1, \ldots, w_n\}$$

4. *Expand accessibility relation*: Once the new possible worlds have been created, the `InitAlg` must create the new initialization relations expressing that, by imagining formula δ at the world of reference w^R, the agent has crated new imaginary worlds $w_1, \ldots, w_n$. This defines the *expanded set of initialization relations* as follows:
$$R_I^+ = R_I \cup \Big(\bigcup_{i=1\ldots n} \{(w^R, w_i, \delta)\} \Big)$$

5. *Expand nominal structure*: Now, the `InitAlg` must add a set of new nominals to refer to the newly created imaginary worlds. This defines both the *expanded set of nominals*, by adding one new nominal k_i for each new possible world w_i created during the current execution of the `InitAlg`, and the *expanded nominal function*, which is a functional extension of N relating the new pairs of nominals and possible worlds:
$$\mathrm{NOM}^+ = \mathrm{NOM} \cup \{k_1, \ldots, k_n\}$$
$$N^+ = N \cup \Big(\bigcup_{i=1\ldots n} \{(k_i, w_i)\} \Big)$$

6. *Expand valuation function*: Last but not least, the `InitAlg` must expand the valuation function to account for the new imaginary possible worlds. In order to do so, the algorithm must account for both the atoms that are present in each δ_i, and also for the atoms that are true in the world of reference w^R and which should be imported to the new imaginary worlds, provided they do not appear in δ_i; this is so because any atom appearing in δ_i would have already been set in the new imaginary world. Therefore, the definition of the *expanded valuation function* involves two different phases:

   (a) *Clamp new atoms*: Firstly, the `InitAlg` must set the new valuation functions according to the atoms p appearing in δ_i, for each new imaginary world w_i:

$$V_1^+(p) = V(p) \cup \Big(\bigcup_i \{w_i \mid p \text{ is a positive literal appearing in } \delta_i\}\Big)$$

(b) *Import existing atoms*: Then, it must import all the atoms that are true at the world of reference w^R, provided they do not appear in δ_i, for each new imaginary possible world w_i:

$$V^+(p) = V_1^+(p) \cup \Big(\bigcup_i \{w_i \mid w^R \in V_1^+(p) \text{ and } p \text{ is not a literal of } \delta_i\}\Big)$$

7. The `InitAlg` has finished its execution: a new set of imaginary possible worlds satisfying δ has been created, and these worlds are now accessible through the initialization relation R_I from the world of reference w^R.

Note that, due to the fact that we no longer have an epistemic indistinguishability relation (as relation R_K in the Logic of Imaginary Scenarios defined in Chapter 4), the intuitions behind the newly created imaginary worlds are slightly different. Whereas in the Logic of Imaginary Scenarios we created different imaginary worlds accounting for δ, and about which the agent was "unable" to tell the difference (because they were all epistemically indistinguishable, as far as she knew), we now interpret the creation of multiple imaginary worlds differently. If δ can be accounted for more than one possible world, then the agent can obviously reach any of them by imagining δ, but she no longer fails to distinguish them; in fact, they are not even related.

Imagining δ, therefore, does *not* lead to an epistemically vague state in which the agent is unsure about which specific world she meant by imagining δ, but it rather draws different possible courses of action, each one leading to a different world satisfying δ; the agent is aware of that and, if she wants to keep elaborating on one of the imaginary worlds, she must *choose* which particular world to follow.

The difference between this approach and the former one may seem unimportant; in fact, the agent did already choose one of the possible imaginary worlds to execute a new act of imagination in the Logic of Imaginary Scenarios, even if she was not able to tell the difference between them. Nevertheless, we believe that the difference between both approaches is worth mentioning. Using the epistemic indistinguishability relation on the imaginary worlds implicitly meant, in some sense, that the agent created a set of vague, indistinguishable possibilities when imagining δ, but with an uncomfortable flavor of uncontrolled imagination, as if she was unable to tell the difference between them, and so about to get lost in her own imaginings.

In the present setting, the fact that the different new imaginary worlds are not indistinguishable between them, represents that the agent recognizes that they are all valid possibilities when it comes to imagine δ, but at the same time it captures the fact that, if she wants to keep elaborating them, she must *choose* (with all the agentiveness it implies) between different distinguishable courses of action. The same phenomenon happens with the rest of the algorithms: if the agent imagines something (or elaborates an imaginary world in a certain way) that could be accounted for more than one possible imaginary world, then they are all represented as distinguishable courses of action, but without being related between them.

7.4.2 The `Descr` Algorithm

The `DescrAlg` corresponds to the Description process of the Common Frame for Imagination Acts, and it is responsible for elaborating on the static details of an imaginary scenario by following certain factual rules, which accounting for what else could be the case in the imagining, given certain conditions. The `DescrAlg` needs an imaginary possible world as the world of reference w^R and a certain factual rule $\zeta \in \text{FACT}$, and it is called by evaluating the dynamic formula $\text{Descr}(\zeta)$ as follows:

$$\texttt{DescrAlg}(\zeta, w^R)$$

The algorithm follows these steps:

1. *Check initial conditions*: If world w^R is a real world (that is, there is no v such that $(v, w^R, \#)$ is in neither R_I, R_D, R_E, nor R_A, being $\#$ one of the corresponding formulas required by each accessibility relation), do nothing[9].

2. *Compute DNF*: Formula $\zeta \in \text{FACT}$ is required to be of the form $\zeta = \varphi\langle\rightarrow\rangle\psi$ (where $\{\varphi, \psi\} \subseteq \text{FORM}^*$). Now, the `DescrAlg` must check whether the antecedent of such formula is true at the world of reference and, if it is, then it must elaborate on the description of the imaginary world by clamping its consequent to a new imaginary world (or several new ones, depending on the form of the consequent); aside from this new addition, though, the new imaginary world will be as the world of reference

[9] In this case, the agent can only elaborate on the description of an imaginary world that has *already* been created: the agent cannot imagine "how things would be like" in the actual world!

is. If $\mathcal{M}, w^R \vDash \varphi$ (that is: the antecedent of ζ holds at the world of reference w^R), then the `DescrAlg` must compute $\text{DNF}(\psi) = \psi_1 \vee \ldots \vee \psi_n$ (that is, the DNF of the consequent of ζ).

3. *Create imaginary worlds*: The algorithm must create n new imaginary possible worlds $w_1, \ldots, w_n$. This defines the *expanded set of possible worlds* as follows:

$$W^+ = W \cup \{w_1, \ldots, w_n\}$$

4. *Expand accessibility relation*: Create new R_D relations linking the world of reference w^R with each new world $w_1, \ldots, w_n$ created in the previous step through the factual rule ζ. This defines the *expanded set of description relations* as follows:

$$R_D^+ = R_D \cup \Big(\bigcup_{i=1 \ldots n} \{(w^R, w_i, \zeta)\} \Big)$$

5. *Expand nominal structure*: The `DescrAlg` must now add a set of new nominals to refer to the newly created imaginary worlds. This defines both the *expanded set of nominals*, by adding one new nominal k_i for each new possible world w_i created during the current execution of the present algorithm, and the *expanded nominal function*, which is a functional extension of N relating the new pairs of nominals and possible worlds:

$$\text{NOM}^+ = \text{NOM} \cup \{k_1, \ldots, k_n\}$$
$$N^+ = N \cup \Big(\bigcup_{i=1 \ldots n} \{(k_i, w_i)\} \Big)$$

6. *Expand valuation function*: The `DescrAlg` must expand the valuation function to account for the new imaginary possible worlds. In order to do so, the algorithm must account for both the atoms that are present in each ψ_i, and also for the atoms that are true in the world of reference w^R and which should be imported to the new imaginary worlds, provided they do not appear in ψ_i. Therefore, the definition of the *expanded valuation function* involves two different phases:

   (a) *Clamp new atoms*: Firstly, the `DescrAlg` must set the new valuation functions according to the atoms p appearing in ψ_i, for each new imaginary possible world w_i:

$$V_1^+(p) = V(p) \cup \left(\bigcup_i \{w_i \mid p \text{ is a positive literal appearing in } \psi_i\} \right)$$

(b) *Import existing atoms*: Then, it must import all the atoms that are true at the world of reference w^R, provided they do not appear in ψ_i, for each new imaginary possible world w_i:

$$V^+(p) = V_1^+(p) \cup \left(\bigcup_i \{w_i \mid w^R \in V_1^+(p) \text{ and } p \text{ is not a literal of } \psi_i\} \right)$$

7. The `DescrAlg` has finished its execution: a new set of imaginary possible worlds has been created as a result of the agent following a factual rule ζ describing her belief about how things would be like, if things were as represented by the world of reference; moreover, these new imaginary worlds are now accessible through the description relation R_D from the world of reference w^R.

About the Restriction on Factual Rules

As we have already specified in Section 7.2 while defining the set FACT of factual rules, we only allow this kind of formulas to contain the $\langle\rightarrow\rangle$, but not the $[\rightarrow]$ operator; why is that so?

Likewise to the Logic of Imaginary Scenarios introduced in Chapter 4, in the Logic of Imagination Acts we want to focus on the dynamics of imagination and the processes involved in creating and developing an imaginary world, step-by-step. Therefore, we want to keep trace of each "imaginative process" that the agent performs in a specific model: we do not want to replace or update any part of a model, as we do not want to lose information about any of the processes that have been executed in there.

The way we treat $\langle\rightarrow\rangle$-formulas in the `DescrAlg` conforms to these requirements: when the agent executes the `DescrAlg` in a certain world of reference w^R, the factual rule ζ_1 applied states that, if the antecedent is true at w^R, then a new world satisfying the consequent can be imagined: nevertheless, the world of reference w^R is *not* modified in any way. Even though the factual rules describe what *else* could be the case in a certain imaginary world, which keeps representing a static state of affairs, our system captures this internal reasoning step by creating a new imaginary possible world conforming to the world of reference w^R in every way, except for the new information that should be added in there, according to ζ_1. This way, our system can keep track of *every* process executed by the agent, and how each process has updated a certain imaginary world —not by explicitly updating the world itself, but by creating an updated copy of it.

If, after applying a certain factual rule to a world of reference w^R, which creates a new imaginary world called w_1, the agent wants to keep elaborating on what else could be the case in the resulting world, then she should do so by taking w_1 as the new world of reference. If, conversely, the agent wants to "backtrack" her static elaboration of w^R and consider how a different factual rule ζ_2 would change the imaginary world, then she could execute the `DescrAlg` again by taking w^R as the world of reference, and using the new factual rule ζ_2.

Note, however, that this new execution would *not* erase not replace the already existing world w_1, created using the factual rule ζ_1; instead, it would create a new imaginary world w_2, based on w^R, but developed following a different factual rule. This conforms with the fact that the new factual rule, as we require, is characterized by the $\langle\rightarrow\rangle$ modality, stating that world w^R could lead to w_2 by following ζ_2; this factual rule, thus, does not say anything about any other possible world created from w^R, such as w_1.

What would happen, though, if the agent wanted to backtrack its steps to w^R and then consider a certain factual rule ζ_3 based on a $[\rightarrow]$ operator, instead? Formula ζ_3 would be saying that a certain set of consequences must be applied to *every* imaginary world accessible from w^R. If we follow the same example, we already have a pair of worlds w_1 and w_2, which are accessible from w^R, and which have already been elaborated following certain factual rules. If we now wanted to consider a $[\rightarrow]$-formula from w^R we would need to "update" what is the case at both worlds w_1 and w_2. However, this presents two different issues.

First of all, the way worlds w_1 and w_2 have been elaborated has already been determined by factual rules ζ_1 and ζ_2: those worlds are the way they are *because* of such rules. If we now modify what is the case in them by following the factual rule ζ_3, they would no longer be representing the worlds that result from following rules ζ_1 and ζ_2, respectively, but rather representing a combination of two different factual rules, which could result in different consequences. We would be losing information regarding some of the processes performed by the agent within the whole act of imagination, and thus we would not be able to keep track anymore of every process performed by her. As we already said, we do *not* want such thing to happen in our system.

Moreover, there is still one further technical reason why we do not want to do so, and which follows the reason presented in the previous paragraphs. While creating a new imaginary world w_1 from w^R by using the factual rule ζ_1, the `DescrAlg` creates a new accessibility relation $(w^R, w_1, \zeta_1) \in R_D$, stating that world w_1 has been created *as a result of* formula ζ_1. However, if we now wanted to use a $[\rightarrow]$-formula (call it ζ_3) that would change what is the case in world w_1, then we would also have to add a new relation $(w^R, w_1, \zeta_3) \in R_D$. If such was the case, we would end up having two different accessibility relations from w^R to w_1, stating that world w_1 has been created as a consequence of rule ζ_1, while the other relation would be saying that w_1 results from ζ_3. The truth, however, would be that w_1 would not probably conform to what ζ_1 expresses, nor to what ζ_3 expresses. It would be a sort of mixture between both, but which would not entirely conform to any of the two. The order of executing such formulas would alter the results as well; if ζ_1 and ζ_3 conflict in any of their consequences, executing them in a different order would lead to a different resulting imaginary world.

As we can see, allowing for such kind of formulas would not only hide procedural information in our model, but it would also compromise the correctness of how our model represents the way our agent develops an imaginary scenario. Therefore, for both motivational and technical reasons, we do not allow such kind of formulas to appear in our set of factual rules. This could be seen as a limitation in our system, but it follows a decision on the way we treat the `DescrAlg`. Nevertheless, we could have chosen to define our algorithms in a different way, and this would lead to a different way of representing the processes we are interested in. For now, though, we stick with our decision.

7.4.3 The `Evo` Algorithm

Note that this algorithm has some particularities, with respect to the other ones. The argument needed for this algorithm is an action $\alpha \in$ ACT; however, and as we explain in brief, the algorithm uses the scripts $\xi \in$ SCRIPT in order to develop the imaginary scenario. This is unlike what happened with the `DescrAlg`, in which a factual rule $\zeta \in$ FACT was directly used to infer the contents of the new imaginary world; why is that?

In the previous case, the agent just needs to consider what it is already the case in the imaginary world, in order to imagine what *else* would be the case in there: the development is static and concerns a scenario which is not "moving", nor anything is happening in there. Nevertheless, in the present case things go different: in this case, the agent imagines what events could happen in the imaginary scenario, or which actions could be carried out in there. Therefore, the agent is *not* just imagining the consequences of something being already the case in the imagining, but she is imagining the consequences of a certain action α, while taking into account what is already the case in the scenario.

The development of the imaginary scenario, therefore, should not be done by using rules of the kind "if something was the case, then something else would also be the case", but rather by using scripts saying that "if something was the case and a certain action happened, then a series of consequences would also be the case". Due to this, what the agent imagines is that a certain action α takes place, and then develops what the consequences of such action would be, using her beliefs about how α would affect the imaginary world, regarding certain antecedents being true in there. The `EvoAlg`, therefore, corresponds to the *Default Evolution* process of the Common Frame for Imagination Acts, and it is responsible for computing which consequences would follow from a certain action or event α happening in an imaginary world, given certain conditions.

It is worth noting that, during the development of the algorithm, the elements of the expanded model will be defined recursively over a series of loops. Due to this, there is a convention that, for the sake of simplicity, we introduce in our notation. In particular, and due to the fact that the algorithm involves looping over possibly many formulas, the expansions of the corresponding elements $W^+, R_E^+, V^+, \dots$ are begin accumulated at each loop. Therefore, we assume that, whenever we refer to any element $W, R_E, V, \dots$ we are referring to the most "updated" version of that element, in the sense of already including whatever has been added to it in the previous loop[10].

For instance, and as we will see in brief, the first time we go over the loop and evaluate the first script of the queue, we define the expanded set of possible worlds as $W^+ = W \cup \{\dots\}$; in this case, the set W refers to the original, unexpanded set of possible worlds belonging to the model. Nevertheless, when we perform another loop by taking, say, the second script in the queue, the set W of the definition $W^+ = W \cup \{\dots\}$ is meant to correspond to the "current" or "most updated" state of the set of possible worlds; that is, including any possible world that has been added in the first loop, and similarly for any other loop and any other expansion of elements of the model.

The `EvoAlg` needs an imaginary world as the world of reference w^R, and an action $\alpha \in \mathrm{ACT}$, which the agent would imagine happening, and it is called by evaluating the dynamic formula $\mathrm{Evo}(\alpha)$ as follows:

$$\texttt{EvoAlg}(\alpha, w^R)$$

[10] We can draw a parallelism with the way variables are usually handled in programming languages. In there, it is typical to override the value of a variable by using its own value; for instance, one can increase the value of an integer index i by saying $i = i + 1$.

Now, the algorithm must follow these steps:

1. *Check initial conditions*: If world w^R is a real world, do nothing (this algorithm to be executed upon an imaginary world); if α does not appear in any formula within SCRIPT, do nothing (it would mean that the agent has no beliefs at all about the consequences of an action α).

2. *Create queue of scripts*: Each formula $\xi \in$ SCRIPT is either of the form $\varphi\langle\to\rangle^\alpha\psi$ or of the form $\varphi[\to]^\alpha\psi$, for some $\alpha \in$ ACT. Create a queue S^α of formulas sorted in the following way (in Section 7.4.3.1 we argue why we sort the scripts in such way):

   - Firstly, look for all the diamond-formulas in SCRIPT which are about α, and add them to the queue S^α while prioritizing the ones with the least complex antecedent (that is, the ones whose antecedent has less atomic formulas); for example, a formula $p\langle\to\rangle^\alpha$... has more priority than a formula $p \vee q\langle\to\rangle^\alpha$.... Formulas with the same antecedent complexity are sorted sequentially.

   - Secondly, look for all box-formulas in SCRIPT that are about α, and add them to the queue while prioritizing as well the ones with the least complex antecedent.

3. *Loop through the scripts*: This loops forms the central part of this algorithm. The loop starts by evaluating the first formula $\xi_1 \in S^\alpha$, and keeps looping until it has evaluated every script in the queue[11]. Recall that diamond-formulas with the least complex antecedent will be evaluated first, and box-formulas with the most complex antecedent will be evaluated last:

   (a) *Evaluate diamond-formula*: If the current script $\xi \in S^\alpha$ being evaluated is of the form $\varphi\langle\to\rangle^\alpha\psi$, and if its antecedent φ holds at the world of reference w^R (that is, if $\mathcal{M}, w^R \vDash \varphi$), do the following:

      i. *Compute DNF*: Compute the DNF of the consequent; that is, DNF(ψ) = $\psi_1 \vee \ldots \vee \psi_n$.

[11]The idea behind this loop is that, conversely to what happened with the Description process (in which the agent elaborated the scenario step by step by picking a single factual rule each time), in this case the agent imagines performing an action. Therefore, the agent must check for *all* the consequences of such action, which are described (according to their preconditions) by the formulas in SCRIPT; as we have already argued in this chapter, we claim that the reason for doing so is that one cannot imagine that she performs an action, and then that only *some* of its consequences happen.

ii. *Create imaginary worlds*: Create n new imaginary worlds $w_1, \ldots, w_n$, one for each ψ_i in $\mathrm{DNF}(\psi) = \psi_1 \vee \ldots \vee \psi_n$. This defines the expanded set of possible worlds as follows:

$$W^+ = W \cup \{w_1, \ldots, w_n\}$$

iii. *Expand accessibility relation*: Create new R_E relations from w^R to each new imaginary world w_i, signed with action α:

$$R_E^+ = R_E \cup \Big(\bigcup_{i=1\ldots n} \{(w^R, w_i, \alpha)\} \Big)$$

iv. *Check nominal structure*: Similarly to what happened in the other algorithms, the `EvoAlg` must now add a set of new nominals to refer to the newly created imaginary worlds. This defines both the *expanded set of nominals*, by adding one new nominal k_i for each new possible world w_i created during the current execution of the present algorithm, and the *expanded nominal function*, which is a functional extension of N relating the new pairs of nominals and possible worlds:

$$\mathrm{NOM}^+ = \mathrm{NOM} \cup \{k_1, \ldots, k_n\}$$

$$N^+ = N \cup \Big(\bigcup_{i=1\ldots n} \{(k_i, w_i)\} \Big)$$

v. *Expand valuation function - Clamp new atoms*: Set the valuation of each new imaginary world w_i according to the consequences of the corresponding clause ψ_i in $\mathrm{DNF}(\psi)$; note that, in this case, we do *not* yet import the atomic formulas of the world of reference, as we first need to keep evaluating the consequences of action α according to the other formulas in S^α: we do this because box-formulas may also affect the atomic valuation of the new imaginary worlds being created right now, and so this must be given priority over importing atomic formulas from the world of reference:

$$V^+(p) = V(p) \cup \Big(\bigcup_i \{w_i \mid p \text{ is a positive literal appearing in } \psi_i\} \Big)$$

(b) *Evaluate box-formulas*: If the current script $\xi \in S^\alpha$ being evaluated is of the form $\varphi[\rightarrow]^{\alpha}\psi$, and if its antecedent φ holds at the world of reference w^R (that is, if $\mathcal{M}, w^R \vDash \varphi$), then:

i. *Compute DNF*: Compute the DNF of the consequent; that is, $\text{DNF}(\psi) = \psi_1 \vee \ldots \vee \psi_m$.

ii. *Loop over already existing worlds*: When evaluating box-formulas within the script queue, there are some considerations that are worth taking into account. Whereas diamond-formulas express that certain outcomes *could* follow from certain antecedent, box-formulas state that certain outcomes *must* follow. Therefore, box-formulas should also take into account those new imaginary worlds that have already been created when evaluating diamond-formulas, and apply the corresponding consequences described by the current box-formula to them as well.
Thus, if there exists at least one world w_i such that $(w^R, w_i, \alpha) \in R_E$ (that is, if at least one new imaginary world has been already created during the current execution of this algorithm), then each possible consequence ψ_j in $\text{DNF}(\psi)$ must be handled while taking in account those already existing imaginary worlds. In order to handle this, the algorithm must loop over the already existing possible worlds w_i created during the current execution of the algorithm, and, for each w_i, do the following:

A. *Create new imaginary worlds*: For the already existing imaginary world w_i being considered, which we will call w_{i_m} throughout the current loop, the algorithm must create $m-1$ new imaginary possible worlds w_{i_j} (for $j = 1, \ldots, (m-1)$, and being m determined by $\text{DNF}(\psi) = \psi_1 \vee \ldots \vee \psi_m$); this defines the expanded set of possible worlds as follows[12]:

$$W^+ = W \cup \{w_{i_1}, \ldots, w_{i_{(m-1)}}\}$$

B. *Expand accessibility relation*: Create, for each new possible imaginary

[12]In order to clarify the intuitions behind this step, suppose that the algorithm is evaluating a box-formula that represents m alternative outcomes: by following the way our other algorithms have been working, we should create m new possible imaginary worlds to account for each one of those outcomes. However, if there already exists an imaginary possible world (possibly as a result of evaluating a diamond-formula), then one of such outcomes should be represented in the world that already exists —for, otherwise, if we created m new imaginary worlds, aside from the already existing one, we would have $m+1$ new possible worlds to account for just m alternative outcomes represented in the current box-formula. Therefore, when evaluating a box-formula upon an already existing imaginary world, we only have to create $m-1$ new worlds, as the already existing one is also taken into account while evaluating the outcomes of the current box-formula.

