Uniwersytet Warszawski
Wydział Filozofii i Socjologii

Jędrzej Kołodziejski
Student no. 345076

Bisimulational categoricity

Praca magisterska
na kierunku Filozofia
w zakresie logiki i filozofii nauki

Praca wykonana pod kierunkiem:
dra Bartosza Wcisły
w Instytucie Filozofii UW

Warszawa 2018

Oświadczenie kierującego pracą

Oświadczam, że niniejsza praca została przygotowana pod moim kierunkiem i stwierdzam, że spełnia ona warunki do przedstawienia jej w postępowaniu o nadanie tytułu zawodowego.

Data | Podpis kierującego pracą

Oświadczenie autora pracy

Świadom odpowiedzialności prawnej oświadczam, że niniejsza praca dyplomowa została napisana przez mnie samodzielnie i nie zawiera treści uzyskanych w sposób niezgodny z obowiązującymi przepisami.

Oświadczam również, że przedstawiona praca nie była wcześniej przedmiotem procedur związanych z uzyskaniem tytułu zawodowego w wyższej uczelni.

Oświadczam ponadto, że niniejsza wersja pracy jest identyczna z załączoną wersją elektroniczną.

Data | Podpis autora

Abstract

We introduce and investigate a notion of *bisimulational categoricity* – the property of having a unique model *up to bisimulation*, which is analogous to the well-studied notion of *categoricity* – the property of having a unique model *up to isomorphism*. Bisimulational categoricity turns out to be well-behaved, which is reflected in a nice characterisation we give: a complete modal theory has a unique model up to bisimulation iff all its models are bisimilar to finitely branching ones iff it has at least one finitely branching model. We develop a topological framework that allows us to provide a complete proof of the theorem which consists of (i) proofs of so far unknown facts, as well as (ii) new proofs of already known ones.

Keywords

bisimulation, modal logic, categoricity

Dziedzina pracy

[08.1] Filozofia

Tytuł pracy w języku polskim

Kategoryczność Bisymulacyjna

Contents

Chapter 1

Introduction

Modal logic is a surprisingly universal tool that allows us to formally describe a wide range of phenomena. It was originally designed to describe the notions of necessity and possibility with mathematical precision – which is reflected in this formalism's name: *modal* logic. However, this kind of problems do not have much to do with what the modal logic is as we know it today, as they are only a very special example of application of this much richer framework. A number of seemingly unrelated fields, such as formal epistemology, proof theory, reasoning about time, behaviour of programs or ethical duties, can be analysed with help of an elegant toolbox of modal formalism (e.g. see [1] for more details on the application of modal logic to modelling of epistemic notions, or [19] for its use in the formal proof theory, where the author calls the ease with which modal logic describes complex proof systems 'a miracle').

In the light of the great success of modal logic, it is natural to ask where do the number of its nice properties come from. Are there any deep reasons for that? It appears that the answer is that the modal logic only takes some relevant part of information into account. Namely, some redundant part of the structure of what is analysed is skipped and only some interesting fragment of the data is kept.

This statement can be made precise when we introduce the notion of *relational semantics* for modal logic. As [3] highlights:

> 'Revolutionary' is an overused word, but no other word adequately describes the impact relational semantics [...] had on the study of modal logic.

Indeed, it turned modal logic – which up to that point involved only syntactic derivation systems with some intended interpretations – into a robust tool, general enough to talk about a large class of relational structures. But what are the relational semantics all about? Let us start with an example.

One way of explicating the notions of *necessity* and *possibility* is by appealing to so called *possible worlds*. In brief, one can say that φ is *possible* whenever *there exists a possible world* in which it holds. Dually, φ is *necessary* if it is true in *all the possible worlds*. But now let us make what we just said a bit more precise. In fact, we introduced – apart from the actual world – some possible worlds that we can *observe* from the actual one and which may differ from it. Of course, this relativises the notion of sentence truth, as given a sentence we have to specify which world is the *actual* one. Moreover, we could imagine that worlds we can see from the actual world p are not the same as the ones accessible from some q that can be

observed from p. Going back to logic, it simply means that we do not exclude *a priori* the possibility that the sentence 'It is possible that it is possible that φ.' is not equivalent to simple 'It is possible that φ.'
What is the essence of the example described above? We have some range of objects – in this case called *possible worlds* – together with an information about which atomic sentences (e.g. φ) hold in which object. Moreover, we also have an information about which objects can be seen from which ones – which can be just viewed as a relation linking the mentioned objects.