world, an R_E relation as follows:

$$R_E^+ = R_E \cup \Big(\bigcup_{j=1\ldots(m-1)} \{(w^R, w_{i_j}, \alpha)\} \Big)$$

C. *Expand nominal structure*: Create and associate new nominals to these new imaginary worlds. This defines both the expanded set of nominals, and the expanded nominal function as follows:

$$\mathrm{NOM}^+ = \mathrm{NOM} \cup \{k_{i_1}, \ldots, k_{i_{(m-1)}}\}$$

$$N^+ = N \cup \Big(\bigcup_{j=1\ldots(m-1)} \{(k_{i_j}, w_{i_j})\} \Big)$$

D. *Expand valuation function - Clamp new atoms*: Set the valuation of each new imaginary world w_{i_j} according to the consequences of the corresponding clause ψ_j in DNF(ψ). Note that, in the previous step, we have created $m-1$ new imaginary worlds; this results from the fact that, before evaluating the current box-formula, we already had at least one existing imaginary world created in a previous step of the current execution of this algorithm (probably while evaluating a diamond-formula). Now, as DNF(ψ) $= \psi_1 \vee \ldots \vee \psi_m$, and as we have used index $j = 1, \ldots, (m-1)$ for the new imaginary worlds created in this loop, we must associate the world w_{i_m}, which was the one that already existed when entering into the current loop, with the clause ψ_m of DNF(ψ), and the rest of the newly created imaginary worlds w_{i_j} with the rest of the clauses in DNF(ψ_j), for $j = 1, \ldots, (m-1)$. The clamping of new atoms, in this case, must proceed in three different steps or phases.
Firstly, as the newly created worlds w_{i_j} are meant to be copies of the original diamond-world w_{i_m}, the algorithm needs to ensure that those worlds w_{i_j} satisfy the same atomic propositions as w_{i_m}: in other words, it should add every world w_{i_j} to the valuation function of any atom satisfied by w_{i_m} as follows:

$$V^+_{copy}(p) = V(p) \cup \Big(\bigcup_{j=1\ldots(m-1)} \{w_{i_j} \mid w_{i_m} \in V(p)\} \Big)$$

Secondly, the algorithm must proceed and clamp the positive atoms in each clause φ_j to their corresponding imaginary world w_{i_j}, as usual:

$$V^+_{add}(p) = V^+_{copy}(p) \cup \Big(\bigcup_{j=1...m} \{w_{i_j} \mid p \text{ is a positive literal appearing in } \psi_j\} \Big)$$

Thirdly, it could be the case that the new box-formula being evaluated forces certain atomic formulas to be false at certain imaginary world; therefore, our algorithm should account for that[13]. In order to do that, we follow the "inverse" of the process we have been following, when clamping new atoms: this time, we look for any atom appearing as a *negative* literal in the corresponding clause φ_j, and, if the imaginary world w_{i_j} has been added to the valuation function of such atom, we *remove* it (in order to force that atom to being false in that world):

$$V^+_1(p) = V^+_{add}(p) \setminus \Big(\bigcup_{j=1...m} \{w_{i_j} \mid p \text{ is a negative literal appearing in } \psi_j\} \Big)$$

iii. *Create witness world*[14]: Conversely, if there are no new imaginary worlds created during the current execution of the algorithm (because, for instance, there are no diamond-formulas in the current set of scripts), then the box-formula being evaluated must create a so-called "witness-world" to account for the consequences described by it[15].
Therefore, if there exists no world w_i such that $(w^R, w_i, \alpha) \in R_E$ (which would be because there are no diamond-formulas referring to action α),

[13]This consideration is related with the fact that an agent could have an inconsistent set of scripts she believes in, as we discuss in Section 7.4.3.1.

[14]The relation between this step and step 3(b)ii could be understood, in terms of programming languages, as an "if ... else" statement. Namely, the algorithm first must check whether there already exists any imaginary world and, if it does, go through the corresponding branch of the algorithm; otherwise (else), if there are no already existing imaginary worlds, the algorithm must go through this current branch. Note, therefore, that both branches are *never* going to be executed for the same script, but rather just one of the two branches.

[15]It is worth mentioning that, in modal logic, the box operator $\Box$ has a sort of "vacuous" or "trivial" truth-condition: namely, if a world w has no accessibility relations at all, then every formula of the form $\Box\varphi$ would be vacuously true in there; as there are no worlds accessible from w, then *every* world accessible from w satisfies φ. That being said, we do not want our `EvoAlg` to conform to this fact, when evaluating a box-formula. One may argue that, if no diamond-formula about α has been previously evaluated, and so no new worlds have been created, then every box-formula about α could be true without the need of creating any world at all as a consequence of the agent imagining it. Nevertheless, this is not the way we reason when evaluating these kind of formulas, and so we still require our `EvoAlg` to create, at least, one witness world for a box-formula, in case there exist none yet.

then do the following:

A. *Create imaginary worlds*: Create m new imaginary possible worlds, one for each ψ_i in DNF$(\psi) = \psi_1 \vee \ldots \vee \psi_m$. This defines the expanded set of possible worlds as follows:

$$W^+ = W \cup \{w_1, \ldots, w_m\}$$

B. *Expand accessibility relation*: Create new R_E accessibility relations from the world of reference w^R to these new imaginary worlds as follows:

$$R_E^+ = R_E \cup \Big(\bigcup_{j=1\ldots m} \{(w^R, w_i, \alpha)\} \Big)$$

C. *Expand nominal structure*: Create and associate new nominals to these new imaginary worlds. This defines both the expanded set of nominals, and the expanded nominal function as follows:

$$\mathrm{NOM}^+ = \mathrm{NOM} \cup \{k_1, \ldots, k_m\}$$

$$N^+ = N \cup \Big(\bigcup_{i=1\ldots m} \{(k_i, w_i)\} \Big)$$

D. *Expand valuation function - Clamp new atoms*: Set the valuation of each new imaginary world w_i according to the consequences of the corresponding clause ψ_i in DNF(ψ):

$$V_1^+(p) = V(p) \cup \Big(\bigcup_i \{w_i \mid p \text{ is a positive literal appearing in } \psi_i\} \Big)$$

iv. *Expand valuation function - Import existing atoms*: After that, the `EvoAlg` must import all the atoms that are true at the world of reference w^R, provided they do not appear in ψ_i, for each new imaginary possible world w_i. Note that, even in the case of imaginary worlds created by using diamond-formulas, this must be done after having evaluated the box-formulas, as they may require to clamp further atoms in those worlds; importing existing atoms *before* evaluating box-formulas would probably import atomic formulas that would need to be removed afterwards as a requirement of the box-formulas being processed. The last step of the expanded valuation function, therefore, is as follows:

$$V^{+}(p) = V_1^{+}(p) \cup \Big(\bigcup_i \{w_i \mid w^R \in V_1^{+}(p) \text{ and } p \text{ is not a literal of } \psi_i\}\Big)$$

4. The `EvoAlg` has finished its execution: a new set of imaginary worlds have been created as a result of evaluating the consequences of a certain action or event α, according to the scripts believed by the agent, and which details what would likely happen in the imaginary world, if α took place in there. Conversely to the way the `DescrAlg` works, the `EvoAlg` handles the whole set of scripts referring to action α, and thus its execution is rather more complex than that of the `DescrAlg`.

7.4.3.1 Dealing with inconsistent beliefs

It may happen, when determining the set of scripts the agent believes in, that they turn out to be inconsistent, when put together. For example, the agent may believe that, after pouring tea into a cup, the kettle *can* be empty; this would be encoded by using a diamond-formula and would describe a new imaginary world. At the same time, it could happen that the agent believes in a script saying that, after pouring tea into a cup, the kettle *must* always end up being empty; this would be encoded by using a box-formula, and would affect every imaginary world created during the evaluation of the action "pour tea into a cup". As it can be seen, evaluating the box-formula would require to add the fact that "the kettle is empty" into an imaginary world containing the fact "the kettle is not empty". What should the `EvoAlg` do, in such cases?

First of all, there is a choice we have to make: do we want to allow our agent to have contradicting beliefs, or should we, as modelers, prevent that from happening? Even though belief revision is one of the topics that has received more attention in the literature of epistemic and doxastic logics (see, for instance, Chapter 7 in [50]), the tons of works and approaches about it makes it fall outside the scope of our work. As modelers, we ourselves define the model we want to represent, which captures what the agent believes in: if we include in the agent's beliefs, for whichever reasons, a set of scripts contradicting each other, so be it. If this is the model we want to represent, our logic should not take care of it and "clean up" our mess. It is worth noting that, as the set of scripts is only used through the execution of the `EvoAlg`, instead of being formulas holding at some specific possible worlds, then the fact that they may be altogether inconsistent will *not* make our model inconsistent, at least initially and before any execution of the `EvoAlg`. Therefore, for both motivational and technical reasons, there are no prior issues in allowing our model to represent a set of inconsistent scripts.

Once this has been made clear, a further question pops up into scene: even if the set of scripts does not make the model inconsistent at the beginning, what happens when the `EvoAlg` should use two different scripts ξ_1 and ξ_2, which would be altogether inconsistent? Before digging deeper, there is still one consideration we have to keep in mind regarding the form of the scripts:

- Inconsistencies will *never* come from diamond-formulas: by the way the algorithm handles them, each new diamond-formula generates a brand-new imaginary world; therefore, two different diamond-formulas would lead to two different imaginary worlds, and so no contradiction would arise. For instance, if the agent has both scripts $\varphi\langle\rightarrow\rangle^{\alpha}\psi$ and $\varphi\langle\rightarrow\rangle^{\alpha}\neg\psi$ in her beliefs, executing action α will create two different new imaginary worlds: a world w_1 in which ψ is the case, and a different world w_2 in which $\neg\psi$ is. This is neither a contradiction, nor a problem: the agent sees that both outcomes could be possible, but they would not be simultaneously true.

- Inconsistencies, however, may arise from box-formulas: by the way the algorithm handles them, the consequences of a box-formula must be added to every imaginary world already created in the previous step. This may turn into a problem if either a diamond-formulas has created a φ-world, and a box-formula says that every world must be a $\neg\varphi$-world, or if two different box-formulas state that every world must be a φ-world, and that every world must be a $\neg\varphi$-world, respectively.

How should the `EvoAlg` behave, once we state that we initially allow for sets of inconsistent scripts to be altogether in our model, and once we see where could inconsistencies come from?

So far, the way the `EvoAlg` works prevents inconsistencies from arising. In particular, note that, when processing a certain script ξ_1, the consequences ψ_1 described by it are prioritized and "clamped" into the imaginary world; however, as soon as script ξ_1 has been processed, consequences ψ_1 lose their "privileged" status. In particular, if the `EvoAlg` was now evaluating a box-formula ξ_2, describing a certain set of consequences amounting to $\psi_2 \equiv \neg\psi_1$, then this new consequences $\neg\psi_1$ would be prioritized and clamped into the imaginary world where ψ_1 was initially set; as ψ_1 is no longer given priority, it would simply be overridden by $\neg\psi_1$, and so no inconsistent imaginary world would arise. In other words: as the `EvoAlg` prioritizes each new script being evaluated within the loop, the consequences of each new script will override whatever has already been set in the imaginary worlds that are being created in the current execution. If it turns out that two scripts ξ_1 and ξ_2 describe contradictory consequences, then the one which is evaluated the last will override the first one.

Let's take a closer look from a more motivational point of view: does it make sense that the scripts that are evaluated the last override the ones evaluated the former? By the way we sort our queue of scripts in the Step 2 of the `EvoAlg`, we argue that it does. We require every diamond-formula to be sorted before any box-formula; among them, we require formulas with more simple antecedents to be sorted before the ones with more complex antecedents.

Regarding diamond-formulas and box-formulas, we claim that, if one kind of formulas were to override what has already been stated by the other one, then the box-formulas should be the ones given priority. Consider an agent who believes that, after pouring tea into a cup, then the cup could still be empty (maybe because there is no tea in the kettle), and who also believes that, after pouring tea into a cup, there must necessarily be some tea in the cup (even if it's just a few drops). It may seem challenging to accept that an agent could simultaneously believe these two scripts to be true, specially when put together, but let's just suppose so for the sake of the example. Whereas the first script describes a situation that *could* arise, the second one describes a situation that *must* follow; if both scripts are together inconsistent, and we have to take only one of the two, we argue that it makes more sense to stick with the rule stating what necessarily follows, rather than the one stating what could follow.

Regarding the complexity of the antecedent, we also argue that the way we sort the scripts makes sense, when considering the possible override of consequences coming from contradictory scripts. Take the previous example, and consider an agent who believes that "if she pours tea into a cup, then the cup could still be empty"; similarly, the agent also believes that "if she pour tea into a cup *and* the kettle is not empty, then the cup must *not* be empty". While evaluating these pair of scripts sorted in that way, the agent first considers a scenario that could follow, based on a certain conditions; then, when moving to another scripts, the agent considers a formula that takes into account more conditions of the initial situation than the former script did, and this new script describe a consequence that differs from the former one. While focusing on just certain details of the initial scenario (say: the agent pours tea into a cup), she recognized that it could be that the cup is still empty after that; nevertheless, while taking into account but the fact that the she pours tea into a cup, and the fact that the kettle is not empty, then she realizes that the cup cannot, after this, be empty. The second script is more specific, as it captures more conditions holding at the initial situation, and so it must be given priority over the first one, if they are altogether contradictory.

Therefore, and recalling eveyrthing we have just said in these last paragraphs, it is now impossible, taking into account how the `EvoAlg` works, that two different scripts clamp two contradictory consequences into an imaginary world being created; this is not to say, however, that this is the only stance we could take regarding this problem. In Section 7.8.2 we go back to this topic and discuss an alternative way we could use to detect inconsistent sets of scripts and prevent the agent from using them in an execution of the `EvoAlg`.

7.4.4 The `Add` Algorithm

Note that this algorithm is almost identical to the `InitAlg`. This is consistent both with what the algorithm (and the process it represents) does, which is to voluntarily clamp a new premise into the imagining, and with Langland-Hassan's theory (see Section 2.3.3), which claims that these kind of additions are, in fact, a cyclical involvement of the initial addition of a premise that characterizes an imaginary scenario.

Even though they may be similar, there is still an important difference that justifies that we have defined two distinct algorithms: whereas the `InitAlg` creates new imaginary worlds by using a *real world* as its referent, the `AddAlg` creates new imaginary worlds, but by adding a new premise into an *already existing* imaginary world. Consequently, we can say that the range of possible worlds available to both algorithms is complementary, in the sense that the `InitAlg` can only use real possible worlds, and the `AddAlg` can only use imaginary possible worlds.

The `AddAlg`, thus, corresponds to the *Unscripted Addition* process of the Common Frame for Imagination Acts, and it is responsible for adding a new premise into an imaginary world that would not be derived by following neither any factual rule, nor any script. Unlike the `InitAlg`, the `AddAlg` needs an imaginary possible world as the world of reference w^R, and a certain formula $\delta \in \mathrm{FORM}^*$. The `AddAlg` is called by evaluating the dynamic formula $\mathrm{Add}(\delta)$ at a possible world w^R as follows:

$$\texttt{AddAlg}(\delta, w^R)$$

The algorithm follows these steps:

1. *Check initial conditions*: If world w^R is a real possible world, do nothing. Similarly, if δ is contradictory (that is, if $\delta \equiv \bot$), do nothing.

2. *Compute DNF*: In order to handle the formula in an efficient way, we compute the *Disjunctive Normal Form* (DNF) of δ, to which we refer as DNF(δ). In the following steps of the `ImgAlg`, we refer to the clauses that form the DNF(δ) as follows: DNF(δ) $= \delta_1 \vee \ldots \vee \delta_n$.

3. *Create imaginary worlds*: Create a new imaginary possible world $w_1, \ldots, w_n$ for each clause $\delta_1, \ldots, \delta_n$ in DNF(δ). This defines the *expanded set of possible worlds* as follows:
$$W^+ = W \cup \{w_1, \ldots, w_n\}$$

4. *Expand accessibility relation*: Once the new possible worlds have been created, the `AddAlg` must create the new addition relations expressing that, by imagining formula δ at the world of reference w^R, the agent has crated new imaginary worlds $w_1, \ldots, w_n$. This defines the *expanded set of addition relations* as follows:
$$R_A^+ = R_A \cup \Big(\bigcup_{i=1 \ldots n} \{(w^R, w_i, \delta)\} \Big)$$

5. *Expand nominal structure*: Now, the `AddAlg` must add a set of new nominals to refer to the newly created imaginary worlds. This defines both the *expanded set of nominals*, by adding one new nominal k_i for each new possible world w_i, and the *expanded nominal function*, which is an extension of N:
$$\mathrm{NOM}^+ = \mathrm{NOM} \cup \{k_1, \ldots, k_n\}$$
$$N^+ = N \cup \Big(\bigcup_{i=1 \ldots n} \{(k_i, w_i)\} \Big)$$

6. *Expand valuation function*: The `AddAlg` must expand the valuation function to account for the new imaginary possible worlds. In order to do so, the algorithm must account for both the atoms that are present in each δ_i, and also for the atoms that are true in the world of reference w^R and which should be imported to the new imaginary worlds, provided they do not appear in δ_i. Therefore, the definition of the *expanded valuation function* involves two different phases:

(a) *Clamp new atoms*: The `AddAlg` must set the new valuation functions according to the atoms p appearing in δ_i, for each new imaginary possible world w_i:

$$V_1^+(p) = V(p) \cup \left(\bigcup_i \{w_i \mid p \text{ is a positive literal appearing in } \delta_i\} \right)$$

(b) *Import existing atoms*: Then, it must import all the atoms that are true at the world of reference w^R, provided they do not appear in δ_i, for each new imaginary possible world w_i:

$$V^+(p) = V_1^+(p) \cup \left(\bigcup_i \{w_i \mid w^R \in V_1^+(p) \text{ and } p \text{ is not a literal of } \delta_i\} \right)$$

7. The `AddAlg` has finished its execution: a new set of imaginary possible worlds satisfying δ has been created, and these worlds are now accessible through the addition relation R_A from the world of reference w^R.

7.5 Semantics

Having presented the four algorithms of the Logic of Imagination Acts, we define its semantics as follows, for a Model for Imagination Acts $\mathcal{M}$, a possible world $w \in W$, and being $p \in \text{ATOM}$, $i \in \text{NOM}$, $\{\varphi, \psi\} \subseteq \text{FORM}$, $\delta \in \text{FORM}^*$, $\zeta \in \text{FACT}$, and $\alpha \in \text{ACT}$:

Propositional formulas:

$\mathcal{M}, w \vDash p$	iff	$w \in V(p)$
$\mathcal{M}, w \vDash \neg\varphi$	iff	$\mathcal{M}, w \nvDash \varphi$
$\mathcal{M}, w \vDash \varphi \wedge \psi$	iff	$\mathcal{M}, w \vDash \varphi$ and $\mathcal{M}, w \vDash \psi$
$\mathcal{M}, w \vDash \varphi \vee \psi$	iff	$\mathcal{M}, w \vDash \varphi$ or $\mathcal{M}, w \vDash \psi$
$\mathcal{M}, w \vDash \varphi \rightarrow \psi$	iff	$\mathcal{M}, w \vDash \neg\varphi$ or $\mathcal{M}, w \vDash \psi$

Hybrid and modal formulas:

$\mathcal{M}, w \vDash i$	iff	$N(i) = w$ and, for every $v \in W$, if $\mathcal{M}, v \vDash i$, then $v = w$
$\mathcal{M}, w \vDash @_i\varphi$	iff	there exists a world $v \in W$ such that $N(i) = v$ and $\mathcal{M}, v \vDash \varphi$
$\mathcal{M}, w \vDash \langle * \rangle\varphi$	iff	there exists a world $v \in W$ such that either $(w, v, \delta) \in R_I$, $(w, v, \zeta) \in R_D, (w, v, \alpha) \in R_E$ or $(w, v, \delta) \in R_A$ and $\mathcal{M}, v \vDash \varphi$

Dynamic imagination formulas:

$\mathcal{M}, w \vDash \mathrm{Init}(\delta)$	iff	δ is not contradictory ($\delta \not\equiv \bot$) and either there already exists $v \in W$ such that $(w, v, \delta) \in R_I$ or, after executing `InitAlg`(δ, w), $\mathcal{M}$ is expanded into $\mathcal{M}^+$
$\mathcal{M}, w \vDash \mathrm{Descr}(\zeta)$	iff	either there already exists $v \in W$ such that $(w, v, \zeta) \in R_D$ or, after executing `DescrAlg`(ζ, w), $\mathcal{M}$ is expanded into $\mathcal{M}^+$
$\mathcal{M}, w \vDash \mathrm{Evo}(\alpha)$	iff	either there already exists $v \in W$ such that $(w, v, \alpha) \in R_E$ or, after executing `EvoAlg`(α, w), $\mathcal{M}$ is expanded into $\mathcal{M}^+$
$\mathcal{M}, w \vDash \mathrm{Add}(\delta)$	iff	δ is not contradictory ($\delta \not\equiv \bot$) and either there already exists $v \in W$ such that $(w, v, \delta) \in R_A$ or, after executing `AddAlg`(δ, w), $\mathcal{M}$ is expanded into $\mathcal{M}^+$

Static imagination formulas:

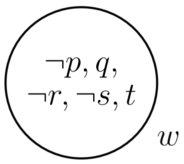
$\mathcal{M}, w \vDash \langle\delta\rangle^I\varphi$	iff	there is some $v \in W$ s.t. $(w, v, \delta) \in R_I$ and it is the case that $\mathcal{M}, v \vDash \varphi$ and $\mathcal{M}, v \vDash \delta$
$\mathcal{M}, w \vDash \langle\zeta\rangle^D\varphi$	iff	there is some $v \in W$ s.t. $(w, v, \zeta) \in R_D$ and it is the case that $\mathcal{M}, v \vDash \varphi$
$\mathcal{M}, w \vDash \langle\alpha\rangle^E\varphi$	iff	there is some $v \in W$ s.t. $(w, v, \alpha) \in R_E$ and it is the case that $\mathcal{M}, v \vDash \varphi$
$\mathcal{M}, w \vDash \langle\delta\rangle^A\varphi$	iff	there is some $v \in W$ s.t. $(w, v, \delta) \in R_A$ and it is the case that $\mathcal{M}, v \vDash \varphi$ and $\mathcal{M}, v \vDash \delta$

Recall that the expanded model $\mathcal{M}^+$ is computed by any execution of either the `InitAlg`, the `DescrAlg`, the `EvoAlg` or the `AddAlg`; elements W^+, R_I^+, R_D^+, R_E and R_A^+ belong to the expanded model $\mathcal{M}^+$, and are also computed by the execution of the previous algorithms. Propositional and hybrid formulas work as usual (see the first sections of Chapter 3 for a brief reminder of how they work).

7.6 An Example

In order to show how the Logic for Imagination Acts works, let's follow an example showing how each one of the four algorithms work. Recall that, aside from the `InitAlg`, which requires a real world in order to be executed, the order in which the algorithms are applied is up to the modeler (or, intuitively, up to the decision of the agent to develop the imagining in such and such way). Besides, and although we keep always moving "deeper" in this example, it is possible to apply an algorithm to an already existing world that has already been used as the world of reference for some other process. For instance, consider and agent who could have already imagined a certain world w; the agent can now perform a certain process creating a new world v, and then, instead of keep imagining over v, she could "go back", focus again on world w, and perform a different process leading to a new world z. In other words: the execution of the algorithms in neither limited by a certain order, nor it is required to be applied in a world of reference which was not used before. Our main goal, in this approach, is precisely to capture acts of imagination characterized by the freedom of choice we actually have when deciding how to elaborate on an imagining, and our logic captures this fact.

Let's start with the example. Consider an initial Model for Imagination Acts $\mathcal{M}$, which includes a single real possible world w that captures the state of affairs the agent believes to be the case. Besides, we have a set ACT of actions that could be carried out in the example, a set FACT, and a set SCRIPT representing the factual rules and the scripts the agent believes in, respectively. For the sake of simplicity, we include, in this example, a small quantity of factual rules, and we only include those scripts related to the action we will be using in it; recall, nevertheless, that such sets of factual rules and scripts are meant to include as many different formulas as beliefs the agent has about how an imaginary situation could typically be elaborated. The initial model, before executing any process of imagination, is as shown in Figure 7.5.



ACT: $\{\alpha\}$

FACT:
$\{p\langle\rightarrow\rangle\neg t,\ \neg s \wedge t\langle\rightarrow\rangle\neg r,$
$r \wedge s\langle\rightarrow\rangle(\neg q \vee t \vee (q \wedge \neg s))\}$

SCRIPT:
$\{p \wedge q\langle\rightarrow\rangle^{\alpha}\neg s,\ s \wedge t\langle\rightarrow\rangle^{\alpha}p \wedge \neg q,$
$t\langle\rightarrow\rangle^{\alpha}p,\ s[\rightarrow]^{\alpha}\neg r\}$

Figure 7.5: The initial model, before executing any process related to imagination.

The `InitAlg`

As we have already explained in this chapter, every act of imagination is required to start by an application of the Initialization process, which is captured by the `InitAlg`, and which requires a world of reference w^R and a formula $\delta \in \text{FORM}^*$. Let's set $w^R = w$ and $\delta = \neg p \rightarrow (r \wedge s)$, which leads to the following call:

$$\texttt{InitAlg}(w, \neg p \rightarrow (r \wedge s))$$

This call to the `InitAlg` goes as follows:

1. *Check initial conditions*: World w must be checked to asses whether it is a real possible world, and indeed it is; there is no relation $(v, w, \#)$, for any other possible world $v \in W$ and any formula $\#$ in any of the accessibility relations R_I, R_D, R_E or R_A, which means that world w cannot be accessed through any process involved in an act of imagination. Similarly, formula $\neg p \rightarrow (r \wedge s)$ is not contradictory (that is: $\neg p \rightarrow (r \wedge s) \nvdash \bot$). Therefore, the `InitAlg` continues with its execution.

2. *Compute DNF*: The `InitAlg` must compute $\text{DNF}(\neg p \rightarrow (r \wedge s))$, which corresponds to $p \vee (r \wedge s)$ (we omit the details of the procedure for computing the DNF).

3. *Create imaginary worlds*: As $\text{DNF}(\neg p \rightarrow (r \wedge s)) = p \vee (r \wedge s)$, the `InitAlg` must create two new imaginary possible worlds (one for each clause), which we will call v_1 and v_2, and which define the expanded set of possible worlds as follows:

$$W^+ = W \cup \{v_1, v_2\}$$

4. *Expand accessibility relation*: The newly created imaginary worlds must be made accessible through the R_I relation, expressing that, by imagining $\neg p \to (r \wedge s)$ while taking world w as the world of reference, the agent creates two different imaginary worlds; this defines the expanded set of initiation relations as follows:

$$R_I^+ = R_I \cup \{(w, v_1, \neg p \to (r \wedge s)), (w, v_2, \neg p \to (r \wedge s))\}$$

5. *Expand nominal structure*: In order to name those newly created imaginary worlds, the `InitAlg` must create new nominals and associate them to the new worlds. Let's call j_1 and j_2 the two new nominals created in this step; this defines the expanded set of nominals and the expanded nominal function as follows:

$$\mathrm{NOM}^+ = \mathrm{NOM} \cup \{j_1, j_2\}$$

$$N^+ = N \cup \{(j_1, v_1), (j_2, v_2)\}$$

6. *Expand valuation function*: At this points, the `InitAlg` must determine the atomic valuation of the new imaginary worlds. This must be done in two different steps.

   (a) *Clamp new atoms*: First, the `InitAlg` must clamp the atomic formulas represented in each clause for each corresponding world. In this particular example, we associate clause p to world v_1, and clause $r \wedge s$ to world v_2; this defines the first step of the expanded valuation function as follows[16]:

$$V_1^+(p) = V(p) \cup \{v_1\}$$

$$V_1^+(r) = V(r) \cup \{v_2\}$$

$$V_1^+(s) = V(s) \cup \{v_2\}$$

   (b) *Import existing atoms*: Once the atomic formulas of δ have been clamped in the new imaginary worlds, the `InitAlg` must import those atomic formulas that hold in the world of reference w, as long as they do not appear in the clause that has been associated to the corresponding new imaginary world. As v_1 has been associated to clause p, the `InitAlg` must look for the value of atomic

[16] Recall that, although the algorithm loops over every atomic formula, we only write, in this example, the expansions of V that concern the atoms clamped in the new imaginary worlds.

propositions q, r, s and t; conversely, as v_2 has been associated to clause $r \wedge s$, the algorithm must look for the value of propositions p, q and t. This defines the second step of the expanded valuation function in the following way (note that we do not need to set the valuation function for negated atomic formulas; in particular, as p is false in the world of reference, we keep it false in the new imaginary worlds by not adding those worlds to the valuation function for p):

$$V^+(q) = V_1^+(q) \cup \{v_1, v_2\}$$

$$V^+(t) = V_1^+(t) \cup \{v_1, v_2\}$$

7. The `InitAlg` has finished its execution. The model $\mathcal{M}$, at this point, has been expanded into model $\mathcal{M}^+$, and corresponds to the model shown in Figure 7.6. For the sake of clarity, we also write the negated atomic formulas at each corresponding world, and we highlight in bold font the formulas clamped in step 6a.

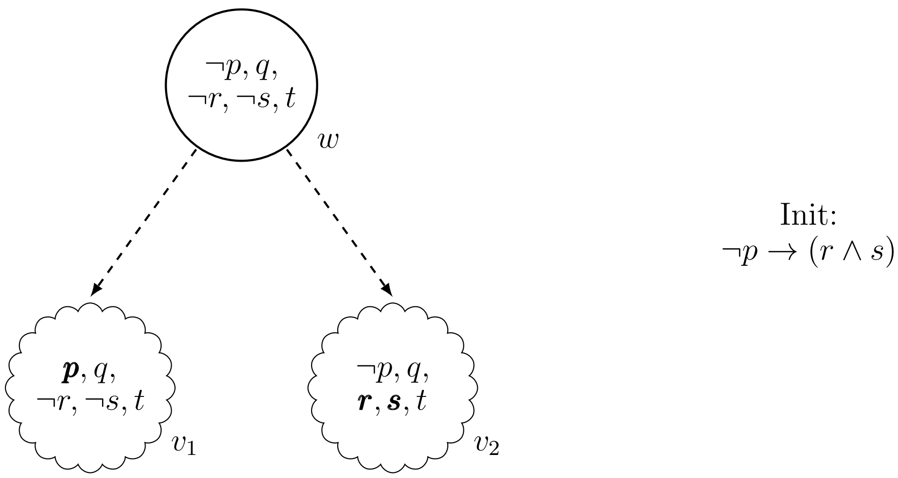


Figure 7.6: The model after the execution of the `InitAlg`.

As soon as the agent has initiated an act of imagination through an execution of the `InitAlg`, a certain set of imaginary possible worlds is created (in this case, worlds v_1 and v_2). Once an imaginary world exists, the agent can start developing it by applying any of the other three processes distinguished within acts of imagination.

The `DescrAlg`

In this particular example, our agent will keep elaborating the static details of one imaginary world by following a certain factual rule she believes in. As we have shown in Figure 7.5, in this model the agent believes in the following set of factual rules:

$$\text{FACT} = \{p\langle\rightarrow\rangle\neg t, \neg s \wedge t\langle\rightarrow\rangle\neg r, r \wedge s\langle\rightarrow\rangle(\neg q \vee t \vee (q \wedge \neg s))\}$$

In order to develop the static details of an imaginary world, the `DescrAlg` needs a world of reference w^R to elaborate on, and a factual rule $\zeta \in \text{FACT}$. In the example, our agent chooses to elaborate on the imaginary world v_2, thus making it the world of reference, and chooses to follow the factual rule $r \wedge s\langle\rightarrow\rangle(\neg q \vee t \vee (q \wedge \neg s))$. This calls the `DescrAlg` in the following way:

$$\texttt{DescrAlg}(v_2, r \wedge s\langle\rightarrow\rangle(\neg q \vee t \vee (q \wedge \neg s)))$$

The execution of the algorithm follows these steps:

1. *Check initial conditions*: World v_2 is not a real world, as there exists a relation $(w, v_2, \neg p \rightarrow (r \wedge s)) \in R_I$ making v_2 an imaginary world. Therefore, the `DescrAlg` can continue with its execution.

2. *Compute DNF*: The factual rule the agent chooses to elaborate the imaginary scenario is $\zeta = r \wedge s\langle\rightarrow\rangle(\neg q \vee t \vee (q \wedge \neg s))$, which indeed has the required form $\varphi\langle\rightarrow\rangle\psi$, for $\{\varphi, \psi\} \in \text{FORM}^*$. The `DescrAlg` must check whether the antecedent of ζ, which is $r \wedge s$, is true at v_2: and, indeed, it is. Therefore, the factual rule can be applied to elaborate on the state of affairs described by v_2. The `DescrAlg` must now compute the DNF[17] of the consequent of ζ; that is, $\text{DNF}(\neg q \vee t \vee (q \wedge \neg s)) = \neg q \vee t \vee (q \wedge \neg s)$ (in this case, it is already in DNF, so no computations need to be done).

3. *Create imaginary worlds*: As $\text{DNF}(\neg q \vee t \vee (q \wedge \neg s))$ has 3 different clauses, the `DescrAlg` must create 3 new imaginary possible worlds, which we will call u_1, u_2 and u_3. This defines the expanded set of possible worlds as follows:

$$W^+ = W \cup \{u_1, u_2, u_3\}$$

[17]For the sake of simplicity, we have chosen to use, during the rest of this example, formulas that are already in DNF. Nevertheless, computing the DNF of any propositional formula can be done within the algorithm without any problem.

4. *Expand accessibility relation*: Once these new imaginary worlds have been created, the `DescrAlg` must make them accessible from the world of reference through relation R_D, thus defining the expanded description relation:

$$\begin{aligned} R_D^+ = R_D \cup \{ \quad & (v_2, u_1, r \wedge s\langle\rightarrow\rangle(\neg q \vee t \vee (q \wedge \neg s))), \\ & (v_2, u_2, r \wedge s\langle\rightarrow\rangle(\neg q \vee t \vee (q \wedge \neg s))), \\ & (v_2, u_3, r \wedge s\langle\rightarrow\rangle(\neg q \vee t \vee (q \wedge \neg s)))\} \end{aligned}$$

5. *Expand nominal structure*: Similarly to what happened in the `InitAlg`, the `DescrAlg` must create new nominals and associate them to the new imaginary worlds. Let's call k_1, k_2 and k_3 the new nominals created in this step; this defines the expanded set of nominals and the expanded nominal function as follows:

$$\mathrm{NOM}^+ = \mathrm{NOM} \cup \{k_1, k_2, k_3\}$$

$$N^+ = N \cup \{(k_1, u_1), (k_2, u_2), (k_3, u_3)\}$$

6. *Expand valuation function*: Now, the `DescrAlg` must expand the valuation function for the new imaginary worlds, according to the clauses of the consequent of ζ that has been used to create them. This process works the same way it does in the case of the `InitAlg`, with the difference that, instead of the whole formula ζ, in this case the algorithm only uses the consequent of ζ to determine the atomic formulas that must be clamped at each imaginary world.

   (a) *Clamp new atoms*: First, and for each new imaginary world and each clause in $\mathrm{DNF}(\neg q \vee t \vee (q \wedge \neg s))$, the algorithm must set the first step in the expansion of the valuation function as follows (recall that we do not added negated atomic formulas in the valuation function):

$$V_1^+(t) = V(t) \cup \{u_2\}$$

$$V_1^+(q) = V(q) \cup \{u_3\}$$

   (b) *Import existing atoms*: Second, the algorithm must import those atomic formulas that hold at the world of reference v_2, but which do not appear in the corresponding clause of $\mathrm{DNF}(\neg q \vee t \vee (q \wedge \neg s))$, for each new imaginary world.

This second step finishes the expansion of the valuation function as follows:

$$V^+(q) = V_1^+(q) \cup \{u_2\}$$

$$V^+(r) = V_1^+(r) \cup \{u_1, u_2, u_3\}$$

$$V^+(s) = V_1^+(s) \cup \{u_1, u_2\}$$

$$V^+(t) = V_1^+(t) \cup \{u_1, u_3\}$$

7. The `DescrAlg` has finished its execution: the imaginary world v_2 has been further elaborated by following a certain factual rule the agent believes in, and which specifies what else could be the case in that imaginary world. The resulting model corresponds to Figure 7.7. Note that we only draw the part of the model that is involved in this particular step for the sake of simplicity. Similarly, we write the negated atomic formulas that hold at each possible world, and highlight in bold font those atomic formulas that were clamped in the new imaginary worlds in step 6a.

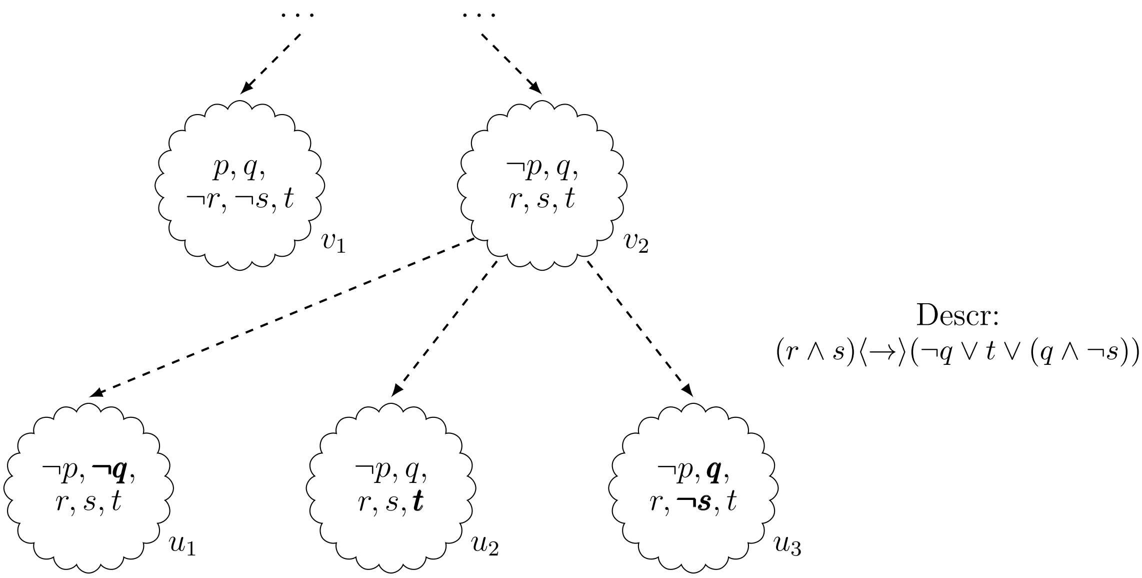


Figure 7.7: The agent elaborates on the static details of an imaginary world.

The new set of imaginary possible worlds created by this execution of the `DescrAlg` represent what else could be the case at the imaginary world v_2, according to a factual rule the agent believes.

Notice that, in this particular example, worlds v_2 and u_2 are equivalent, in terms of their atomic valuation: it could happen that, either as a result of the atomic formulas we import from the world of reference, or as a result of the new atoms being clamped in the new imaginary world, the agent ends up imagining a possible world which is, in turn, equivalent to the one she took as her reference. Our stance towards this is the same we took in the Logic of Imaginary Scenarios (detailed in Section 4.6.1), in which we could also obtain equivalent worlds.

The `EvoAlg`

Let's say that now our agent wants to imagine what would happen if she performed action α, or if event α took place. This corresponds to a dynamic elaboration of the imagining in which an action α takes place, and in which the scripts the agent believes in determine what would be the consequences of such action in the imagining.

Conversely to the `DescrAlg`, which used factual rules in FACT, the `EvoAlg` does *not* use one of the believed scripts in SCRIPT as an argument (although they are used by the algorithm later), but rather it uses the action that the agent imagines to happen, and which belongs to ACT. Let's say the agent wants to elaborate on world u_2 by imagining that α happens in there. This calls the `EvoAlg` as follows:

$$\texttt{EvoAlg}(u_2, \alpha)$$

The algorithm follows these steps:

1. *Check initial conditions*: We know that world u_2 is an imaginary world, as it appears in $(v_2, u_2, (r \wedge s)\langle \rightarrow \rangle(\neg q \vee t \vee (q \wedge \neg s))) \in R_D$. Action α does appear in some formulas in SCRIPT, which means that the agent has certain beliefs about the outcomes that action α would carry with it. Therefore, the `EvoAlg` continues with its execution.

2. *Create queue of scripts*: Now, the `EvoAlg` must take every formula $\xi \in$ SCRIPT in which α appears (either as $\langle \rightarrow \rangle^\alpha$ or as $[\rightarrow]^\alpha$), and create a queue of scripts S^α sorted according to the following criteria:

   - First, all the diamond-formulas about α, prioritizing those with the simplest antecedent.
   - Second, all the box-formulas about α, prioritizing those with the simplest antecedent.

Considering the scripts we have in this example, and which appear in Figure 7.5, queue S^α is defined as follows:

$$S^\alpha = \langle t\langle\rightarrow\rangle^\alpha p, p \wedge q\langle\rightarrow\rangle^\alpha \neg s, s \wedge t\langle\rightarrow\rangle^\alpha p \wedge \neg q, s[\rightarrow]^\alpha \neg r\rangle$$

3. *Loop through the scripts - 1*: Take the first script $\xi \in S^\alpha$, which corresponds to $t\langle\rightarrow\rangle^\alpha p$:

   (a) *Evaluate diamond-formulas*: As $t\langle\rightarrow\rangle^\alpha p$ is a diamond-formula, the `DescrAlg` must check whether its antecedent, t, holds at the world of reference u_2; in other words, the algorithm must check whether:

   $$\mathcal{M}, u_2 \vDash t$$

   In this case, t does hold at u_2, so the algorithm continues within the loop.

   i. *Compute DNF*: The `DescrAlg` must compute the DNF of the consequent of $t\langle\rightarrow\rangle^\alpha p$; that is, $\mathrm{DNF}(p) = p$.

   ii. *Create imaginary worlds*: For each clause in $\mathrm{DNF}(p)$, the algorithm must create one new imaginary possible world. As there is only one clause in $\mathrm{DNF}(p)$, the algorithm creates one new possible world, which we call z_1, and defines the expanded set of possible worlds as follows:

   $$W^+ = W \cup \{z_1\}$$

   iii. *Expand accessibility relation*: The algorithm must expand R_E in order to make the new imaginary world z_1 accessible as a consequence of action α taking place in u_2:

   $$R_E^+ = R_E \cup \{(u_2, z_1, \alpha)\}$$

   iv. *Expand nominal structure*: The `DescrAlg` must create a new nominal to name the new imaginary world; let's call this nominal l_1:

   $$\mathrm{NOM}^+ = \mathrm{NOM} \cup \{l_1\}$$

   $$N^+ = N \cup \{(l_1, z_1)\}$$

   v. *Expand valuation function - Clamp new atoms*: The algorithm must clamp the atomic formulas appearing in the consequent of the script being evaluated right now; that is, it must add the atomic proposition p into the valuation of the new imaginary world z_1 as follows:

$$V^+(p) = V(p) \cup \{z_1\}$$

4. *Loop through the scripts - 2*: Take next script $\xi \in S^\alpha$ corresponding to $p \wedge q\langle\rightarrow\rangle^\alpha \neg s$:

   (a) *Evaluate diamond-formulas*: As $p \wedge q\langle\rightarrow\rangle^\alpha \neg s$ is a diamond-formula, the `DescrAlg` must check whether its antecedent, $p \wedge q$, holds at the world of reference u_2; in other words, the algorithm must check whether:

   $$\mathcal{M}, u_2 \vDash p \wedge q$$

   In this case, $p \wedge q$ does *not* hold at u_2, so the algorithm must skip this script and look for the next one in S^α.

5. *Loop through the scripts - 3*: The algorithm must now take the next script $\xi \in S^\alpha$, which corresponds to $s \wedge t\langle\rightarrow\rangle^\alpha p \wedge \neg q$:

   (a) *Evaluate diamond-formulas*: As $s \wedge t\langle\rightarrow\rangle^\alpha p \wedge \neg q$ is a diamond-formula, the `DescrAlg` must check whether its antecedent, $s \wedge t$, holds at the world of reference u_2; in other words, the algorithm must check whether:

   $$\mathcal{M}, u_2 \vDash s \wedge t$$

   In this case, $s \wedge t$ does hold at u_2, so the algorithm continues within the loop.

   i. *Compute DNF*: The `EvoAlg` must compute the DNF of the consequent of $s \wedge t\langle\rightarrow\rangle^\alpha p \wedge \neg q$; that is, $\text{DNF}(p \wedge \neg q) = p \wedge \neg q$.

   ii. *Create imaginary worlds*: For each clause in $\text{DNF}(p \wedge \neg q)$, the algorithm must create one new imaginary possible world. As there is only one clause in $\text{DNF}(p \wedge \neg q)$, the algorithm creates one new possible world, which we call z_2, and defines the expanded set of possible worlds as follows:

   $$W^+ = W \cup \{z_2\}$$

   iii. *Expand accessibility relation*: The algorithm must expand R_E in order to make the new imaginary world accessible as a consequence of action α taking place:

$$R_E^+ = R_E \cup \{(u_2, z_2, \alpha)\}$$

iv. *Expand nominal structure*: The `DescrAlg` must create a new nominal to name the new imaginary world; let's call this nominal l_2:

$$\text{NOM}^+ = \text{NOM} \cup \{l_2\}$$

$$N^+ = N \cup \{(l_2, z_2)\}$$

v. *Expand valuation function - Clamp new atoms*: The algorithm must clamp the atomic formulas appearing in the consequent of the script being evaluated right now. As $\text{DNF}(p \wedge \neg q) = p \wedge \neg q$, formulas p and $\neg q$ must hold at world z_2; therefore, the algorithm must add the atomic proposition p into the valuation of the new imaginary world z_2 as follows. Note, however, that unlike the previous case, q is *required* to be false at z_2, whereas its truth-value does not matter in z_1; when importing atomic formulas from u_2, we will see how this ends up altering the truth-value of q in both worlds:

$$V^+(p) = V(p) \cup \{z_2\}$$

6. *Loop through the scripts - 4*: Take next script $\xi \in S^\alpha$ corresponds to $s[\rightarrow]^\alpha \neg r$:

(a) *Evaluate box-formulas*: Unlike the previous cases, $s[\rightarrow]^\alpha \neg r$ is a box-formula, so it must be processed differently by the `EvoAlg`. Firstly, and as it happened before, the algorithm must check whether the antecedent of $s[\rightarrow]^\alpha \neg r$ (that is: s), holds at world u_2:

$$\mathcal{M}, u_2 \vDash s$$

As s holds at u_2, the algorithm continues its execution within the loop.

i. *Compute DNF*: Compute the DNF of the consequent of $s[\rightarrow]^\alpha \neg r$, which corresponds to $\text{DNF}(\neg r) = \neg r$.

ii. *Loop existing worlds - 1*: Now, if there exists at least one world z_i such that $(u_2, z_i, \alpha) \in R_E$ (that is, if the evaluation of diamond-formulas has created at least one new imaginary possible world), then each possible consequence of $\text{DNF}(\neg r)$ must be clamped in the imaginary worlds. There are, indeed, two different possible worlds z_1 and z_2 accessible from u_2 through an α-relation. The EvoAlg must consider the first one, z_1, and do the following:

A. *Create imaginary worlds*: Create $j = m - 1$ (for $\text{DNF}(\psi) = \psi_1 \vee \ldots \vee \psi_m$) new imaginary possible worlds. In this case, as $\text{DNF}(\neg r) = \neg r$, the algorithm must create $j = 1 - 1 = 0$ new possible worlds (the script being evaluated has only one possible outcome, so there is no need to consider different results). Therefore, the set of possible worlds does not need to be expanded in this case.