As we said, this structure (which is called a *modal model* and is formally defined in the Definition 2.1) turns out to be very universal and allows us to capture many different phenomena. It consists of: a set of *objects* or *points*; a *valuation* telling us which atomic propositions are true in which objects (these atomic propositions are supposed to represent all the information about a single point – all the facts about it that we take into account) and an *accessibility relation* representing which points 'can see' which points. This last component is perhaps the most abstract, as it may represent many different things (in the above example it was the relation of 'being possible from the perspective of').

Note that although some part of the information about such a modal model – e.g. the valuation – is crucial, there is a lot of redundancy there as well. First of all, we do not care what the underlying objects are, as long as it does not affect the structure (i.e. the valuation of objects and the accessibility relation linking them). Putting it more precisely, we do not want to distinguish two models that are *isomorphic*.

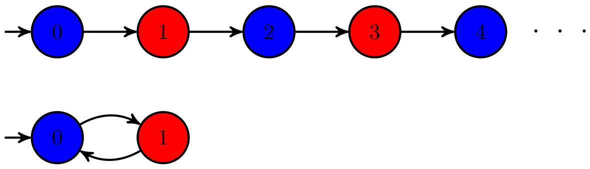
Identification of models that have the same structure – isomorphic models – is therefore something desirable. However, it may still be not enough. This is because a structural comparison requires an 'external', omniscient point of view which makes it somewhat strong. Let us illustrate it with an example of two models $\mathcal{M}$ and $\mathcal{M}'$ (see Example 2.5 for a formal description):

(Here the nodes, arrows between them and their colours represent (respectively): points, accessibility relation and the atomic propositions satisfied by a node. The short arrow distinguishes the *actual* node – the one in which we live).

On the one hand, it is clear that these models are not isomorphic, for $\mathcal{M}$ has three elements while $\mathcal{M}'$ has only two. However, if the only thing I can see is that the current node 0 is blue and the nodes accessible from 0: 1 and $1'$ (resp. 1 alone) are all red, I cannot really tell the difference between $\mathcal{M}$ and $\mathcal{M}'$. There is nothing that allows me to detect that 1 and $1'$ are actually *two different objects*, as long as they are *indistinguishable*.

The above example is special as in both considered models it is only possible to go at most one step from the initial node. Let us take a look at another example, where one can make an infinite number of steps from the current node.

Consider the two models $\mathcal{M}$ and $\mathcal{M}'$ (the convention is the same as with the previous picture; for a formal description see Example 2.4):



If I am observing the models from the perspective of the initial point 0, the only thing I can see is that the initial point 0 is blue, the point accessible from 0 is red, the one accessible from it is again blue and so on. As with the previous example, I cannot distinguish the two models because it would require some external means that would allow me to see that 2 and 0 are different objects *despite being indistinguishable.* Looking at the models from an 'internal' perspective that replaces the 'external', omniscient knowledge with an 'observation' made 'from inside', I do not have such means of comparison.

Let us now generalise the two above examples. If I am an agent inhabiting a point p_0 and all I can see is the information about p_0, points accessible from p_0, points accessible from points accessible from p_0 etc., I only have restricted means of distinguishing p_0 from other points. In fact, since the atomic propositions are supposed to represent all the facts and properties we take into account, the only way for me to see any difference between my p_0 and some other point p_1 is to observe that either:

- p_0 and p_1 satisfy different atomic propositions; or
- p_0 and p_1 differ with respect to the accessible points – i.e. there is some q_i accessible from p_i s.t. there is no q_{1-i} *indistinguishable* from q_i and accessible from p_{1-i}.

This statement is clearly circular. We define what it means that points are *indistinguishable* in terms of *indistinguishability* of the points accessible from them. However, it is a recursion rather than a vicious circle.[1] A relation of *observational* or *behavioural* equivalence is *any* relation that satisfies the two conditions: equivalent points p_0 and p_1 satisfy the same atomic sentences and for any point q_i accessible from p_i there exists a corresponding point q_{1-i} accessible from p_{1-i} that is equivalent to q_i.

This kind of relation is called a *bisimulation* relation (see the Definition 2.6 for a formal description). In many contexts it is more natural to ask about behavioural rather than structural equivalence. In such contexts an isomorphism is too restrictive and the actual relaxed equivalence we are interested in is precisely the bisimulation.

The concept of a bisimulation was introduced by van Benthem in order to characterise the expressive power of modal logic – which turned out to be precisely the fragment of first order logic invariant under bisimulation (see Theorem 2.21). It was also designed independently by D. Park in the context of computer science (see [15]). However, as further research revealed, the use of this notion is not restricted to the two mentioned contexts. For instance, as Janin & Walukiewicz showed (see [10]), an extension of modal logic – called the *modal μ-calculus*

[1] More precisely, it is more of a fixpoint than a simple recursion.

– also turned out to be closely related to bisimulation, as it is precisely the bisimulation invariant fragment of MSO[2]. Another example could be the fact that bisimulation can also be used to characterise equivalence of automata on finite words ([17]) and infinite ones ([5]).

Given how natural the notion of a bisimulation is and how closely it is related to modal logic, it is also extremely natural to ask when modal logic can describe a model *uniquely*. That is, given a set of modal sentences, when does it happen that it has a unique model *up to bisimulation*? A reader familiar with the model theory of first order logic may realise that this question is in fact analogous to the notion of *categoricity*, the property of having a unique model *up to isomorphism*. Although some problems of that sort were already investigated, it is surprising that this question was never asked before explicitly and only some partial answers are present. It is a classical result of Hennessy & Milner (Theorem 3.4) that if two models are finitely branching (i.e. every point has at most finitely many points accessible from it) and they satisfy the same modal sentences, they must be bisimilar. A strengthening of this theorem, which appears to be a folklore result, tells us that it suffices that just *one* of the models is finitely-branching. Another extension of the Hennessy–Milner theorem is that in fact it suffices that the considered models are in some sense saturated, which is a bit weaker requirement than finite branching (see Definition 4.9 for more details). Interestingly, these two extensions cannot be achieved at once: it does not suffice that a theory has *at least one saturated model* in order to have a unique model up to bisimulation, as every satisfiable set of modal sentences is satisfied in a saturated model (Proposition 4.15).

The contribution of this thesis is therefore twofold. First, most importantly, we state and prove a characterisation that can be seen as a completion of the Hennessy–Milner theorem: a complete modal theory has a unique model up to bisimulation iff all its models are bisimilar to finitely branching ones iff it has a model bisimilar to a finitely branching one (Theorem 3.9). The really new implication is the one telling us that if a modal theory has a model $\mathcal{M}$ that is not bisimilar to a finitely branching one, then it has another model $\mathcal{M}'$ which is not bisimilar to $\mathcal{M}$.
Second, we give a clean, uniform presentation of the known results – which adapts some of the known concepts (e.g. the Canonical Model – see Definition 4.14, or the modal version of saturation – see Definition 4.9), but is to some extent independent. In particular, the idea of interpreting the set of logical types as a topological space is well known in the context of classical first order logic, but introducing it to the modal logic appears to be a new idea that allows us to simplify proofs and see our problems from a uniform, high-level perspective.

The thesis is organised as follows. We start with this introduction. Then, in Chapter 1 we explain the necessary basics of modal logic – which could be skipped by a reader already familiar with it. Next, in Chapter 2 we discuss the main topic of this thesis, namely the problem of uniqueness of a model of a modal theory. We also state our Main Theorem (Theorem 3.9). Chapter 4 is devoted to the proof of the Main Theorem which consists of two parts: introduction of proof-specific tools and the actual proof. Chapter 5 contains a brief sketch of the possible further investigations and their limitations. Apart from this, we also include a short appendix containing remarks on the notation we use.

[2] MSO – the *monadic second order* logic – is the fragment of second order logic where second order quantification is restricted to sets (i.e. quantification over relations of arity higher than 1 is excluded).

Chapter 2

Standard definitions and facts

This section contains the very basic definitions and facts that can be skipped by anyone familiar with modal logic. The reader may find more details in a nice introductory handbook [3].

2.1. Models

For the rest of this thesis, let us fix an arbitrary set of *atomic propositions* Σ, whose elements will be usually denoted by $a_1, a_2, ...$ and $b_1, b_2, ...$ (or sometimes just a, b).