B. *Expand accessibility relation*: As no new possible worlds are created, the algorithm must skip this step.

C. *Expand nominal structure*: Similarly, as there are no new possible worlds to name, the algorithm must skip this step.

D. *Expand valuation function - Clamp new atoms*: The algorithm must now clamp the new atomic formulas in the imaginary possible worlds. Note that the expansion of the valuation function is defined, in Step 3(b)iiD of Section 7.4.3, in different steps or phases.

   First, and through V^{+}_{copy}, the EvoAlg must replicate, in the newly created imaginary possible worlds, the valuation of the already existing imaginary world, created as a result of evaluating diamond-formulas (which is z_1 in this case); as there are no new imaginary worlds created as a consequence of the current box-formula, this step can be omitted. Once this has been done, the algorithm must loop over every new imaginary world (which, in this case, means only z_1, as no new worlds has been created as a result of the box-formula), and define V^{+}_{add} by clamping the corresponding atoms appearing as positive literals in their related clause. In this case, clause $\neg r$ corresponds to z_1, so no positive atoms need to be clamped in the imaginary world.

   Finally, the EvoAlg must loop over every atom appearing as a negative literal in $\neg r$ (in this case, r) and, if the imaginary world associated to the current clause (which would be z_1 in this case) satisfies such atom, remove it from $V^{+}_{add}(r)$. Nevertheless, this is neither the case, so the valuation function remains the same as is was.

iii. *Loop existing worlds - 2*: Now the EvoAlg must consider the next world that appeared in the accessibility relation, z_2, and repeat the same steps:

A. *Create imaginary worlds*: Similarly, as DNF($\neg r$) $= \neg r$, and so $j = 1 - 1 = 0$, the `EvoAlg` does not have to create any new imaginary world.

B. *Expand accessibility relation*: As no new possible worlds are created, the algorithm must skip this step.

C. *Expand nominal structure*: Similarly, as there are no new possible worlds to name, the algorithm must skip this step.

D. *Expand valuation function - Clamp new atoms*: As no new possible worlds has been created in this case, the `EvoAlg` should, again, expand the valuation function of only the already existing imaginary world being evaluated in this loop, which is z_2. Similarly to what happened in the previous loop with world z_1, nothing needs to be done in the valuation function, as there are no positive literals in $\neg r$ (so z_2 should not be added to the valuation function for any atom), and z_2 does not appear in $V(r)$ (and so it neither has to be removed from it).

iv. *Create witness worlds*: As there already exist possible worlds created as a result of processing diamond-formulas, the `EvoAlg` must skip this whole branch (it is only intended to ensure that, if there are no diamond-formulas available, the box-formulas will at least create their own worlds).

v. *Expand valuation function - Import existing atoms*: Once the box-formula has been evaluated, the `EvoAlg` must import the atomic formulas holding at the world of reference, as long as they do *not* appear in the clauses that were used to define the valuation function of the new worlds up to this point: notice that this means both the clauses of the diamond-formulas and the box-formulas used in each imaginary world. In this case, the clauses used in the DNF of the diamond-formulas and the box-formula have set the corresponding truth-value for atoms p and r at world z_1, and for p, q and r at world z_2. Therefore, the `EvoAlg` should look for the values of q, s and t for z_1, and s and t for z_2. This defines the following expansion of the valuation function:

$$V^+(q) = V(q) \cup \{z_1\}$$

$$V^+(s) = V(s) \cup \{z_1, z_2\}$$

$$V^+(t) = V(t) \cup \{z_1, z_2\}$$

(b) *Evaluate box-formulas*: There are no more scripts $\xi \in S^\alpha$, so the algorithm can exit this loop.

7. The `EvoAlg` has finished its execution. While evaluating the consequences of action α, the algorithm has created two new imaginary possible worlds by following two different scripts the agent believes in, and which describe different possible resulting states of affairs that *could* follow from action α; then, by evaluating a script about the consequences that would *necessarily* follow from action α, the agent has added more information to those possible states of affairs. Note how the way this algorithm handles the Default Evolution process is far more complex than the way it handles the Description process. This is because, as we have already discussed in Section 7.1.3, we want to consider all the consequences of an action as defined by the scripts believed by the agent, whereas we give her the chance to choose which factual rule to apply, in a step-by-step way, while elaborating on the static details of the imaginary world. Figure 7.8 shows the model after this execution of `EvoAlg`; note how, in the new imaginary worlds created in this case, we clamp both the formulas specified by the diamond-scripts and the box-script evaluated by the `EvoAlg`.

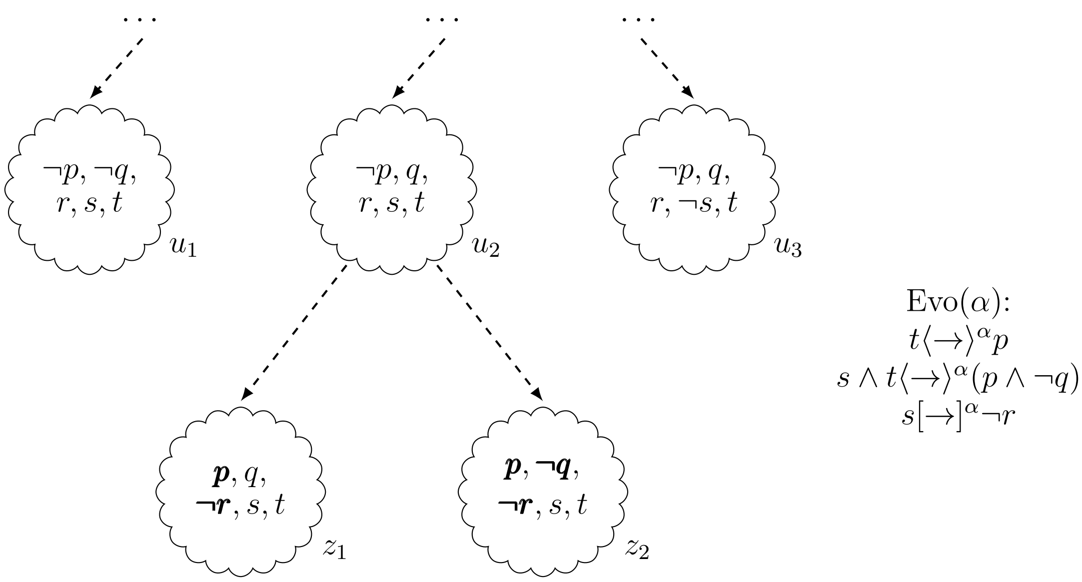


Figure 7.8: The agent imagines all the consequences of an action α in her imagining.

The `AddAlg`

At any time during an imagining the agent can decide to add a new premise into it, which would neither follow from her beliefs about factual rules, nor from her beliefs about scripts. This process, which is similar to the Initialization of a new imagining and the creation of new imaginary worlds, is handled by the `AddAlg` algorithm.

Similarly to what happens with the `InitAlg`, the `AddAlg` needs a world of reference w^R (which, in this case and conversely to what happened with the `InitAlg`, is required to be an imaginary world) and a formula $\delta \in \text{FORM}^*$. In this example, the agent takes world z_1 as the world upon which she wants to add a new premise, which she chooses to be $\neg p \vee r$. This executes a call to the algorithm in the following way:

$$\texttt{AddAlg}(z_1, \neg p \vee r)$$

The algorithm follows these steps:

1. *Check initial conditions*: World z_1 is an imaginary world as required, as it appears in $(u_2, z_1, \alpha) \in R_E$. Formula $\neg p \vee r$ is not contradictory (that is, $\neg p \vee r \nvdash \perp$), so the `AddAlg` continues its execution.

2. *Compute DNF*: Following what we did in the case of the `InitAlg`, the `AddAlg` must compute $\text{DNF}(\neg p \vee r)$ which, in this case, remains the same.

3. *Create imaginary worlds*: For each clause in $\text{DNF}(\neg p \vee r)$, the `AddAlg` must create one new imaginary world; in this case, the `AddAlg` creates 2 new possible worlds, x_1 and x_2, and defines the expanded set of possible worlds as follows:

$$W^+ = W \cup \{x_1, x_2\}$$

4. *Expand accessibility relation*: Once these worlds have been created, the `AddAlg` must make them accessible from the world of reference through the R_A relation, thus defining the expanded addition relation:

$$R_A^+ = R_A \cup \{(z_1, x_1, \neg p \vee r), (z_1, x_2, \neg p \vee r)\}$$

5. *Expand nominal structure*: In order to name the new imaginary possible worlds, the `AddAlg` must add two new nominals, which will be m_1 and m_2, and define the

expanded set of nominals and the expanded nominal function:

$$\mathrm{NOM}^+ = \mathrm{NOM} \cup \{m_1, m_2\}$$

$$N^+ = N \cup \{(m_1, x_1), (m_2, x_2)\}$$

6. *Expand valuation function*: The `AddAlg` must define the valuation function for the newly created imaginary worlds in two steps, as usual:

   (a) *Clamp new atoms*: Firstly, it must set the truth-value of the atomic formulas of each clause in DNF$(\neg p \vee r)$, and for each corresponding possible world. This defines the first step of the expanded valuation function as follows:

   $$V_1^+(r) = V(r) \cup \{x_2\}$$

   (b) *Import existing atoms*: Secondly, the `AddAlg` must import the atomic formulas that hold at the world of reference z_1, but which do not appear in the corresponding clause used to create the new imaginary world. This defines the second and last step of the expanded valuation function in the following way:

   $$V^+(p) = V_1^+(p) \cup \{x_2\}$$

   $$V^+(q) = V_1^+(q) \cup \{x_1, x_2\}$$

   $$V^+(s) = V_1^+(s) \cup \{x_1, x_2\}$$

   $$V^+(t) = V_1^+(t) \cup \{x_1, x_2\}$$

7. The `AddAlg` has finished its execution. The part of the model which is relevant for this execution can be seen in Figure 7.9. As usual, we include the negated atomic propositions in the figure, and we also highlight those atomic formulas clamped during the creation of the new imaginary possible worlds.

The Whole Imagination Act

In this example we show how each one of the algorithms defined in the Logic for Imagination Acts work. We have done it step by step and provided the relevant fragment of the model at each step, showing how each algorithm expands an already existing model.

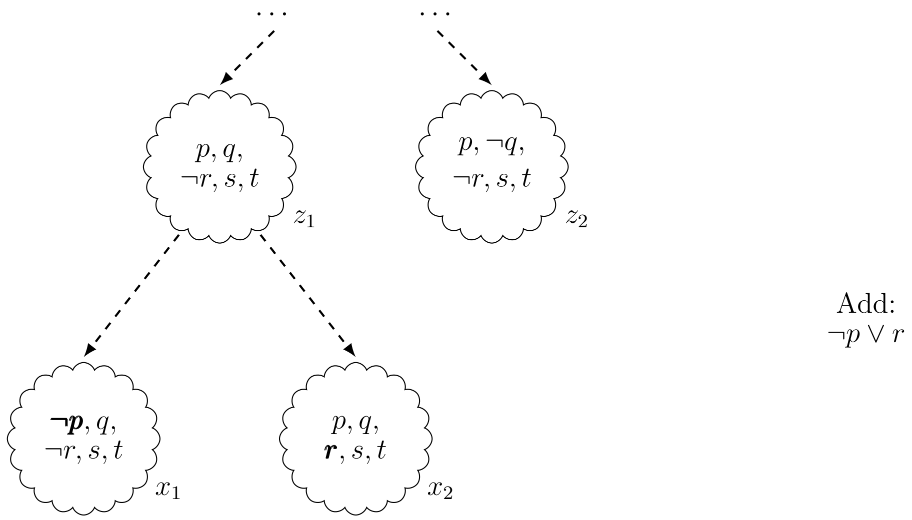


Figure 7.9: The agent adds a new premise into the imagining which would not follow by using factual rules or scripts.

The final model at the end of the example is shown in Figure 7.10. For the sake of readability, we also write the negated atomic formulas at each possible world, and the accessibility relations in the figure are *not* labeled; besides, we include, at the right hand of the figure, the name of the process corresponding to each step of the example, together with the formulas that were used in them. Furthermore, we highlight in bold font the atomic formulas that have been clamped during the execution of the corresponding algorithm; recall that those formulas were given priority while defining the new imaginary worlds, and thus the remaining of the atomic formulas holding in there (that is, those formulas not written in bold font) were imported from the world of reference in each case.

7.7 Plugging in the Rhombus of Imagination Acts

At this point, we have just shown how our Logic of Imagination Acts captures the four processes identified in the Common Frame for Imagination Acts, which we introduced in Chapter 6. At this point, therefore, it is only natural to wonder where the Rhombus of Imagination fits within the current picture. If the Rhombus was specially conceived as an analysis tool for the Common Frame, could we use it for the Logic of Imagination Acts as well?

Indeed, we can, and in order to link the Rhombus of Imagination with the Logic of Imagination Acts, we have to focus on its models. A Model for Imagination Acts $\mathcal{M}$

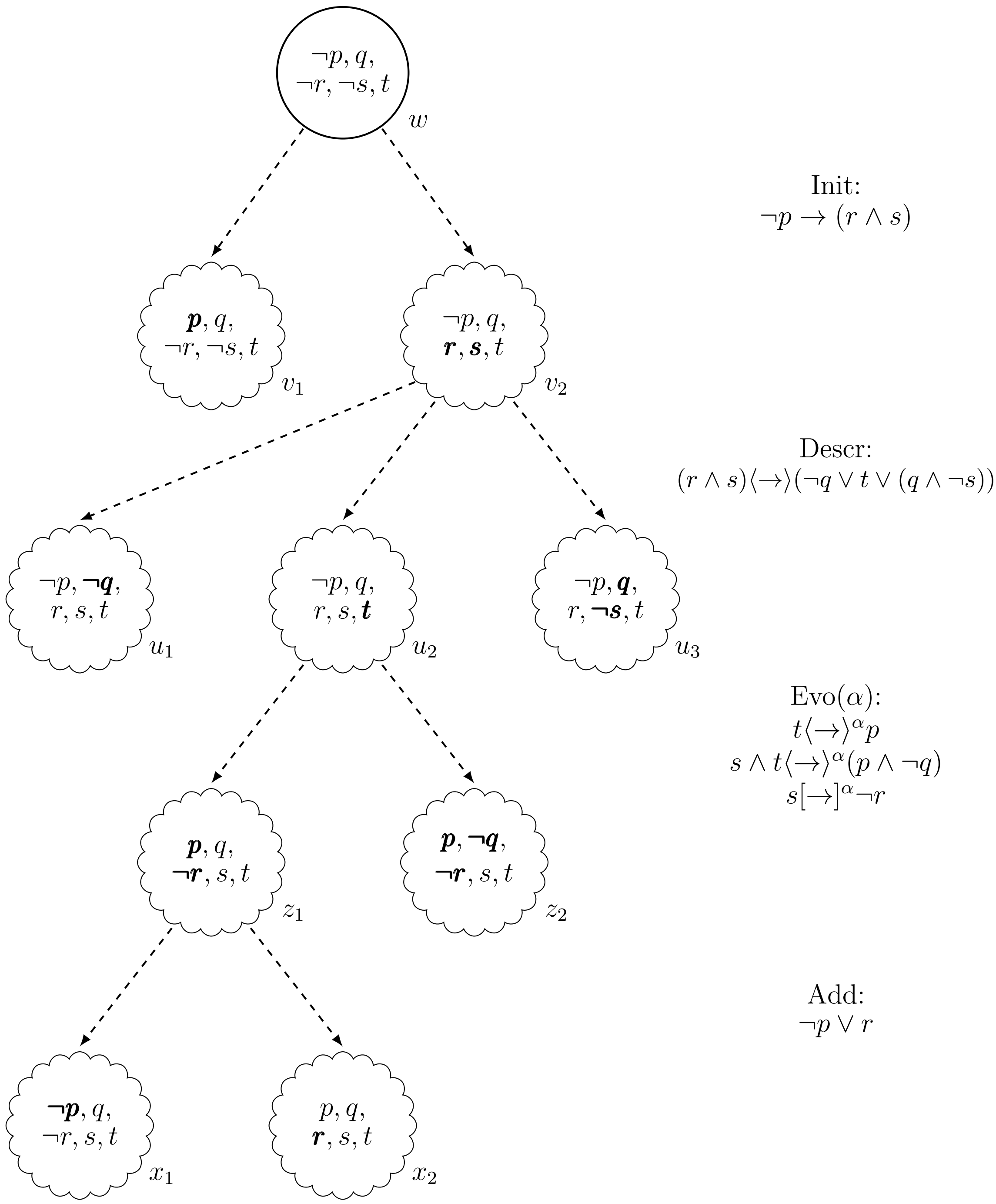


Figure 7.10: The model for the example, with the execution of each algorithm.

contains four different accessibility relations R_I, R_D, R_E and R_A, each one standing for one of the four processes of the Common Frame, and each one being expanded through executions of the four algorithms. In particular, a single execution of one of the algorithms will only create one further "step" or "level" of imagination, as we do not allow the agent to imagine "nested" imagination acts (that is: formulas containing a dynamic imagination formula inside them). Being the representation of a particular act of imagination, a model $\mathcal{M}$ can be used for computing its quantitative blue-print by using the quantitative version of the Rhombus of Imagination (see Section 6.2.2). Taking into account that the accessibility relations capture the executions of each one of the processes that the Rhombus can measure, we need to take into account the elements in them to compute the Rhombus. Before doing so, nevertheless, there are three important considerations we have to make beforehand:

1. A Model for Imagination Acts can, in fact, represent different imagination acts. The agent may initiate a certain imagination act, elaborate on it, and then toss it away and begin anew with another imagination act that has nothing to do with the first one. Even though both imagination acts will be part of the same model (and will be represented, specifically, by the accessibility relations of the model), if we want to create the blue-print of a single imagination act, we cannot simply take *every* element in *every* accessibility relation, as we would be considering more than one distinct imagination acts together.

2. Even if we only take the elements in the accessibility relations concerning one particular imagination act, we should not count them all without any kind of filter. An execution of one of the algorithms corresponds to a single call of one of the processes of the Common Frame, but nevertheless it might create more than one new imaginary possible world, and thus more than one new accessibility relation. The Rhombus of Imagination does not count how many alternative imaginary worlds are created by the different executions of each process, but it rather counts how many executions of each process there have been, regardless of how many alternative imaginary worlds they can represent.

3. It is possible that, during an imagination act, the agent explores or elaborates on different stories or chains within it, and then backtracks to elaborate on a different one. For instance, if I imagine that I am in a fancy restaurant and I keep fantasizing about being the diner, but suddenly I get bored and I decide that it would have been more fun to be the waiter, so I backtrack to the very beginning of my imagining and I begin my fantasy anew by following a different, independent course of action. Should both independent paths be considered as being part of the same imagining,

or should we distinguish between both? Our stance is that, as the two imaginary stories are, in fact, different and independent, we should evaluate them separately, and so should the Rhombus. When computing the blue-print of an imagining, therefore, we want to compute the blue-print of a single imaginary story, rather than of a whole bunch of independent stories that were created by using the same initial premise.

Each one of these previous considerations, thus, must be taken into account and "translated" into the formalities of the models of our logic to see how they can be fed into the Rhombus of Imagination (we refer to the Section 7.3.2 for the definition of some specific terms we use in here):

1. We must restrict the elements of the accessibility relation that we are going to consider by using a single root; that is, a single pair of world of reference and initial premise (w^R, δ). An act of imagination is initiated by a *single* execution of the `InitAlg`, which has taken a single possible real world w^R as a world of reference, and which has used a single initial premise δ. If world w^R has been used as the world of reference for some other execution of `InitAlg`, but using a different initial premise γ, then it belongs to a different imagination act, and we should not take it into account (and similarly for a different world of reference using the same initial premise). Therefore, we must limit the counting of the elements appearing in the accessibility relations by considering only those elements for which there exists a chain leading to one of the elements of the form $(w^R, x, \delta) \in R_I$. Another way of putting it would be that, if there exists an imaginary possible world z from which we cannot backtrack, by following the accessibility relations, until we reach an element of the form $(w^R, z, \delta) \in R_I$, then it is because world z belongs to a different act of imagination, and thus should not be taken into account when computing the current Rhombus of Imagination.

2. We should not count the horizontal worlds in the elements of an accessibility relation, but rather just the vertical ones. Our interest does not lie in how many different imaginary worlds an execution of (say) the `DescrAlg` has created, but just in that the `DescrAlg` has been executed and that it has created *some* world (no matter if just one, or a thousand different worlds). In order to account for this, we should only count look for elements in the accessibility relations where the world of origin w and the formula # signing the element is different, regardless of the world

of destination x that appears in them. For instance, if we find a set of elements $\{(w, x_1, \zeta), (w, x_2, \zeta), \ldots, (w, x_n, \zeta)\} \subseteq R_D$, they should just be counted as *one* execution of the `DescrAlg`, even if such execution has generated n different imaginary possible worlds (and similarly for any other accessibility relation and algorithm).

3. We must ensure not only that the root of the imaginative episode is a single pair of world of reference and initial premise, and that the elements of the accessibility relations are only considered in their vertical "deepness", but we also have to consider whether a particular execution of an algorithm belongs to the same imaginary story or chain. In order to do so, the easiest way is to determine which executions we have to take into account and which not by following a sort of backwards induction method: when computing the blue-print of a certain imaginary story, we have to locate one of the leaf-worlds of such imaginary story, and then backtrack all the way by following the backwards-accessible worlds until we reach a real world w. This way, we make sure that each and every execution involved in this backtracking process will be part of the same imaginary story.