A modal model is just a directed graph together with a colouring of its nodes. Since in this context we are particularly interested in logic, instead of assigning each node a particular *colour*, we say which – possibly many – *atomic propositions* (from Σ) this node satisfies.
This is a subtle difference. On the one hand, one could always treat the set of propositions (an element of $\mathcal{P}(\Sigma)$) satisfied in a node as a single colour of that node – and in this sense the two approaches are equivalent.
However, there are reasons to choose the first representation to be the primitive one, as it is more natural for logic. To see this, assume that Σ is infinite and imagine we define a logic capable of describing whether a node has a particular colour. Then, the set of statements $\Phi \overset{df}{=} \{$"The node p does not have colour c."$\mid c \in \mathcal{P}(\Sigma)\}$ could not be satisfied (as every node has to have *some* colour), but every finite fragment $\Phi_{\text{fin}} \subseteq_{\text{fin}} \Phi$ would remain satisfiable, as Φ_{fin} only excludes a finite number of colours. Therefore such logic would not be compact.

Formally, a model consists of an arbitrary universe, a binary accessibility relation linking its elements and a valuation telling us, for each node, which atomic propositions are true in that node.

Definition 2.1. A *model* $\mathcal{M}$ consists of:

- A universe $U_{\mathcal{M}}$, which is an arbitrary set. We will call the elements of the universe *points* and typically denote them by p, q, r, s and o.

- A binary *accessibility* relation $R_{\mathcal{M}} \subseteq U_{\mathcal{M}} \times U_{\mathcal{M}}$. Given $pR_{\mathcal{M}}q$ we will say that p is a *predecessor* of q and q is a *son*[1] or *successor* of p. For convenience, we will often encode

[1] Note that despite the terminology we use, we do not assume that the model has to be a tree. That is, we do not exclude the possibility that a point has more than one predecessor. However, as we will soon show (in Proposition 2.10) , every model is in some sense *equivalent* (i.e. bisimilar) to a tree.

the accessibility relation $R_{\mathcal{M}}$ as a function $f_{\mathcal{M}} : U_{\mathcal{M}} \to \mathcal{P}(U_{\mathcal{M}})$ returning all sons of a given point, i.e. $f_{\mathcal{M}}(p) \stackrel{df}{=} \{q \mid pR_{\mathcal{M}}q\}$.

- A valuation $V_{\mathcal{M}} : \Sigma \to \mathcal{P}(U_{\mathcal{M}})$, for every atomic proposition a determining the set $V_{\mathcal{M}}(a)$ of all the points where it *holds* (or *is satisfied*). Similarly to the previous case, it will be convenient to encode the valuation as a *colouring* map $c_{\mathcal{M}} : U_{\mathcal{M}} \to \mathcal{P}(\Sigma)$ s.t. $c_{\mathcal{M}}(p) \stackrel{df}{=} \{a \in \Sigma \mid p \in V_{\mathcal{M}}(a)\}$ returning all the atomic propositions satisfied at a point. We will call $c_{\mathcal{M}}(p) \in \mathcal{P}(\Sigma)$ the *colour* of p.

A *pointed model* is a model $\mathcal{M}$ together with a point $p \in U_{\mathcal{M}}$ (called the *root*). Following the standards of modal logic notation, we will skip the parentheses and simply denote a pointed model (with a model $\mathcal{M}$ and a point p) by $\mathcal{M}, p$ instead of $(\mathcal{M}, p)$. ◁

If the model is clear from the context, we will skip the subscripts and simply write U, R, V and c. We will also abuse notation and write $q \in \mathcal{M}$ instead of $q \in U_{\mathcal{M}}$.

We define the restriction of a model in an obvious way.

Definition 2.2. Given a model $\mathcal{M} = (U, R, V)$ we define its *restriction* to the set $U' \subseteq U$, denoted $\mathcal{M}_{|_{U'}}$, to be the model $\mathcal{M}_{|_{U'}} \stackrel{df}{=} (U', R', V')$ where R' is just R restricted to $U' \times U'$ and $V'(a) \stackrel{df}{=} V(a) \cap U'$ for every a.

We call a model $\mathcal{M}'$ a *submodel* of $\mathcal{M}$ if it is a restriction of $\mathcal{M}$ to $U_{\mathcal{M}'}$. ◁

Definition 2.3. Given a model $\mathcal{M}$ and a point $p \in \mathcal{M}$, the set of points *reachable* from p is the least set containing p and closed under the accessibility relation $R_{\mathcal{M}}$.

A submodel $\mathcal{M}'$ of $\mathcal{M}$ generated by p – denoted $<p>_{\mathcal{M}}$ – is the model consisting of all the points reachable from p.