By translating our initial considerations into our formal setting, we have already defined how computing the Rhombus of Imagination for a particular imaginary story works, using our Logic of Imagination Acts (and, in particular, one of its models) as the input.

7.7.1 An Algorithm for the Rhombus of Imagination

In this section, we present an algorithm that can be used to automatically compute the Rhombus of Imagination of an imaginary story, given a Model for Imagination Acts $\mathcal{M}$. We call it the `RhombusAlg` algorithm and, although it should not be considered as being part of our formal logic, it is an additional layer that can be applied upon it. We define it as follows:

1. The `RhombusAlg` requires a Model for Imagination Acts $\mathcal{M}$ and a leaf-world v, which is required to be an imaginary possible world, as arguments. If $\mathcal{M}$ is not a model of the required kind, or if v is either not an imaginary possible world or a leaf-world, the algorithm finishes its execution. Otherwise, we define our *current world* as v, and the algorithm proceeds.

2. In order to compute the blue-print of that particular imaginary story, the `RhombusAlg` must keep track of the number of executions of each algorithm by using the following counters (the value of which is initialized as being 0):

- $c_I := 0$
- $c_D := 0$
- $c_E := 0$
- $c_A := 0$

3. *Loop backtracking relations:* The `RhombusAlg` must browse every accessbility relation R_I, R_D, R_E, R_A and look for an element of the form $(u, v, \#)$ (where v is the current world being considered, and u any other possible world in W). If such an element is found, the algorithm must check:

   (a) If $(u, v, \#) \in R_I$, then $c_I := c_I + 1$. As $(u, v, \#) \in R_I$, it means that this step corresponds to the Initialization of the imagination act, and so it means that world u is a real world: the `RhombusAlg` must now get out of the current loop by jumping to Step 4.

   (b) If $(u, v, \#) \in R_D$, then $c_D := c_D + 1$.

   (c) If $(u, v, \#) \in R_E$, then $c_E := c_E + 1$.

   (d) If $(u, v, \#) \in R_A$, then $c_A := c_A + 1$.

   (e) In any case of the previous, set $v := u$ and go back to Step 3 to continue with the backtracking loop; in other words, we set the new world u we just found as the current world, and the loop begins anew.

4. *Compute the blue-print*: The `RhombusAlg` reaches this step once it has found to real world used to initiate the imaginary story[18]. As the counters have been increased while backtracking from the initial leaf-world v towards the beginning of the imaginary story, we know how many executions of each imagination algorithm have been in that particular imaginary story, and so we can directly apply the formula for computing the weight for each vertex of the Rhombus of the v-Imaginary Story (again, see Section 6.2.2 for details and an informal example on how we compute the Rhombus of a particular imagination act).

   We use c_i to refer to any of the counters used in the `RhombusAlg`, and we take $i \in \{I, D, E, A\}$. Furthermore, we use $c_\uparrow$ to refer to the particular counter c_i which has the highest value (that is: the counter referring to the process which has been

[18]Note how, as we require $\mathcal{M}$ to be a Model for Imagination Acts, we know that, by the way such models work, there will always be an element belonging to R_I at the beginning of each imaginary story, so this ensures that the loop within the `RhombusAlg` will always finish by that condition.

executed the most in the imaginary story being analyzed); we use it to compute the proportion constant k (which is used to scale the resulting figure accordingly). We compute each value and weight in the following way (for $\{i, j\} \subset \{I, D, E, A\}$):

$$\text{Set counter with higher value as: } c_{\uparrow} = c_i \text{ such that } c_i \geq c_j$$

$$\text{Compute proportion constant: } k = \frac{1}{c_{\uparrow}}$$

$$\text{Compute each weight as: } w(c_i) = \frac{c_i}{c_I + c_D + c_E + c_A} \cdot k$$

5. Once all the weights for each algorithm (that is, for each vertex) have been calculated, the four values $w(c_i)$ for each c_i correspond to the values that can be used to draw the Rhombus of Imagination for the imaginary story represented in model $\mathcal{M}$ that ends up at the leaf-world v.

As we can see, adding the Rhombus of Imagination as a support tool for the Logic of Imagination Acts can be easily done. As the logic captures the processes recognized by the Common Frame for Imagination Acts, the Rhombus is a natural extension for the logic; furthermore, the procedure required to compute it can be expressed by using a simple algorithm as well.

7.8 Thoughts on the Logic of Imagination Acts

Last, but not least, we devote the last section of this chapter to consider and discuss the features of the Logic of Imagination Acts. As we argue in the following paragraphs, this new approach to a dynamic logic of imagination has many positive features, but it also has certain shortcomings that could be improved in a further refinement of it.

7.8.1 The Good

Probably the most striking feature of the Logic of Imagination Acts is its dynamic modularity, which amounts to dynamic versatility. An act of imagination is no longer seen, as it was in the Logic of Imaginary Scenarios introduced in Chapter 4, as a *single* execution of a *single* algorithm: in this setting, an act of imagination is seen as a sequence of possibly many executions of different processes, each of them being part of the act of imagination as a whole, but each one of them distinct enough to account for a particular way in which the imagining can be elaborated. In this setting, we can see beyond an act of imagination and we can identify and split it into the different processes that take part in it. The depth of analysis and, as a result, the level of understanding that this approach gives us when studying the formation and elaboration of imaginary worlds is much more higher than the one we would have, if imagination acts were handled by a single black-box algorithm taking care of the whole process —even if that algorithm could account for all the processes within itself. Furthermore, this feature fulfills one of the initial goal we had while defining this logic: to create a formal system capable of capturing the Common Frame for Imagination Acts.

The next positive feature of our logic follows directly from the previous considerations: being able to develop imagination acts in a modular, non-brute-forced way, allows us to represent acts of imagination that can create and develop imaginary worlds in a wide variety of different ways. We no longer aim to creating a sort of exhaustive algorithm detailing each and every possible alternative in an act of imagination, as if it was a chess game being analyzed by a computer in order to check each and every possible move. Instead, we now aim to define a system that captures the way human being elaborate on their imagining in a sort of step-by-step way, partially exploring one or another possibility, and leaving many options aside. This kind of selective, partial development of imagination acts allows our system to represent how, when using, for instance, two different factual rules, the same imaginary scenario could be developed in radically different ways.

Inevitably, when we mention the fact that imaginary scenarios could develop in different ways, a question pops up into our minds: why is that? Because involuntary imagination is, in fact, *partially* involuntary: the consequences that follow from the factual rules and the scripts are indeed implied, but the *choosing* of which factual rules and scripts to use, or on which one of the possible alternatives keep elaborating, involves a certain degree of agentiveness, as we argue in Section 6.1.6. When a certain imaginary scenario states that there is, for instance, a huge tree and a small cabin by it, the agent can choose whether she wants to elaborate on the tree (how does it look like? What kind of tree is it? How huge is it?), or on the cabin (how small is it? Is it made of wood? Is there anyone inside?). Similarly, when in an imaginary scenario about being in a tea party the agent comes up with two possible alternative situations following from it (say, pouring tea into a cup, or cutting a cake), it is up to the agent to choose which she actually wants to follow in her imagining. In order to account for this kind of agentiveness that we find "hidden" within processes that, otherwise, develop the imagining in an involuntary mode, we need to be able to identify first where they take place, within those processes. The

Logic of Imagination Acts does so by allowing the agent (or the modeler, as she is the one accounting for the agent's decisions within the model) to select which premises, factual rules and scripts she wants to use in order to develop the imagining; later on, the logic also locates this hidden agentiveness (and retrieves its control to the modeler) when having to decide which one, among all the possible alternative imaginary worlds, is selected as the input of the next imagination algorithm to apply. Therefore, the modular way in which this logic handles imagination acts allows, at the same time, to identify these checkpoints of agentiveness within the acts, and to give the control back to the modeler in such moments.

There is still one important feature of our logic that we claim represents a great advancement, specially with respect to our previous approach to a formal system; namely, the Logic of Imagination Acts can now account for the notion of "reality-oriented rules", which already appeared in the theories of imagination we reviewed in Chapter 2, and which is critical when capturing in which way imaginary worlds are reality-like. This is a feature that our first formal approach, the Logic of Imaginary Scenarios, could not account for, and which we already identified, in Chapter 5, as one of the requirements we wanted to account for in a new formal setting. Through the sets of factual rules and scripts, our logic captures this notion and reproduces how human beings determine what would typically be the case in an imaginary world, *if it was real*. With this, we can account for both those premises that are voluntarily clamped into an imagining (either via the Initialization or the Unscripted Addition process), but also for those developments that follow rules and scripts based on how things work in the real world, and which are believed by the agent and then used to elaborate on the imaginings.

Regarding the positive features of our approach, therefore, we claim that the Logic of Imagination Acts represents a valuable contribution in the analysis and understanding of how imaginary worlds are created and developed, and also that this logic is more suited than the previous Logic of Imaginary Scenarios to be used for such goal. Moreover, the features discussed in the previous paragraphs account for those requirements we identified while evaluating the Logic of Imaginary Scenarios, both by the end of Chapter 4 and in Chapter 5.

7.8.2 The Bad (and the Ugly)

The first shortcoming the reader may notice in the Logic of Imagination Acts with respect to the Logic of Imaginary Scenarios concerns the depth of their formal approach. Whereas in the former logic we put more attention into its syntax and certain validties that could be expressed with formulas, we do not do such thing in this latter approach. In the Logic of Imagination Acts, we decided to build our logic from scratch, precisely because we wanted to focus on the algorithmic account of imagination, while avoiding technical constraints that could be derived from "external" systems; nevertheless, the cost of doing so is reflected in a logic that is richer at a procedural level, but more limited in its language. In the Logic of Imagination Acts, we get a system with a much more detailed procedural account of the dynamics involved in imagination acts, but at the cost of devoting less attention to the study of the formal properties of the system.

Still related with the formal setting of the Logic of Imagination Acts, we believe that having an explicit ways of considering beliefs would be a very interesting addition. Even though we have already argued, in Section 7.3.1, how we interpret beliefs in this logic, the lack of an explicit operator accounting for the "believe" attitude limits the interaction our logic can have with this other mental attitude. The current version of our system provides great insights regarding how imaginary worlds are created and develop; if we expanded our logic in order to include explicit representation for beliefs, we could then explore how those imaginary worlds would potentially affect the agent's beliefs. For instance, if we had a way of explicitly representing the agent's beliefs in the real worlds, we could study how one further algorithm, corresponding to the "realization" of an agent through performing a thought experiment, could then alter her beliefs about the real world. In such case, in would also be interesting to compare our approach to other systems able to deal with belief revision or internal reasoning steps, such as some applications of Dynamic Epistemic Logic (see Chapter 6 of [50] for an overview) and other works devoted to represents the dynamics of awareness and realization (like in [51]).

Regarding the way we have defined our algorithms to work, there have been certain decisions we had to make that could have been taken differently. Even though we argue that our current approach shows good results, we could also consider how those decisions affected the overall system, and how changing them would alter the system's behavior.

While introducing the `EvoAlg`, we devote Section 7.4.3.1 to discuss how the algorithm would handle contradictory sets of scripts. Whereas, in the current approach, inconsistent states could never appear as a result of an execution of the `EvoAlg`, we also argued in there how an alternative approach that recognized and signaled inconsistent sets of believed scripts could also be interesting. In order to do so, we would need to alter our `EvoAlg` to keep track of all the formulas that are being clamped in every new imaginary world; then, if at any step during the algorithm, and as a result of evaluating a certain script, the algorithm needed to clamp a formula contradicting another formula already clamped into the same world, then the algorithm would have found an inconsistency. As we pointed out earlier, nevertheless, our current focus is not to have our algorithm "solve" the consistency problems that an agent may have in her beliefs about scripts. If the model is defined in such a way that certain scripts contradict each other, then so be it: our system should not "repair" it. However, it could indeed identify the contradiction and, instead of prioritizing certain formulas over the other (which is still a valid solution, and the one we have taken), rollback the whole execution of the `EvoAlg` and signal that, while trying to evaluate a certain action α, there are contradictory beliefs. How could we do that?

In order to do so, we could use a kind of "special" behavior for the accessibility relation. In the Logic of Imagination Acts, it is *not* possible that an execution of any algorithm points out to an already existing world; even if the new imaginary world that is being defined is, formula-wise, completely equivalent to an already existing one, the algorithm will still create a new world. This also implies that it is not possible, for any execution of any algorithm, to create a reflexive relation going from the world of reference back to it. Therefore, we could use this "restricted behavior" that, following their corresponding algorithms, all accessibility relations have, and use it in our own benefit. In this case, let's focus on relation R_E, which is the one that concerns the `EvoAlg`. We have already explained how, by keeping track of every formula that the `EvoAlg` clamps into an imaginary world while looping through the queue of scripts, it is possible to identify contradictory scripts. This realization, nevertheless, is sort of "black-boxed" to the modeler: it happens within the execution of the algorithm and, unless the modeler is the one actually following the algorithm, it is not explicitly expressed outside the algorithm itself. The solution to signal this kind of situations, then, would be to use a reflexive R_E relation. Whenever the `EvoAlg` is executed with a world of reference w^R and an action α, and finds a contradiction among the set of scripts that it has to use, the algorithm should rollback the whole process (that is: remove every new imaginary world, nominal, accessibility relation and valuation created in this execution) and draw a single reflexive accessibility relation going from the world of reference to itself, and signed with the corresponding action: (w^R, w^R, α). This would semantically account for the fact that, while trying to imagine the consequences of α in world w^R, the agent stumbles upon a contradictory set of scripts (which, obviously, would be some of the formulas in SCRIPT whose modal operator is signed with action α), which prevents her from actually being able to conceive the world that would result from it[19]. Our system would not be

[19]Recall that, as we already mentioned before during this work, we do not want to represent an agent

preventing the contradiction to appear, as it actually does now, nor it will be repairing the set of contradictory scripts, but it would, nonetheless, express explicitly within the model that the agent indeed has contradictory beliefs, regarding the consequences of action α.

7.8.3 Separate Ways

Aside from interesting features and shortcomings, the Logic of Imagination Acts and its algorithms left us with a pool of interesting ideas for future expansions. In the following paragraphs, we briefly introduce them.

One of the features that would be more interesting to add to our logic would be the mechanisms needed to allow imagination processes to affect not only the states of affairs represented by imaginary worlds, but also the sets of rules and scripts holding in there. Currently, each algorithm creates new imaginary worlds by clamping certain atomic formulas in them, as specified by premises or rules; imaginary possible worlds can turn out to be indeed different to the real possible worlds, but they are different in terms of their states of affairs, or what is the case in there. The factual rules and the scripts believed by the agent, nevertheless, are constant throughout the model. However, imagination should be able to alter that as well. It is true that I can imagine that *things* are different, but I can also imagine that the *rules* governing the world are also different. For instance, I could imagine that I have this weird condition in which, whenever I sneeze, I shoot a lighting bolt from my mouth. This would amount to imagining not that the state of affairs of the imaginary world is different from the actual real world (at least not yet), but rather that the scripts governing it are distinct. If, after creating an imaginary world in which such rule holds, I imagine that I sneeze, then the resulting scenario will probably be dramatically different then the one depicted without adding that imaginary script.

In order to account for that, we would need to add still one further layer to our formal models and associate, to each possible world, a particular set of factual rules and scripts believed in there. This way, the real possible worlds would account for the factual rules and scripts that the agent actually believes about the real world, but we could also create an imaginary world in which such rules and scripts were different. After all, and despite the fact that most laws of physics, for instance, are held constant in many imaginary worlds and fictions, there are also a huge amount of factual rules and scripts that change from one imagining to another.

who can actually imagine inconsistent states; nevertheless, other authors like Berto (see 3.2.5) allow that.

Due to the nature of how imaginary worlds are gradually developed, it would also be an interesting contribution to see how our algorithms would accommodate in a *paracomplete* setting[20]. In a nutshell, a paracomplete setting would allow our possible worlds to have certain truth-value gaps, with respect to certain atomic formulas; this, in turn, would allow us to determine, for each possible world and atomic formula, whether the formula is true in there, false in there, or simply undetermined or unspecified. When using this setting, we could avoid having to import *every* possible atomic formula from the world of reference, and thus we could just allow the new imaginary worlds to be developed in a truly step-by-step way, filling up only those details that are specified by the corresponding premise, factual rule or script. We think that it would be interesting to see how our layer of imagination algorithms could be adapted into such setting, and thus it is an interesting topic for future work.

Furthermore, and as a way of complementing the paracomplete setting, we could also consider adding *aboutness*[21] into the system to import only those atomic formulas that were related somehow to the details being clamped into the new imaginary worlds. In short, aboutness introduces a way of establishing a connection between different atomic formulas whose interpretation is related. For instance, we could say that, when imagining myself having breakfast on top of the Empire State Building, I am imagining something about the city of New York, but not about Barcelona, nor about the Moon, for instance. Therefore, adding an aboutness filter determining what else is related with my imagining could allow to filter which facts should the algorithms import, and which not.

[20] Some of the logical system defined by Berto allow for both paracomplete and paraconsistent worlds; see Section 3.2.5 for further details.

[21] Also appearing in works from Berto, in Section 3.2.5.

Chapter 8

Applications

Throughout the previous parts of this work we have been reviewing theories about imagination, we have proposed our own theory for the dynamics involved in creating and developing imaginary worlds, and we have defined two different formal systems able to capture, through the definition of detailed algorithms, how those dynamics work. We want to devote the last part of this work to the applications that we derive, and that could be derived, from our contributions.

Firstly, we present a prototype of a computer program designed to implement the algorithms defined by the Logic of Imagination Acts, and which provides a way of testing, through computer simulations, how our formal system reproduces the way human beings imagine. Then, we discuss and consider the applications that our contributions could have in a field where the creation and the development of imaginary worlds has a critical importance, and thus where a better understanding of the dynamics involved in it can be a valuable contribution: video games.

8.1 SILOGIA: The Simulator for the Logic of Imagination Acts

In the Logic of Imagination Acts, introduced in Chapter 7, we define a formal system in which the dynamic processes involved in creating and developing imaginary worlds, as identified by the Common Frame for Imagination Acts presented in Chapter 6, can be captured by using four distinct algorithms. In this logic, those algorithms can be executed to compute a new, expanded version of the formal model in which new imaginary worlds are created and characterized. Now, we ask ourselves: taking into account that the formal models of our logic can be expressed mathematically, and considering that the four algorithms we define can be expressed by using a modern programming language, would not be possible to capture the dynamics represented by the logic into a computer program?

Yes, indeed: we can. From this answer is born a prototype of the *Simulator for the Logic of Imagination Acts*, or SILOGIA, which is a computer implementation of the algorithms defined by the Logic of Imagination Acts. In this computer program, the algorithms can be tested via computer simulations into different formal models in order to see how they are expanded when the agent executes a certain process of imagination. The SILOGIA, therefore, represents both a way of running tests to asses how the Logic of Imagination Acts works, and also a first approach to a computer simulation of how human beings create and develop imaginary worlds, based on the insights derived from the Common Frame for Imagination Acts.

We follow the usual Waterfall Model methodology for software engineering and, in the following sections, we discuss the analysis, the design and the technologies used to implement this prototype of the SILOGIA. Then, we provide a link to a running version, we follow, step-by-step, the same example introduced in Section 7.6 while defining the Logic for Imagination Acts, and we foresee how the Rhombus of Imagination could easily be introduced in the computer implementation as a support tool for the simulations.

8.1.1 Analysis

We want to implement a computer program capturing the way the algorithms defined by the Logic of Imagination Acts works. The main objective beyond this implementation is two folded:

1. To show that the four algorithms corresponding to the Initialization, the Description, the Default Evolution and the Unscripted Addition processes can indeed be implemented using a modern programming language.

2. To provide a computer-assisted and automated tool for testing the four algorithms upon different instances of Models for Imagination Acts, and also to see how they expand such models.

Requirements

The requirements for the SILOGIA are the following:

1. The user must be able to either load an already existing Model for Imagination Acts, or to create a new one.

    - If the user wants to load an already existing model, she have to select, among all the available models, which one she wants to load.
    - If the user wants to create a new model, she must specify the following elements of the initial model:
        - A set of real possible worlds W.
        - A set of nominals NOM.
        - An exhaustive function N establishing the relation between each nominal and a real possible world.
        - A set of atomic propositions ATOM.
        - A set of actions ACT.
        - A set of factual rules FACT.
        - A set of scripts SCRIPT.

      These elements must conform to the requirements specified by the formal definition of the Models for Imagination Acts, in Section 7.3. Nevertheless, and as in this case we are working over a specific, finite version of the model, the sets of possible worlds, nominals and such have a finite number of elements.

      Furthermore, the form of the formulas in ATOM, ACT, FACT or SCRIPT must conform to the requirements specified in Section 7.2, detailing the syntax of the language. Additionally, in this version of the prototype for the SILOGIA, we impose the following constraint:

        - Both the antecedent and the consequent of every formula in either FACT or SCRIPT *must* be expressed in Disjunctive Normal Form (DNF). That is, for any formula ζ belonging to any of the previous two sets, and whose form corresponds to either $\varphi\langle\rightarrow\rangle\psi$, $\varphi\langle\rightarrow\rangle^{\alpha}\psi$ or $\varphi[\rightarrow]^{\alpha}\psi$, we require that both φ and ψ are expressed in DNF.

2. Once the user has either created a new model or loaded an already existing one, she can execute any of the four algorithms corresponding to the processes related with imagination acts, as defined in the Logic of Imagination Acts. The model upon which the user executes the algorithms will be called the Working Model (as specified in Section 8.1.2), and it is where the expansions that result from the executions of

the algorithms are registered, while the user interacts with the prototype. Recall that, as it happens in our formal system, the user can execute the algorithms in the order she wishes to:

- In order to execute the `InitAlg`, the user must provide the following arguments:
  - A real possible world of reference w^R, selected among the already existing worlds in the Working Model.
  - An initial premise δ, which, in this version of the prototype, must be expressed in DNF.
- To execute the `DescrAlg`, the user must provide the following arguments:
  - An imaginary possible world of reference w^R, selected among the already existing worlds in the Working Model.
  - A factual rule ζ, selected among the formulas belonging to the set FACT.
- In order to execute the `EvoAlg`, the user must provide the following arguments:
  - An imaginary possible world of reference w^R, selected among the already existing worlds in the Working Model.
  - An action α, selected among the actions belonging to the set ACT.
- In order to execute the `AddAlg`, the user must provide the following arguments:
  - An imaginary possible world of reference w^R, selected among the already existing worlds in the Working Model.
  - A new premise δ, which, in this version of the prototype, must be expressed in DNF.

3. The execution of any of these algorithms must expand the Working Model accordingly and draw the expanded version of it. Moreover, the user can execute each algorithm as many times as she wants to, as well as vary their parameters as much as she likes.

4. At any point, the user may save the changes made in the Working Model, with respect to the model she initially created or loaded. In particular, the new set of possible worlds, nominals, valuations and accessibility relations created through a new execution of any algorithm may be saved at any time.

Due to the fact that our current implementation corresponds to a prototype of the SILOGIA, we omit certain features that, although practical, are not directly related to our contributions. In particular, we have not implemented an automated tool for translating a propositional formula into its DNF; this is why we require that all the formulas usable by our program must be given in DNF. We argue that, although adding such feature would enhance the prototype's easiness of use, computing the DNF of a formula is not one of the contributions of our work, and so we have prioritized other features of this prototype.

It is worth noting that the SILOGIA does *not* provide a way of evaluating formulas in the model; that is, the user cannot ask the prototype whether a certain formula holds at a certain world. In this implementation, we are interested in capturing the dynamic processes involved in creating and developing imaginary worlds, and the Logic of Imagination Acts, its algorithms and its formal models provide the perfect starting point. Besides, implementing an automated tool for evaluating the truth-value of any formula of our logic would deviate from the current goals of this prototype, and so we leave it for a future version of the program.

8.1.2 Design

The design of the SILOGIA is briefly explained in this section. Figure 8.1 represents the way our prototype works; the modules involved in the functional architecture are briefly described in the following list:

- *User Interface:* It allows the user to load already existing models or to create new ones, and provides the required interface for calling any of the four algorithms. Additionally, the Working Model is graphically represented in it.

- *Working Model*: It is an "alive" copy of the initial model (be it a new model created by the user, or an already existing one loaded from the User Interface) and stores any change derived from the execution of an algorithm. The Working Model acts as a submodule between the user interface, the algorithms and the database, and prevents that changes made by the user are automatically stored permanently (as the user may be interested in performing different tests over the same initial model, without making any permanent changes into it). Moreover, it is also used as a data source for the algorithms.

- *Algorithms*: There are four different algorithms encoding the behavior of each corresponding process within imagination acts, as defined in the Logic of Imagination Acts. These algorithms are called through the user interface after receiving the required parameters, they get any required data from the Working Model, and then they expand it accordingly.

- *Database*: Stores different instances of Models for Imagination Acts. It is accessed either when loading an already existing model, when creating a new one, or when saving the changes made in the Working Model.

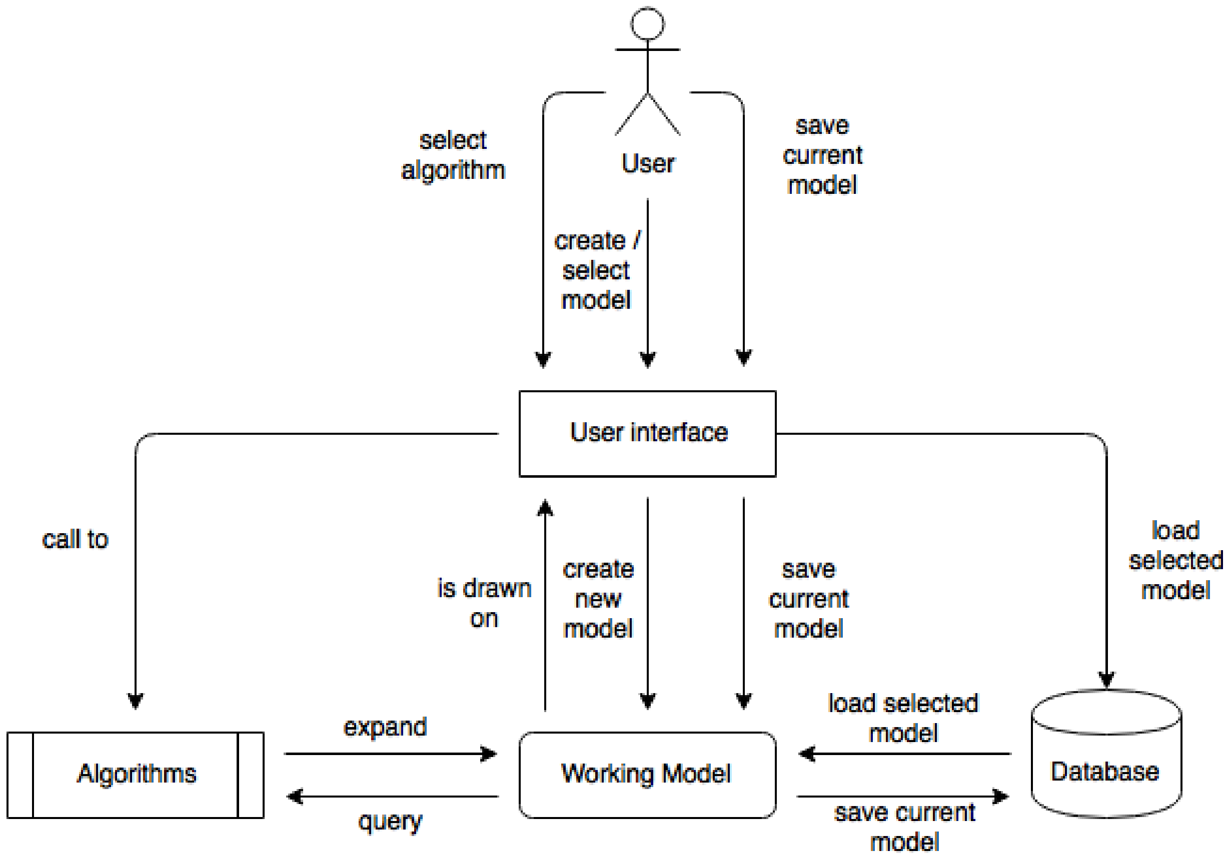


Figure 8.1: The functional architecture of the SILOGIA.

The database contains all the elements required by the formal definition of the Models for Imagination Acts, as in Section 7.3. Figure 8.2 depicts the conceptual schema of the database in an Entity Relationship diagram. In order to allow the SILOGIA to store instances of distinct models, we include a new entity, called "Model", which allows to determine to which specific model belongs any other element. Although this entity must be related to every other one, we do not draw such relationships in the conceptual schema for the sake of clarity. Notice, furthermore, how we define another entity "Script" that is not related to any other ones (besides "Model"). As it happens with the Logic of Imagination Acts and the Models for Imagination Acts, the entity "Script" is only used internally by the Evolution algorithm `EvoAlg`.

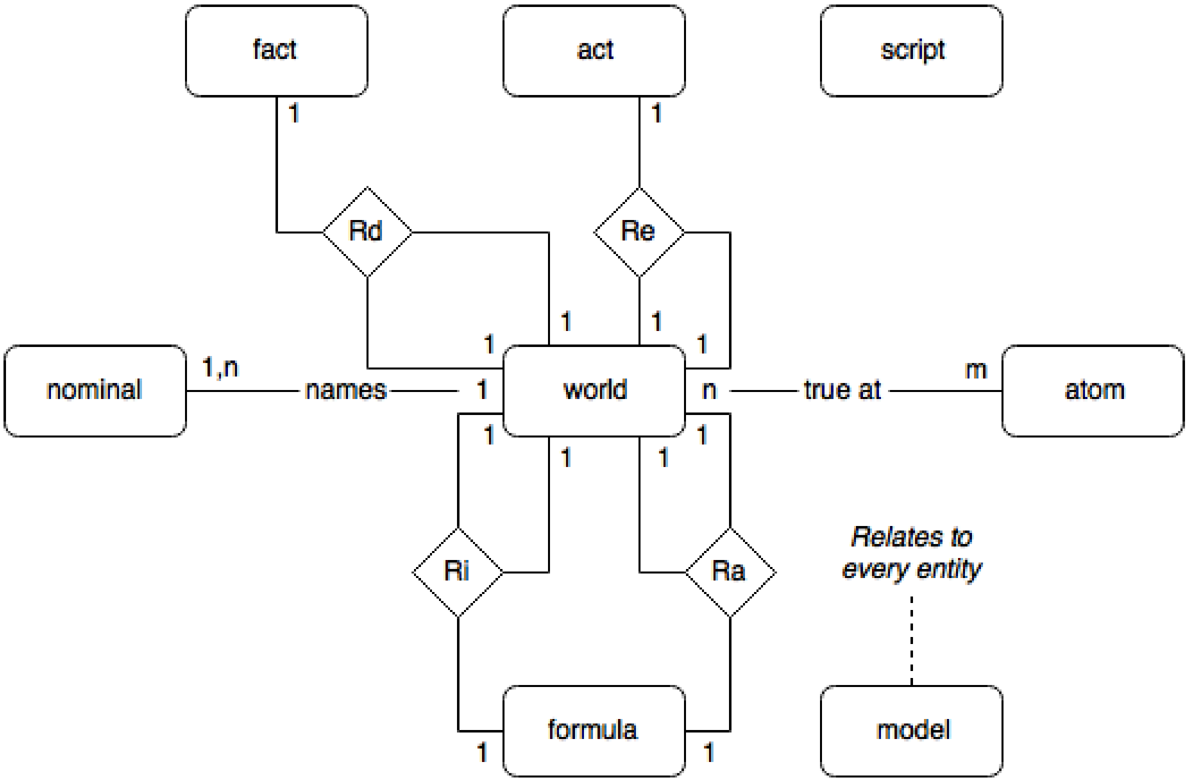


Figure 8.2: The database conceptual schema.

When translating the conceptual schema into the database language, we define the database logical schema depicted in Figure 8.3. Each box represents a table in the database, while lines represent relationships between tables. For each table, we provide the columns (or attributes) that it includes, together with their data type (*int* for an integer number, and *varchar* for text), as well as the primary key (PK) of each table. Finally, the relationships between tables are implemented by means of foreign keys (FK).

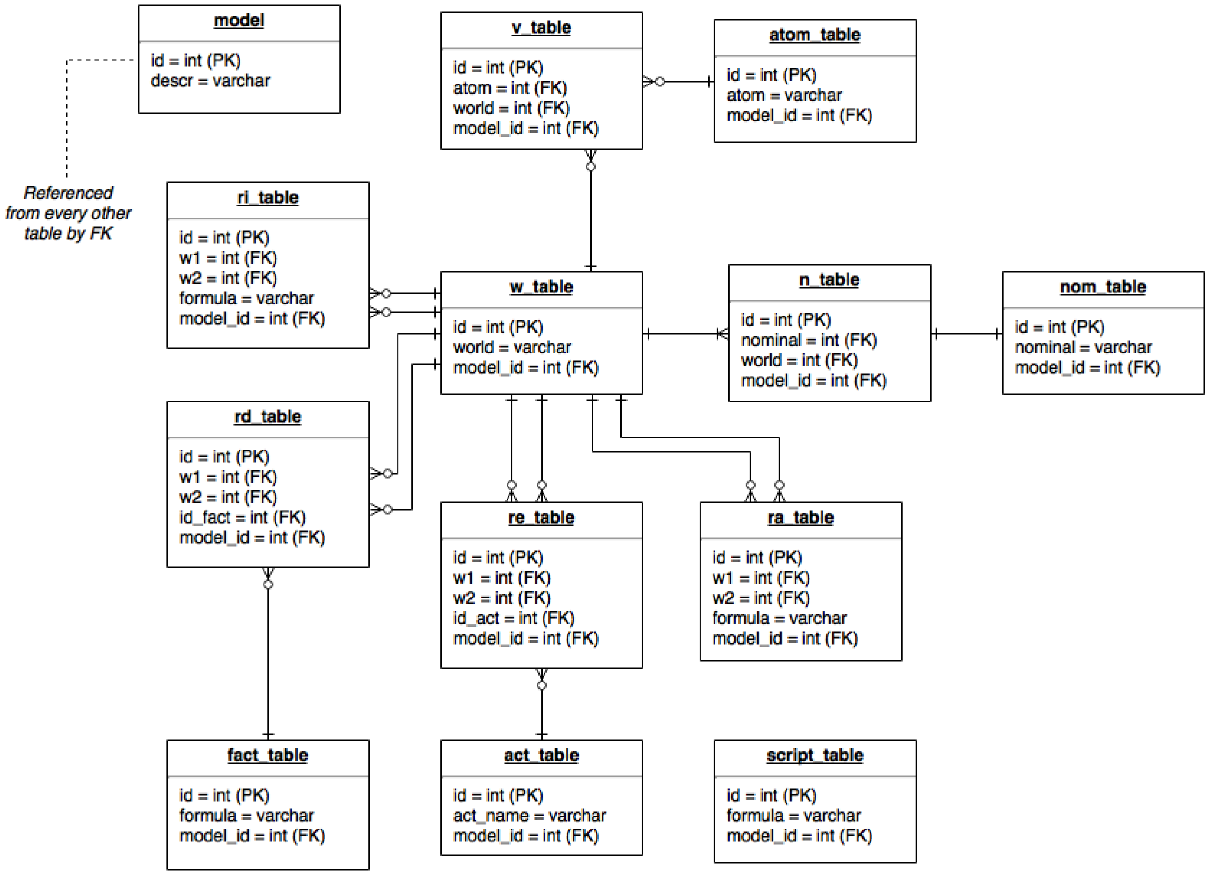


Figure 8.3: The database logical schema.

8.1.3 Implementation

The prototype for the SILOGIA has been developed as a web-based application. The fact that this implementation is web-based provides two clear advantages, with respect to considering a desktop-based application:

1. It makes the prototype reachable by virtually anyone having a computer and Internet connection.

2. It makes it available to be integrated in other web-based environments.

The technologies used in our implementation of the SILOGIA are the following:

- A relational database system based on SQL language (using a MySQL server, version 5.5.42) to store all the data about the structure of the formal models.

- The HTML language (version 5.2), with special emphasis on the Scalable Vector Graphics (SVG) module, to provide the user interface and draw the formal models embedded in it.

- The PHP language (version 7.0.8) to code all the algorithms of the Logic of Imagination Acts, aside from the functionalities needed to interact with the database.

At the time of the publication of this dissertation, a functional version of the SILOGIA is uploaded and accessible at a server from the Universitat Oberta de Catalunya through the following link:

http://einfmlinux1.uoc.edu/jcasasrom/

8.1.4 An Example

In order to show how the SILOGIA works, let's reproduce the same example we used in Section 7.6, while presenting the Logic for Imagination Acts. Throughout the upcoming pages, we reproduce all the steps that need to be followed, from start to end, in order to compute the same model as in the example of the previous chapter. In order to avoid making the example tedious, we will not refer to the previous example further in this section; nevertheless, the way we define the initial model and the parameters we use in every execution of every algorithm follow the same ones as in there. Note, however, that, unlike we did in the example of the previous chapter to make the example easier to follow, the SILOGIA does not write the negated atomic formulas.

First of all, we need to create a new Model for Imagination Acts containing just the real possible worlds, the nominals used to refer to them, their atomic valuation, and the sets of atomic formulas, actions, facts and scripts. Recalling the formal definitions of our logic, we need to specify the following elements (note that, while some of these elements are part of the formal model, others are part of the language, but are required in order to execute some of the algorithms afterwards):

$$
\begin{array}{lcl}
W & = & \{w\} \\
\text{NOM} & = & \{i\} \\
N & = & \{(i, w)\} \\
\text{ATOM} & = & \{p, q, r, s, t\} \\
\text{ACT} & = & \{\alpha\} \\
\text{FACT} & = & \{p\langle\rightarrow\rangle\neg t,\ \neg s \wedge t\langle\rightarrow\rangle\neg r, r \wedge s\langle\rightarrow\rangle(\neg q \vee t \vee (q \wedge \neg s))\} \\
\text{SCRIPT} & = & \{p \wedge q\langle\rightarrow\rangle^{\alpha}\neg s,\ s \wedge t\langle\rightarrow\rangle^{\alpha} p \wedge \neg q, t\langle\rightarrow\rangle^{\alpha} p,\ s[\rightarrow]^{\alpha}\neg r\}
\end{array}
$$

Creating this new model in the SILOGIA requires to fill up the corresponding fields, as it can be seen in Figure 8.4. Note that special symbols, such as logical operators or Greek letters, require to be written in a certain way (specified right below the corresponding fields).

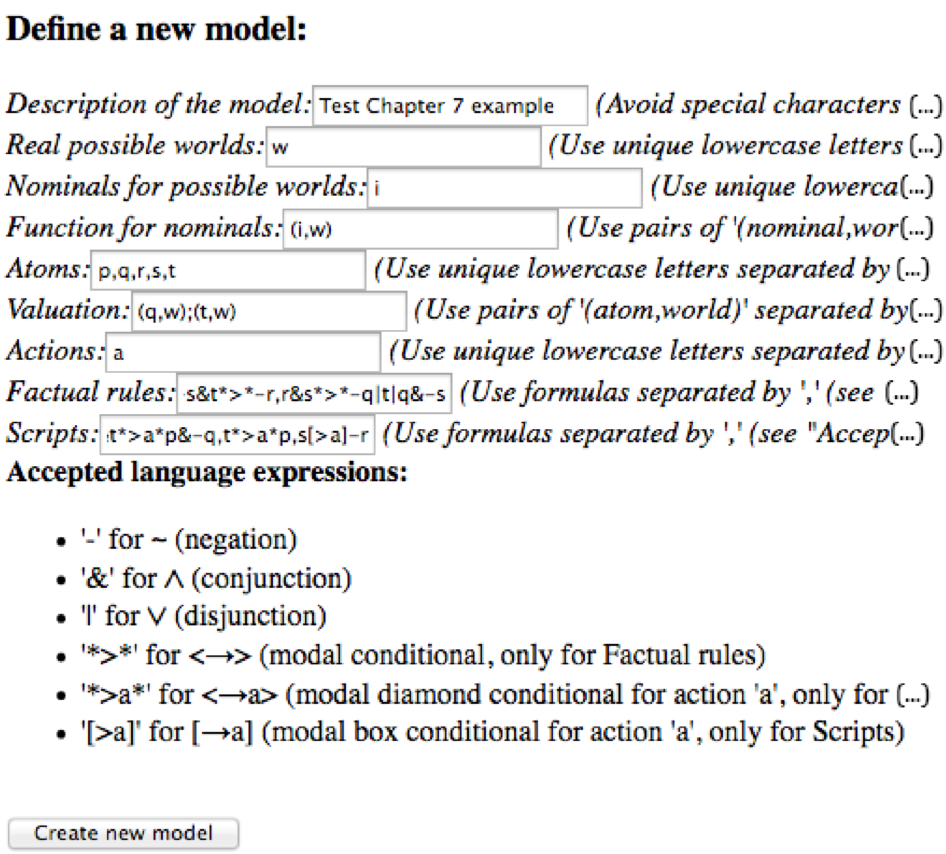


Figure 8.4: Creating a new model in the SILOGIA.

Once the information for the new model is provided, and as a result of clicking the “Create new model” button, the SILOGIA creates the new structure for the model, saves it into the database, and automatically draws it in the screen, as in Figure 8.5, which corresponds to the initial Working Model.

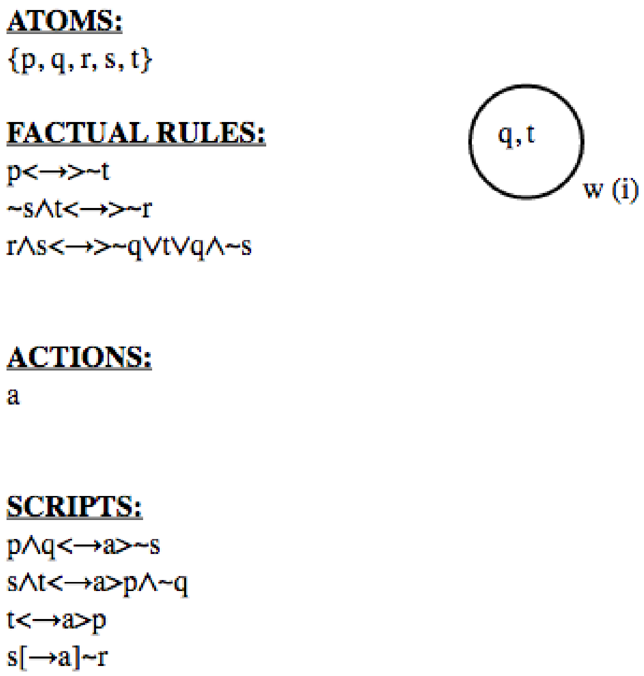


Figure 8.5: An initial model, without the execution of any algorithm in it.

In order to execute the Initialization algorithm, we need to provide a world of reference w^R, which must be selected from a drop-down list containing all available possible worlds, and an initial premise δ. The initial premise must be written, again, by following the convention specified in the application to express the special symbols by using characters available in a standard computer keyboard. Furthermore, and as we have already mentioned before, we require δ to be expressed in Disjunctive Normal Form. Thus, selecting w as the world of reference, and taking into account that $\mathrm{DNF}(\neg p \rightarrow r \wedge s) = p \vee (r \wedge s)$, setting the parameters for the execution of the `InitAlg` corresponds to screenshot in Figure 8.6.

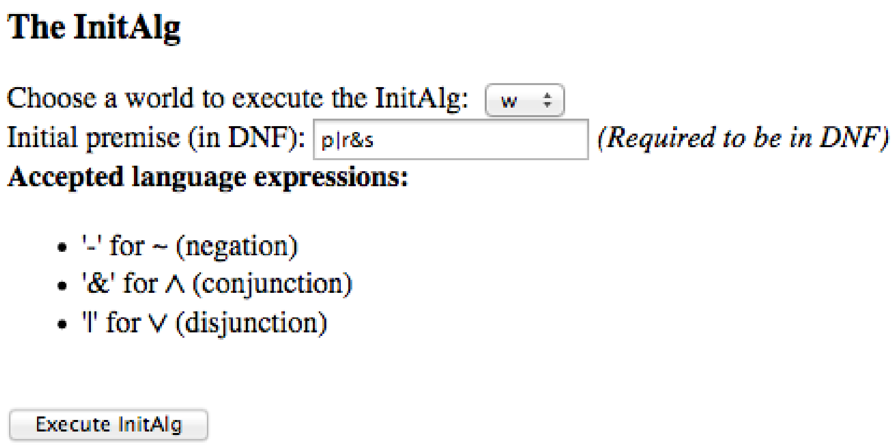


Figure 8.6: Specifying a call to the `InitAlg`.

When clicking the "Execute InitAlg" button, the SILOGIA follows the Initialization algorithm, creates new imaginary worlds, and clamps and imports the atomic formulas required by the `InitAlg`; note that the names of the new imaginary worlds, as well as the order in which the atomic formulas are shown in them, varies with respect to the example of Section 7.6; nevertheless, the model that results is the same as in the previous example. Then, it draws the expanded version of the model as in Figure 8.7. It is worth noting that the expansion of the model is done upon the Working Model, which is used to allow the user to run tests without permanently altering the models saved at the database; nevertheless, the user can also save the current state of the Working Model at the database.

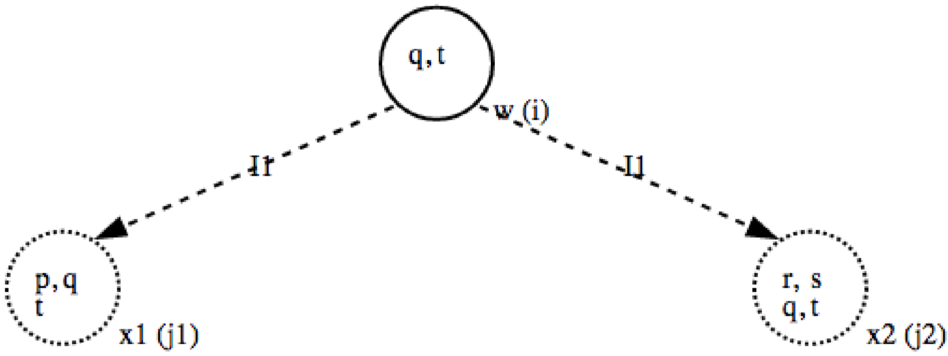


Figure 8.7: The model after executing a call to the `InitAlg`.

As a result of executing the `InitAlg`, the SILOGIA has created two new imaginary possible worlds, x_1 and x_2, and has added them to the model. Besides, and still following the `InitAlg`, it has determined which atomic formulas had to be clamped, according to $\delta = p \vee (r \wedge s)$, and which ones had to be imported from the world of reference w.

Now, we want to call the Description process in order to keep elaborating on the imaginary scenario by following the consequences expressed by a certain factual rule. In order to do so, we must choose, among the expanded set of possible worlds that we have available now, the new world of reference. In this case, we choose the imaginary possible worlds x_2 (which corresponds to world v_2 in the example of Section 7.6); furthermore, we need to select one of the factual rules we specified while creating the initial model, which will be $r \wedge s \langle \rightarrow \rangle \neg q \vee t \vee q \wedge \neg s$. The call to the `DescrAlg` is as shown in Figure 8.8.

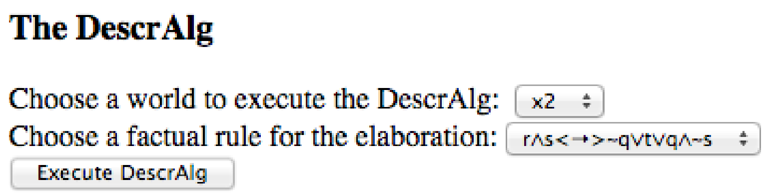


Figure 8.8: Specifying a call to the `DescrAlg`.

Once we click on the corresponding button, the SILOGIA calls the `DescrAlg`, creates and determines the valuation of the new imaginary worlds (that is, in case the antecedent of the factual rule being evaluated holds at the world of reference, which in this case does), and draws the expanded version of the model, as shown in Figure 8.9.

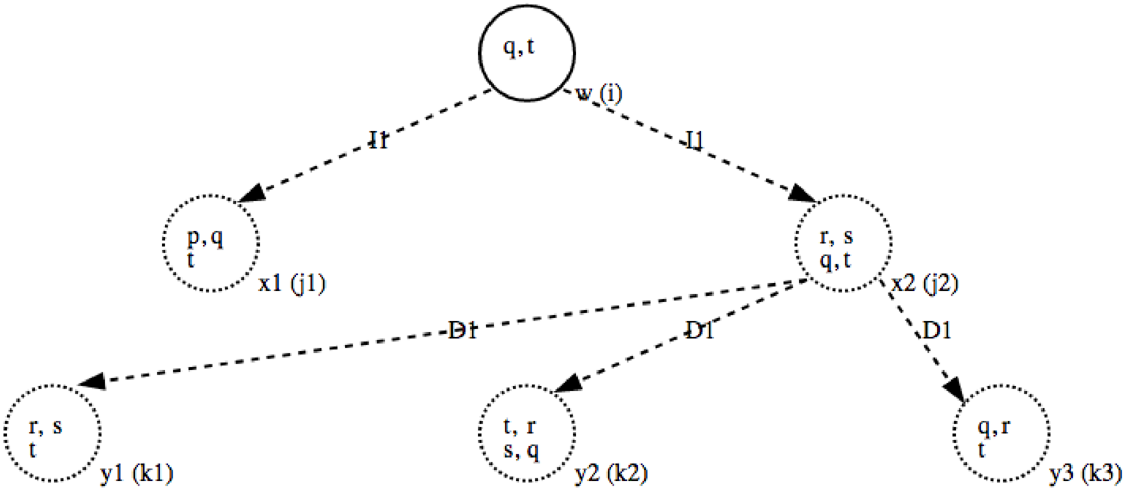


Figure 8.9: The model after a call to the `DescrAlg`.

In this case, the execution of the `DescrAlg` has created three new possible worlds y_1, y_2 and y_3, each one with their corresponding atomic formulas determined both by the factual rule $r \wedge s\langle\rightarrow\rangle\neg q \vee t \vee q \wedge \neg s$ and by the atomic formulas imported from the world of reference x_2.

Now, an execution of the `EvoAlg` requires, as in the previous algorithms, selecting the world of reference upon which we want to keep elaborating, which will be y_2 in this case, as well as one of the actions specified while defining the initial model, which will be a, as we can see in Figure 8.10. Recall that, even though the Evolution process elaborates on the imaginary world by following the consequences expressed by the scripts, those consequences are triggered by the agent imagining that a certain action takes place in there. Therefore, and as we explained while defining the Logic of Imagination Acts, the

`EvoAlg` does not require a script as its argument, but rather an action; then, the algorithm will indeed go through all the scripts belonging to this specific action in order to compute the resulting imaginary worlds.

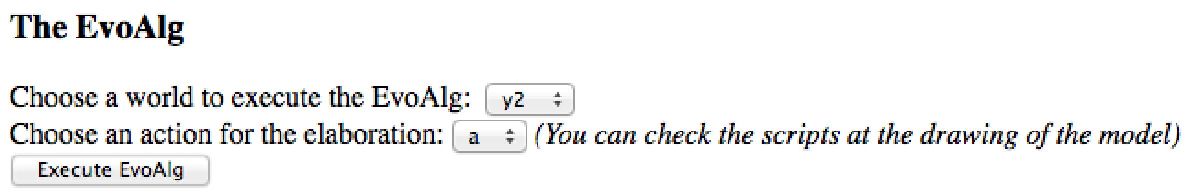


Figure 8.10: Specifying a call to the `EvoAlg`.

When calling the Default Evolution algorithm, the SILOGIA creates a queue with all the scripts related to action a (which, in the current example, are all the available scripts), sorts them accordingly, and creates and determines the valuation functions for the new imaginary worlds, as shown in Figure 8.11. Note that, in this particular case, not all the scripts are used by the algorithm; in particular, script $p \wedge q\langle\rightarrow\rangle^{\alpha}\neg s$ is not used, as the antecedent does not hold at y_2.

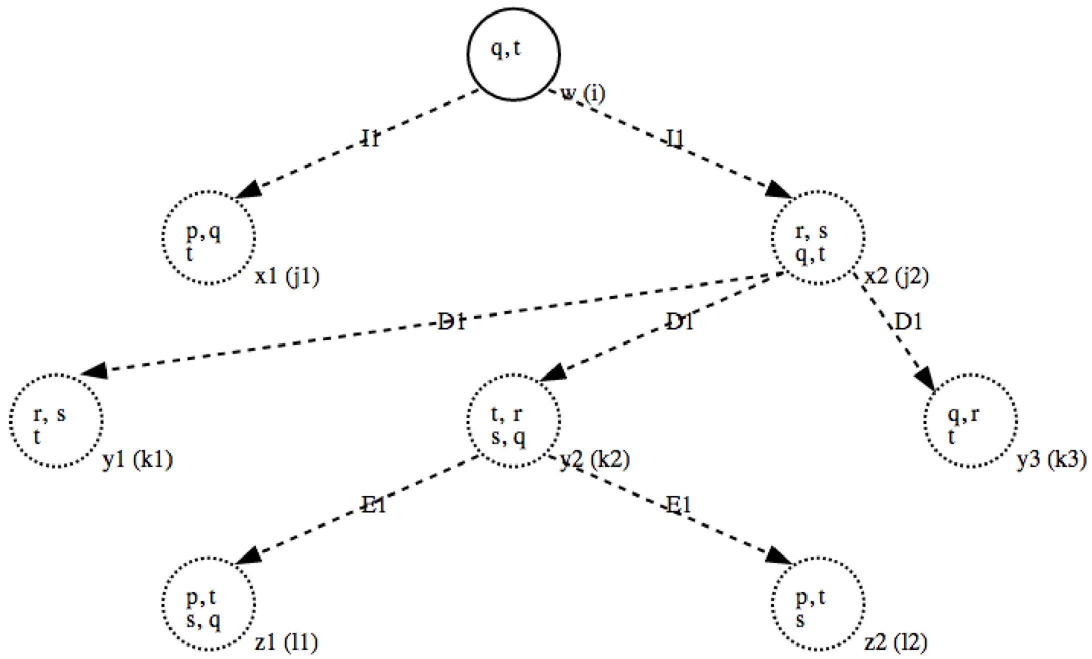


Figure 8.11: The model after a call to the `EvoAlg`.

As a result of executing the `EvoAlg`, the program has created two new imaginary possible worlds z_1 and z_2. In this occasion, and following the way the `EvoAlg` works, the atomic valuation of these new two imaginary worlds has been determined by three different scripts; namely, $s \wedge t\langle\rightarrow\rangle^{\alpha} p \wedge \neg q$, $t\langle\rightarrow\rangle^{\alpha} p$ and $s[\rightarrow]^{\alpha}\neg r$. The evaluation of the first two diamond-formulas has created the two worlds z_1 and z_2 while clamping in them the required atomic formulas of their consequent; in the first case, z_1 is required to satisfy p, whereas it is prevented to satisfy q, and, in the second case, z_2 is required to satisfy p. Then, when evaluating the remaining box-formula, the `EvoAlg` has ranged over worlds z_1 and z_2, while clamping the atomic formulas specified by its consequent (in particular, by preventing r from being true in any of the two worlds).

Lastly, we want to call the Unscripted Addition algorithm in order to represent how the agent voluntarily adds a new premise, which is not derived by following any factual rule nor script. We need to specify, again, a world of reference, which in this case corresponds to z_1, and a new premise (also expressed in DNF), which is $\neg p \vee r$ in this case, as shown in Figure 8.12.

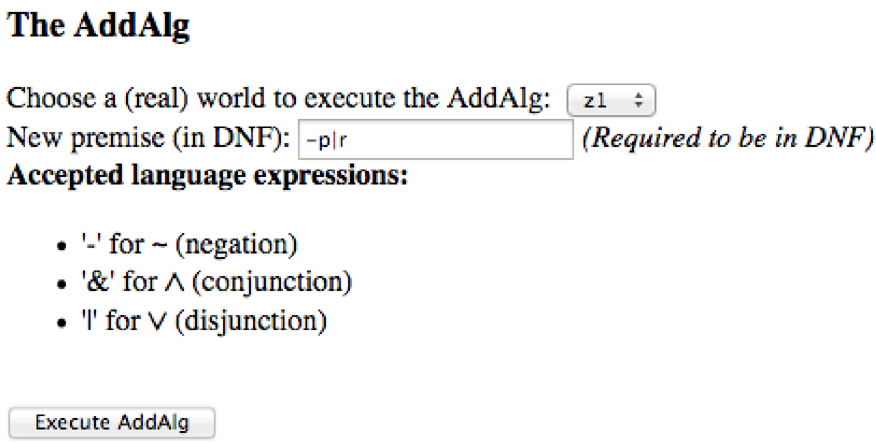


Figure 8.12: Specifying a call to the `AddAlg`.

Once we execute the `AddAlg`, the SILOGIA computes the new imaginary worlds that need to be created and specify their atomic valuation, as in Figure 8.13. At this point, the current example is at the same state as the former example was by the end of Section 7.6. As we already mentioned, throughout the example the user also has the chance to permanently store the changes made in the Working Model by saving them into the database. When clicking the corresponding "Save changes" button, the SILOGIA locates and saves into the database every element in the Working Model that did not exist in the model initially loaded or created.

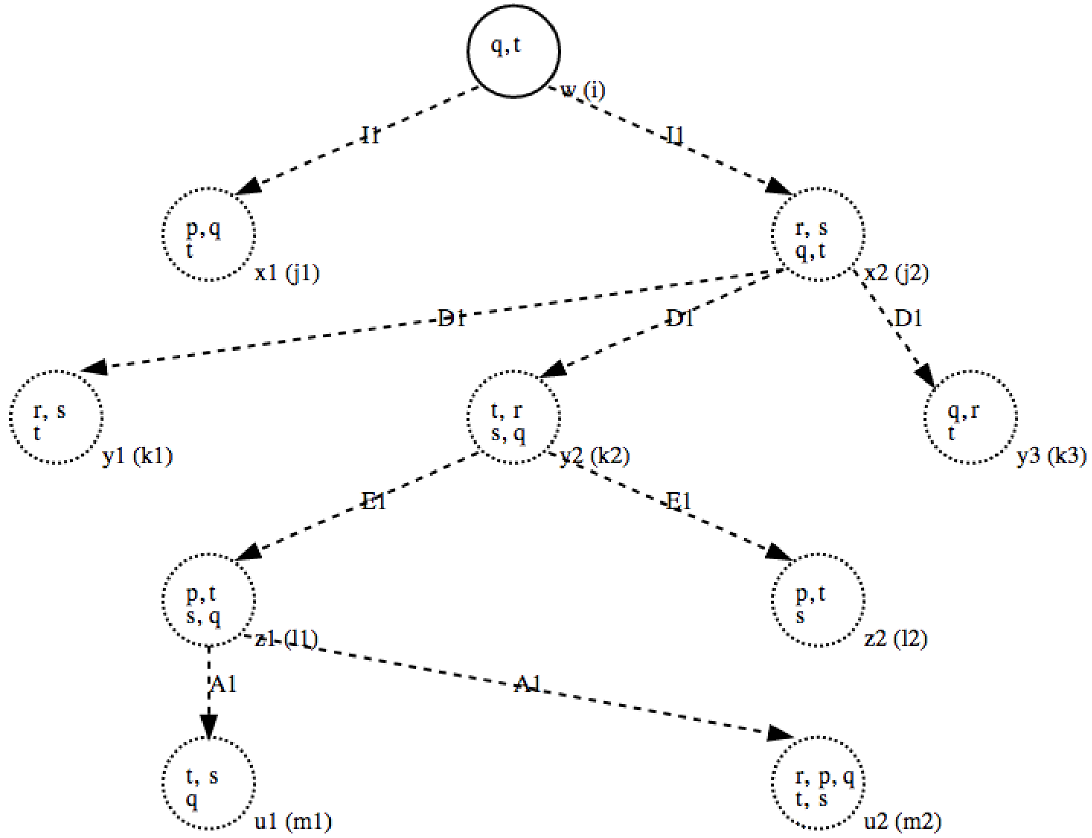


Figure 8.13: The model after a calling the `AddAlg`.

As we can see, we have been able to use the SILOGIA prototype to define and create a new model, execute a call to each one of the algorithms defined in the Logic of Imagination Acts, and we ended up reaching the same expanded model as we did while introducing how the logic works. Therefore, this shows that our implementation of the Logic of Imagination Acts indeed captures the dynamics of imagination defined by it, and so this implementation provides a valuable setting for executing computer-driven tests that simulate the way human being imagine, as described by our logic and the Common Frame for Imagination Acts.

Moreover, and although we have only called each algorithm once in this particular example, we could keep calling the algorithms as many more times as we wanted too, and the model would keep expanding by following the processes captured in each algorithm.

8.1.5 Towards Adding the Rhombus of Imagination

We want to point out to a feature which, although is not implemented in this prototype of the SILOGIA, could easily be added in a future version: namely, the addition of the Rhombus of Imagination as a support tool for computing the blue-print of the models created in the program.

As we already explained in Section 7.7, the Rhombus of Imagination can be easily used in conjunction with the Logic of Imagination Acts by following its own particular algorithm. Specifically, and while using the SILOGIA, it would be possible to add an additional layer giving the user to option to compute, at any time during the expansion of a specific model, the quantitative version of the Rhombus of Imagination corresponding to that particular act of imagination. The user would be required to select a leaf-world within the current model, and the SILOGIA would follow the algorithm defined in the previously mentioned section, which specifies how, given a Model for Imagination Acts, the Rhombus of Imagination of a certain imaginary story could be computed.

8.2 The Dynamics of Imagination in Video Games

Aside from the implementation of the computer-based prototype, the contributions made in the present work can also be a valuable input in other areas. In this section, we discuss and point to some ways our results could be used, specifically, in the field of video games. We do not provide results, though, as engaging in this new topic would fall outside the scope of the present work; nevertheless, we identify it as an interesting line of future work.

First and foremost, why have we suggested video games as a field in which our contributions could be applied? Most video games rely heavily on the immersion of the player within a virtual world. The term "immersion" refers to the subjective experience that players can have of being in one place (specifically, the virtual world), while being physically situated in another[1]. This phenomenon is considered a key factor in video games, as pointed out in [40], [23], [45], [49] or [46], among others. In particular, the authors in [33] state that "successful computer games all have one important element in common: they have the ability to draw people in".

The phenomenon of immersion of players within virtual worlds has received a great

[1]While "immersion" is the term that is often used to discuss this phenomenon in the field of video games, other fields, such as psychology, use the corresponding term "presence" (defined, for instance, in [57]).

deal of attention, and different authors propose alternative conceptions of the the term "immersion" and derived terms, as well as models to study such phenomenon.

For example, in [10], the author proposes a conceptual model for understanding involvement and immersion, and proposes to replace the concept of "immersion" with one of "incorporation"; specifically, the author discusses six kinds of involvement, and how their effects are in relation to the player's immersion and involvement. In [8], the authors identify three different levels of immersion that video game players can experience, each one characterized by a further loss of touch with the real world and a deeper physical and psychological involvement in the virtual world of the video game. Similarly, [23] proposes a different model for immersion that distinguishes three immersion forms, being one of them the *imaginative immersion*. In a nutshell, this form of immersion is related to how the player is absorbed by the game's story, and is characterized by awakening the player's empathy towards other characters in the game.

The motivations that a player has while controlling a character in a virtual world, therefore, often go beyond the goals that have been traditionally associated to most video games, such as achieving a high score (like in *Tetris*[2], or the more recent *Candy Crush Saga*[3]). In certain genres, such as role-playing games (or RPGs for short), the player is often faced with morally-challenging decisions that do not have a simple "good" of "right" solution, and which usually carry out many dilemmas that involve the player emotionally and empathically towards their characters (as in the Mass Effect series[4], for instance): these decisions affect the way other non-playable characters within the game "feel" about such decisions, and what their relation with the player's character is. The way the player thinks about those decisions, thus, is not merely based on a simple rule of thumb about what is the best outcome for the player, but it rather involves how the player would feel like and act like in that same situation, if it was real. Enjoyment, in this kind of games, is directly related to the degree of immersion of the player within the virtual world (as shown in [20], for example).

Having introduced this, the question now is: what determines whether a certain virtual world is successful in immersing the player into it? Or, looking at it from the opposite side, what typically prevents the player from immersing in such world?

[2] First version created by Alexey Pajitnov in June, 1984.
[3] Developed by King, released on April 12, 2012.
[4] Initially developed by BioWare, first game in the saga released in 2007.

8.2.1 Immersion and the Common Frame for Imagination Acts

In [39], the author points out, as one of the three conditions required to immerse a player into a virtual world, that "the conventions of the [virtual] world must be consistent, even if they don't match those of meatspace" (in page 69 of the previous reference); these conventions, as the author points out, are often defined by narrative elements and paradigms that "help the user align their expectations with the logic of the [virtual] world".

We argue that this consistency of conventions is directly related to what we have been calling reality-oriented development in this work; the fact that those conventions do not need to match, according to the previous reference, with those of the real world, would allow us to rephrase the "reality-oriented development" into the term "convention-oriented development", when applied to expectations that are not based on reality. While immersed in the virtual world, the factual rules and the scripts that the player uses in her imaginings and expectations are *not* those about the real world, but rather those that describe the virtual world she is in. By using our understanding of the dynamics of imagination and, more specifically, our Common Frame for Imagination Acts, we can easily represent those virtual conventions as being the sets of factual rules and scripts that the player uses, once she is already immersed within the virtual world of the video game.

Still in [39], the author argues how a sense of realism, which is often required for the immersion, can be divided into social realism and perceptual realism. In particular, social realism is described by [37] as "the extent to which a media portrayal is plausible or 'true to life' in that it reflects events that do or could occur in the nonmediated world"; we argue that, again, this conception of realism, or 'true-to-life', can easily be mapped into our conception of the particular set of reality-oriented scripts detailing how events (in this case, social events) would result, if they were real.

Taking all this into account we claim that there exists a direct correlation between the players' perception of consistency in the virtual worlds' rules (in relation to they expectations), and the processes related to the reality-oriented development of our Common Frame for Imagination Acts; namely, the Description and the Default Evolution processes. Moreover, we claim that our theory is a promising starting point that could be easily adapted to account for the dynamics of imagination involved in the players' immersion into virtual worlds. In order to do so, we would need to replace the sets of factual rules and scripts believed by the player about the real world, by the new sets of factual rules and scripts that define the logic behind the virtual world. Our Common Frame would then provide a detailed tool for analyzing and studying the dynamics involved in the players' immersion in video games and virtual worlds. In particular, by studying those dynamics in detail, we could understand what elements make it easier, for players, to immerse in virtual worlds, or what prevents them from doing so, or what happens (and which particular processes of imagination are involved) when players get "kicked out" of their immersion in a virtual world.

We claim that the answers to such questions can be studied and evaluated, from a dynamic perspective related to how players engage in the virtual worlds being presented, by using our dynamic approach to the creation and development of imaginary worlds. If the player, at any point during the game, disengages from the fiction (while throwing her controller in the air and shouting "come on, this does not make any sense!"), then we argue that it could be because the virtual world, and the events taking place in it, conflict with either the rules previously specified by the game itself, or with some of the player's own expectations about how that virtual world should be like, based on "convention-oriented development". We provide a list summarizing some of the topics we think would be interesting to study, with respect to the relation of our work with video games:

1. Is there any correlation between how different players use the processes involved in their imagination capacities by their own, and the degree in which they immerse in virtual worlds in video games?

2. Are there any kind of virtual worlds which are, generally speaking, more prone for the players to immerse into them than others? If so, what characterizes those worlds, and what makes them different? In particular, do they share more factual rules and scripts with reality, and thus this is what makes them more "easily reachable" or more "close" for the players (in terms of having to vary their real-world rules less than in other cases)?

3. Why do certain events cause some players to "disengage" from a virtual world? How are such events related with the particular set of believed factual rules and scripts of that player? Could this phenomenon be related to the topics of imaginative resistance (see [53], for instance)?

4. Up to which point do players change their behavior, whenever they incarnate a character within an immersive role-playing experience? How does incarnating a character with different motivations and desires affect how the player behaves within

the game? Is there a kind of “line” that players do not usually want to cross in their behavior, even when they are aware that they are just incarnating a video game character?

5. How does virtual reality change the immersion in virtual worlds, with respect to video games shown in a computer or a TV screen? Up to which point constraining the external stimulus of the players by using virtual reality goggles, headphones and controllers affect how easily they immerse in the virtual world, with respect to traditional video game platforms?

Taking into account all the considerations we discuss during this section, and also considering the set of research questions we have just given, we claim that the results of our current work can indeed be a valuable input in the field of video games. By identifying and giving a detailed account of how the dynamics involved in the creation and development of imaginary worlds work, we provide the background setting needed to study the role of the players’ capacity to imagine with respect to the phenomenon of immersion in virtual worlds. This topic can be interesting not only regarding video games, but also when considering broader uses of systems that involve virtual reality.

Chapter 9

Conclusions and Future Work

> "Who are YOU?" said the Caterpillar. [...] Alice replied, rather shyly, "I – I hardly know, Sir, just as present —at least I know who I WAS when I got up this morning, but I think I must have changed several times since then."
>
> —Lewis Carroll
>
> *Alice's Adventures in Wonderland*

In this last chapter, we summarize the main results of the present work. We start by going over our set of initial objectives, and we argue that our work fulfills them; furthermore, we identify the set of contributions that derive from our results, including a summary of publications and communications. After that, we discuss some additional conclusions we extract from the interdisciplinary methodology used in this dissertation, and we end up by pointing out to some interesting future topics of research that can follow from our results.

9.1 Achievement of the Objectives

The main goal in this work was to study and analyze how imaginary worlds are created and developed as a result of voluntary acts of imagination, and which are the particular mechanisms involved in doing so.

While setting this goal, we already established that our analysis was going to be based on three different approaches, related to three different fields: a philosophical analysis of some influential theories about voluntary imagination acts, the definition of a formal system capable of capturing the dynamics of imagination, and the use of algorithmic and programming techniques to both enhance the formal system with algorithms, and to implement a prototype of it using programming languages. In this sense, we claim that our work succeeds in providing a detailed, interdisciplinary analysis of the dynamics of imagination, and that the contributions that result from it do satisfy the main goal of our work.

In relation to the specific objectives we pointed out in the Introduction of this work in Chapter 1, we argue that they all have been fulfilled:

O1. To identify, through a critical review of some of the most influential theories detailing how voluntary acts of imagination work, the mechanisms involved in the creation and development of imaginary worlds.

We began our critical review by studying how the influential theory from Nichols and Stich, the cognitive theory of pretense, identifies the way imaginary worlds are created and elaborated. Then, we compared Nichols and Stich's theory to two recent theories of imagination that also provide and account of the dynamics of imagination; namely, Langland-Hassan's and Williamson's theories. After reviewing them, we showed how the mechanisms they all identify about the dynamics of voluntary imagination acts are closely related. In particular, we showed, by the end of Chapter 2, how all those theories share an underlying procedural structure of voluntary imagination acts.

Furthermore, we discuss, throughout Chapter 6, our concerns regarding the degree of detail of the previously reviewed theories. After arguing that the dynamics involved in developing an imaginary world should be further refined, we propose a new philosophical theory specially accounted to represent such refinement: the Common Frame for Imagination Acts. After introducing the Rhombus of Imagination as a supporting tool specially conceived for our system, we show how our refinement of the processes involved in the development of imaginary worlds is indeed a valuable result and that, without it, the previous theories would not be able to account for the difference between certain kinds of imagination acts.

O2. To define a formal system capable of capturing, through a dynamic process captured by an algorithm, the mechanisms involved in the creation and development of imaginary worlds.

The Logic for Imaginary Scenarios, presented in Chapter 4, defines the syntax and the semantics of a formal logic aimed to capture, by using a single algorithm, how imaginary worlds are created and developed, following the theories previously reviewed. Even though we later decided, due to the insights we got from this logic, to define a distinct system, the Logic of Imaginary Scenarios represents a first approach to the basics of the dynamics of imagination acts, which are represented by combining a formal logic and an algorithm.

Later on, and by using the Common Frame for Imagination Acts as our underlying theory, we define, in Chapter 7, a whole new logical system that captures the refinement proposed in the dynamics of creating and developing imaginary worlds. In particular, we show that our logic allows to elaborate an imaginary world in many different ways, while also accounting for both the notion of reality-oriented factual rules and scripts, and for a hidden agentiveness that we identified, and which can be found embedded at different times throughout an act of imagination.

O3. To consider the applications that this formal system and the algorithm defined in it could have, both when implemented as a computer program, and in relation to other fields where imaginary worlds has a critical importance.

While taking the Logic of Imagination Acts as our referent, in Chapter 8 we introduce the SILOGIA: an implementation of a computer-based prototype that reproduces the dynamics of how imaginary worlds are created and developed, as defined by our logic. This prototype represents both an opportunity to test our formal system and gain still further insights on the way that creating and developing imaginary worlds works; furthermore, the prototype provides a basis for a simulation of the human imagination. In addition, we argue how our contributions can also be a valuable input, in particular, in the field of video games, where immersion in virtual worlds is critical for the enjoyment of the players.

9.2 Our Contributions

The following list presents the contributions made in our work:

C1. A theory-independent layer identifying, through a critical review of some influential theories of imagination, the mechanisms involved in the creation and elaboration of imaginary worlds.

By the end of Chapter 2, our review of Nichols and Stich's, Williamson's and Langland-Hassan's theories allow us to outline the similarities those theories identify in the mechanisms related to voluntary acts of imagination. As a result of that, we provide a comparative table establishing a connection between the mechanisms distinguished by each theory, and we show how they can all be abstracted in a theory-independent layer formed by three different mechanisms.

C2. A new theory specially suited to account for the dynamics of the creation and development of imaginary worlds, with special emphasis on the processes involved in the reality-oriented development.

The Common Frame for Imagination Acts proposed in Chapter 6 represents a novel and detailed account of the dynamics of voluntary imagination acts. This theory is not only useful when defining our new formal approach, but it is also a valuable contribution at a philosophical and analytic level with respect to the previously reviewed theories, being this one specially conceived for capturing the dynamic aspects of imagination. Furthermore, we also define the Rhombus of Imagination, a tool specially suited for the procedural analysis and classification of imagination acts, and which can be used in conjunction with the Common Frame for Imagination Acts.

C3. A dynamic logical system that captures acts of imagination using a single algorithm, and that allows to expand a model based on single-agent epistemic logic.

The Logic of Imaginary Scenarios, introduced in Chapter 4, provides a first approach to a dynamic logic that uses a single algorithm to capture the creation and development of imaginary worlds. Furthermore, we argue that our logic is the first formal approach specifically intended to capture the dynamics of imagination acts, compared to the other existing logics of imagination reviewed in Chapter 3.

C4. A dynamic logical system that captures the creation and development of imaginary worlds in a modular and detailed way by using different algorithms.

The Logic of Imagination Acts, defined in Chapter 7 as a more refined version of a dynamic logic for imagination, provides a detailed account of the distinct processes that take part in a voluntary imagination act. Furthermore, the way it is defined allows to account for the many different ways in which an imaginary world can be elaborated, to capture the notion of factual rules and scripts believed by the agent, and to account for the agentiveness involved at certain moments of imagination acts.

C5. A prototype of a computer program that reproduces the dynamics of the creation and development of imaginary worlds, as defined by our latter logical system.

The prototype for the SILOGIA, which we present in Chapter 8, provides a tool for running computer simulations that mimic the way human beings create and develop imaginary worlds. This is also a valuable source for obtaining still more insights regarding how the dynamics of imagination works, as defined by our latter logical system.

Dissemination

Aside from the contributions listed in the previous paragraphs, the following list of publications and communications derive from the present work:

- J. Casas-Roma, M. E. Rodríguez, and A. Huertas. A dynamic analysis of imagination acts. In *Philosophy of Imagination*, 2018. Oral communication

- J. Casas-Roma, M. E. Rodríguez, and A. Huertas. Towards a dynamic logic for imagination acts. In *Workshop on Logic(s) for Imagination*, 2018. Oral communication

- J. Casas-Roma, M. E. Rodríguez, and A. Huertas. A common frame for imagination acts. *Submitted*

- J. Casas-Roma, A. Huertas, and F. Manyà. Solving maxsat with natural deduction. In I. Aguiló, R. Alquézar, C. Angulo, A. Ortiz, and J. Torrens, editors, *Recent Advances in Artificial Research and Development; Proceedings of the 20th International Conference of the Catalan Association for Artificial Intelligence*, Frontiers in Artificial Intelligence and Applications, pages 186–195. IOS Press, 2017. ISBN: 978-1-61499-805-1 (print) — 978-1-61499-806-8 (online)

- J. Casas-Roma, A. Huertas, and M. E. Rodríguez. An analysis of imagination acts. In *Workshop on Hybrid and Intensional Logic*, 2017. Oral communication

- J. Casas-Roma, A. Huertas, and M. E. Rodríguez. Towards a shared frame for imaginative episodes. In *Fourth Philosophy of Language and Mind Congress*, 2017. Oral communication

- J. Casas-Roma, A. Huertas, and M. E. Rodríguez. Imagination as change: Creating new worlds. In *Non-Classical Logics: Theory and Applications*, 2016. Oral communication

- J. Casas-Roma, A. Huertas, and M. E. Rodríguez. Beyond knowledge: When an agent imagines. In *Advances in Model Logic*, 2016. Oral communication

- J. Casas-Roma, A. Huertas, and M. E. Rodríguez. Towards a semantics for the dynamic imagination logic. In *European Logic Colloquium of the Association for Symbolic Logic*, 2016. Oral communication

9.3 Conclusions

In this work, we aimed to deepen our understanding of imagination by analyzing in detail the mechanisms involved in the creation and development of imaginary worlds. We argue that our dissertation is indeed a valuable step towards a deeper understanding of imagination, through a detailed analysis of their dynamics. The achievement of our initial objectives and the contributions we have just presented support our claim.

Regarding our initial objectives, the spin that our work took around the midpoint resulted in objectives O1 and O2 being achieved in two different steps, or, more precisely, by two different contributions of our dissertation.

In particular, objective O1 has been achieved, firstly, by the critical analysis of the reviewed theories and the identification of the mechanisms they propose for the dynamics of imagination; this, in turn, gave rise to the Logic of Imaginary Scenarios, which is our first approach to a dynamic logic of imagination, and which corresponds to a first step in achieving objective O2.

Afterwards, the refinement of the dynamics of imagination we made through defining the Common Frame for Imagination Acts represents, as well, an accomplishment of objective O1; similarly, we used this new theory to define the Logic of Imagination Acts, which represents a more detailed and modular account to a dynamic logic of imagination, and which also fulfills objective O2 at a more specific level.

Aside from the particular contributions, this work already begun with a deep interdisciplinary aim. We believed that, while working on a topic such as imagination, we could greatly benefit from the resources and methods available at different areas. Philosophy has given us the theories detailing how acts of imagination work; formal logic has given us the perfect setting to define this ideas rigorously and to create formal models to represent them; algorithms have given us the tools needed to put our logic in motion and capture the changes of the dynamics of imagination into our formal models, as well as opening the doors towards computer-based simulations of formal representations of imagination acts.

Furthermore, the benefits of carrying out an interdisciplinary research can be found not only while moving forward with our initial research plan, but also as backwards insights and feedback that can be used to enrich previous step of the work. The moment we finished the definition of our formal system, we were able to gather new and useful insights on the philosophical account of imagination, which ultimately contributed to this field in the form of a new theory. This new theory, in turn, gave us a better setting to engage in a new formal system, more detailed and precise than the previous one, and which ended up being clearly better for capturing the dynamics of imagination acts. Our conclusions, then, also include an awareness of the fact that this work has greatly benefited from the interactions between these three different disciplines.

Considering the bigger picture, we claim that our work constitutes a valuable study of the dynamic processes involved in voluntary imagination acts, and that our contributions are philosophical, formal and applied. Aside from the specific contributions we already detailed, we also point out to other fields where our results could be applied to provide new useful insights.

9.4 Future Work

Last but not least, we consider some lines of future work in which we believe that our results could provide valuable insights for future research:

- By using the Common Frame for Imagination Acts and the Rhombus of Imagination, it is possible to measure and classify different kinds of imagination acts, according to their procedural structure. An interesting line of future work in this direction would be to study and classify every distinct kind of imagination act appearing in the relevant literature, and see up to which point those different kinds of imagination acts can be distinguished, according to their dynamics. Furthermore, it would also be interesting to consider whether there exists a hierarchical classification of the procedural structure of imagination acts, in which certain kinds of imagination acts are identified, with respect to the processes they use, as being subclasses or more specific instances of other more general classes.

- Even though we have argued that the Logic of Imagination Acts is better suited to represent a detailed account of the dynamics of voluntary imagination acts, the Logic of Imaginary Scenarios is still an interesting approach to a dynamic logic for imagination. Besides, the formal properties of the Logic of Imaginary Scenarios has been studied in greater depth than those of the Logic of Imagination Acts. Due to this, and as a result of our formal study for that logic, we are already working on the definition of a calculus for the Logic of Imaginary Scenarios, which we think that will prove our system to be sound and complete.

- While discussing the properties of the Logic of Imaginary Scenarios, by the end of Chapter 4, we already suggested how it would be possible to define alternative versions of the imagination algorithm in order to represent and capture certain psychic disorders related to imagination, such as the Capgras delusion. Even though we did not mention this topic again in Chapter 7, we argue that the Logic of Imagination Acts provides an even better setting for doing so. Therefore, we think that it would be very interesting to define alternative versions of some of the processes involved in the creation and development of imaginary worlds in order to model and simulate psychic disorders affecting our capacity to imagine.

- On a more formal level, it would be interesting to study the properties of the Logic of Imagination Acts, as we did with the former logical system. Moreover, it would also be interesting to define a calculus for it. Once an axiomatic system for this logic is defined, it would be useful to see which consequences can be derived from the way imagination works in it and the axioms defined in the logical system.

- As we already point out by the end of Chapter 8, a field in which the study of imaginary worlds could provide valuable insights is in video games and, more generally, in any field involving virtual worlds and virtual reality. A closer look to the processes followed in the creation and elaboration of such worlds within the mind of the agent could tell things such as why some virtual worlds are easier to engage in than some others, or why certain people disengage from a fiction when something specific takes places in there. It would also be interesting to study how the immersion in a virtual world changes, depending on whether it is presented to the player in a TV or computer screen, or by using virtual reality devices such as goggles.

- Another interesting turn that could be defined using the Logic of Imagination Acts would be to model the way thought experiments are carried out. This would involve reproducing the creation of the imaginary world where the experiment is about to take place, reproducing how the factual rules and scripts determine the next states of that imaginary world, and, whenever the agent is able to reach a conclusion within the imagining, one could define a new dynamic procedure responsible for "importing" the new knowledge back into the real world. This new setting would also need to take into account existing works on belief-revision, and considering how similar and different those approaches are to the Logic of Imagination Acts, when representing thought experiments.

Bibliography

[1] A. Baltag and S. Smets. A qualitative theory of dynamic interactive belief revision. *Logic and the foundations of game and decision theory (LOFT 7)*, 3:9–58, 2008.

[2] F. Berto. On conceiving the inconsistent. In *Proceedings of the Aristotelian Society*, volume 114, pages 103–121. Wiley Online Library, 2014.

[3] F. Berto. Aboutness in imagination. *Philosophical Studies*, 2017. *Forthcoming.*

[4] F. Berto. Impossible worlds and the logic of imagination. *Erkenntnis*, 82(6):1277–1297, 2017.

[5] P. Blackburn. Representation, reasoning, and relational structures: a hybrid logic manifesto. *Logic Journal of the IGPL*, 8(3):339–365, 2000.

[6] P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic.* Cambridge University Press, 2010.

[7] P. Blackburn and J. van Benthem. Modal logic: a semantic perspective. In P. Blackburn, J. van Benthem, and F. Wolter, editors, *Handbook of Modal Logic*, volume 3, chapter 1, pages 1–84. Elsevier, 2007.

[8] E. Brown and P. Cairns. A grounded investigation of game immersion. In *CHI EA '04 CHI '04 Extended Abstracts on Human Factors in Computing Systems*, pages 1297–1300. ACM, 2004.

[9] J. R. Brown and Y. F. Fehige. Stanford encyclopedia of philosophy: Thought experiments, 2014. URL date: 2017-11-16.

[10] G. Calleja. Digital game involvement: A conceptual model. *Games and Culture*, 2(3):236–260, 2007.

[11] J. Casas-Roma, A. Huertas, and F. Manyà. Solving maxsat with natural deduction. In I. Aguiló, R. Alquézar, C. Angulo, A. Ortiz, and J. Torrens, editors, *Recent Advances in Artificial Research and Development; Proceedings of the 20th International Conference of the Catalan Association for Artificial Intelligence*, Frontiers in Artificial Intelligence and Applications, pages 186–195. IOS Press, 2017. ISBN: 978-1-61499-805-1 (print) — 978-1-61499-806-8 (online).

[12] J. Casas-Roma, A. Huertas, and M. E. Rodríguez. Beyond knowledge: When an agent imagines. In *Advances in Model Logic*, 2016. Oral communication.

[13] J. Casas-Roma, A. Huertas, and M. E. Rodríguez. Imagination as change: Creating new worlds. In *Non-Classical Logics: Theory and Applications*, 2016. Oral communication.

[14] J. Casas-Roma, A. Huertas, and M. E. Rodríguez. Towards a semantics for the dynamic imagination logic. In *European Logic Colloquium of the Association for Symbolic Logic*, 2016. Oral communication.

[15] J. Casas-Roma, A. Huertas, and M. E. Rodríguez. An analysis of imagination acts. In *Workshop on Hybrid and Intensional Logic*, 2017. Oral communication.

[16] J. Casas-Roma, A. Huertas, and M. E. Rodríguez. Towards a shared frame for imaginative episodes. In *Fourth Philosophy of Language and Mind Congress*, 2017. Oral communication.

[17] J. Casas-Roma, M. E. Rodríguez, and A. Huertas. A common frame for imagination acts. *Submitted*.

[18] J. Casas-Roma, M. E. Rodríguez, and A. Huertas. A dynamic analysis of imagination acts. In *Philosophy of Imagination*, 2018. Oral communication.

[19] J. Casas-Roma, M. E. Rodríguez, and A. Huertas. Towards a dynamic logic for imagination acts. In *Workshop on Logic(s) for Imagination*, 2018. Oral communication.

[20] G. Christou. The interplay between immersion and appeal in video games. *Computers in Human Behavior*, (32):92–100, 2014.

[21] A. Costa-Leite. Logical properties of imagination. *Abstracta*, 6(1):103–116, 2010.

[22] G. Currie and I. Ravenscroft. *Recreative Minds: Imagination in Philosophy and Psychology*. Oxford University Press, 2002.

[23] L. Ermi and F. Mäyrä. Fundamental components of the gameplay experience: Analyzing immersion. In S. De Castell and J. Jenson, editors, *Worlds in Play: International Perspectives on Digital Games Research*, chapter 3, pages 37–54. Peter Lang Publishing, Inc., 2007.

[24] J. Garson. Stanford encyclopedia of philosophy: Modal logic, 2014. URL date: 2017-10-18.

[25] T. Gendler. Stanford encyclopedia of philosophy: Imagination, 2011. URL date: 2016-07-30.

[26] L. Goble, editor. *The Blackwell Guide to Philosophical Logic*. Wiley-Blackwell, 2001.

[27] D. Harel. Dynamic logic. In D. M. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume II: Extensions of Classical Logic, chapter 10, pages 497–604. D. Reidel Publishing Company, 1984.

[28] J. Hintikka. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell University Press, 1962.

[29] J. Hintikka. The modes of modality. In *Models for Modalities*, pages 71–86. Springer, 1969.

[30] J. Hintikka. The intentions of intentionality. In *Rudolf Carnap, Logical Empiricist*, pages 183–216. Springer, 1975.

[31] W. Hodges. *Logic: An Introduction to Elementary Logic*. Penguin Books, second edition 2001 edition, 1977.

[32] M. Jago. *The Impossible: An Essay on Hyperintensionality*. Oxford University Press, 2014.

[33] C. Jennet, A. L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, and A. Walton. Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studiest*, 9(66):641–661, 2008.

[34] A. Kind, editor. *The Routledge Handbook of Philosophy of Imagination*. Routledge, Taylor & Francis Group, 2016.

[35] P. Langland-Hassan. On choosing what to imagine. In A. Kind and P. Kung, editors, *Knowledge Through Imagination*, chapter 2, pages 61–84. Oxford University Press, 2016.

[36] D. Lewis. *Counterfactuals.* Blackwell Publishing, 1973.

[37] M. Lombard and T. Ditton. At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication*, 2(3), 1997.

[38] M. Manzano and A. Huertas. *Lógica para principiantes.* Alianza Editorial, 2004.

[39] A. McMahan. Immersion, engagement and presence. In M. J. P. Wolf and B. Perron, editors, *The Video Game, Theory Reader*, chapter 3, pages 67–86. Routledge, Taylor and Francis Group, 2003.

[40] L. E. Nacke and A. Drachen. Towards a framework of player experience research. In *Second International Workshop of Evaluating Player Experience in Games, FDG*, 2011.

[41] S. Nichols. Imagination and the puzzles of iteration. *Analysis*, 63(3):182–187, 2003.

[42] S. Nichols, editor. *The architecture of the imagination: New essays on pretence, possibility, and fiction.* Clarendon Press / Oxford University Press, 2006.

[43] S. Nichols and S. P. Stich. A cognitive theory of pretense. *Cognition*, 74(2):115–147, 2000.

[44] I. Niiniluoto. Imagination and fiction. *Journal of Semantics*, 4(3):209–222, 1985.

[45] N. Ravaja, M. Salminen, J. Holopainen, T. Saari, J. Laarni, and A. Järvinen. Emotional response patterns and sense of presence during video games: potential criterion variables for game design. In *NordiCHI '04, Proceedings of the third Nordic conference on Human-computer interaction*, pages 339–347. ACM, 2004.

[46] R. Rouse III. Games on the verge of a nervous breakdown: Emotional content in computer games. *ACM SIGGRAPH Computer Graphics*, (35):6–10, 2001.

[47] G. Ryle. *The Concept of Mind.* Hutchinson, 1949.

[48] C. Strasser. Stanford encyclopedia of philosophy: Non-monotonic logic, 2016. URL date: 2017-12-08.

[49] P. Sweetser and P. Wyeth. Gameflow: A model for evaluating player enjoyment in games. *Computers in Entertainment (CIE) - Theoretical and Practical Computer Applications in Entertainment*, (3), 2005.

[50] H. Van Ditmarsch, J. Y. Halpern, W. van der Hoek, and B. Kooi, editors. *Handbook of Epistemic Logic*. College Publications, 2015.

[51] F. R. Velázquez-Quesada. *Small steps in dynamics of information*. Institute for Logic, Language and Computation, 2011.

[52] K. L. Walton. *Mimesis as make-believe: On the foundations of the representational arts*. Harvard University Press, 1990.

[53] K. L. Walton. On the (so-called) puzzle of imaginative resistance. In S. Nichols, editor, *The Architecture of the Imagination: New Essays on Pretence, Possibility and Fiction*, chapter 8, pages 137–148. Clarendon Press, Oxford, 2006.

[54] H. Wansing. Remarks on the logic of imagination. a step towards understanding doxastic control through imagination. *Synthese*, 194(8):2843–2861, 2017.

[55] T. Williamson. *Knowledge and its Limits*. Oxford University Press, 2000.

[56] T. Williamson. Knowing by imagining. In A. Kind and P. Kung, editors, *Knowledge Through Imagination*, chapter 4, pages 113–123. Oxford University Press, 2016.

[57] B. G. Witmer and M. J. Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence, Teleoperators and Virtual Environments*, 3(7):225–240, 1998.

[58] S. Yablo. Is conceivability a guide to possibility? *Philosophy and Phenomenological Research*, 53(1):1–42, 1993.