A pointed model $\mathcal{M}, p$ is *reachable* iff any point is reachable from the root, i.e. $\mathcal{M}, p = <p>_{\mathcal{M}}, p$. ◁

As with the other definitions, we will skip the subscripts and simply write $<p>$ instead of $<p>_{\mathcal{M}}$ if the model is clear from the context.

2.2. Bisimulations

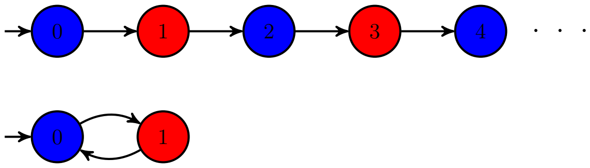
We are now ready to introduce the key notion – a bisimulation relation – which is an equivalence relation being the counterpart of isomorphism in the modal world.

As we explained in the introduction, bisimulation is a very natural notion capturing behavioural, rather than structural equivalence. It can be thought of as expressing the fact that two models are indistinguishable from the internal perspective – although they may have different structure when compared from an external point of view, models behave in exactly the same way.

This approach can be thought of in analogy to the well-known notion of isomorphism, where we ignore some redundant information (e.g. what is the underlying set, does the universe contain any singletons etc.) and only take structural information into account.

Let us take a look at the example we mentioned in the introduction:

Example 2.4. Consider the two models $\mathcal{M}$ and $\mathcal{M}'$:



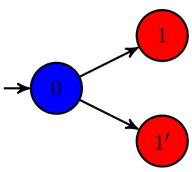
Recall that the arrows represent the accessibility relation and the colours – different colours of the nodes (in the above example blue represents the entire Σ and red – the empty set $\emptyset$). In both cases we chose 0 to be the root.

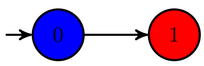
Formally,
$\mathcal{M} \overset{df}{=} (\omega, R_{\mathcal{M}}, V_{\mathcal{M}})$ where $f_{\mathcal{M}}(n) \overset{df}{=} \{n+1\}, V_{\mathcal{M}}(a) \overset{df}{=} \{2n \mid n \in \omega\}$ for all $a \in \Sigma$ and
$\mathcal{M}' \overset{df}{=} (\{0,1\}, R_{\mathcal{M}'}, V_{\mathcal{M}'})$ s.t. $f_{\mathcal{M}'}(n) \overset{df}{=} \{n+1 \mod 2\}, V_{\mathcal{M}'}(a) \overset{df}{=} \{0\}$ for all $a \in \Sigma$.

Although different, the two above pointed models – $\mathcal{M}, 0$ and $\mathcal{M}', 0$ – cannot be distinguished by an agent inhabiting them. The only things that could be observed 'from inside' are: at the current point (i.e. 0) all the Σ is true, in the successor of the current point no atomic proposition is true, in the successor of the successor of the current point again all the atomic propositions are satisfied – and so on. The only way we could distinguish both models is to use some notion of *equivalence* that does not hold between some points, so that we could say that 1 and 3 are not equivalent and thus they are actually different objects despite their identical behaviour. As long as we do not want to introduce any external means of comparison, it is not possible to distinguish $\mathcal{M}$ from $\mathcal{M}'$, for we could identify all the even points with 0 and all the odd points with 1.

The other example from the introduction, illustrating things we would *not* want to distinguish, is even simpler:

Example 2.5. Consider models $\mathcal{M}$ and $\mathcal{M}'$:





Formally, let $\mathcal{M} \overset{df}{=} (\{0,1,1'\}, R_{\mathcal{M}}, V)$ – a model consisting of a point and two its sons, where $f_{\mathcal{M}}(0) \overset{df}{=} \{1,1'\}$ and $f_{\mathcal{M}}(1) \overset{df}{=} f_{\mathcal{M}}(1') \overset{df}{=} \emptyset$; $V_{\mathcal{M}} \overset{df}{=} \lambda a.\{0\}$. Then, we may take its submodel $\mathcal{M}'$ with only one son: $\mathcal{M}' \overset{df}{=} \mathcal{M}_{|_{\{0,1\}}}$.

If we choose 0 to be the initial point, again no behavioural difference between the two models can be observed. Although in $\mathcal{M}$ the root has two sons while in $\mathcal{M}'$ it has only one, these sons (1 and $1'$) are indistinguishable and could be identified – and hence we cannot really tell what is the *number* of them.

Note that the fact that one of our models contained exactly *two* equivalent points does not matter at all. In particular, we could have any – finite or infinite – number of them as well.

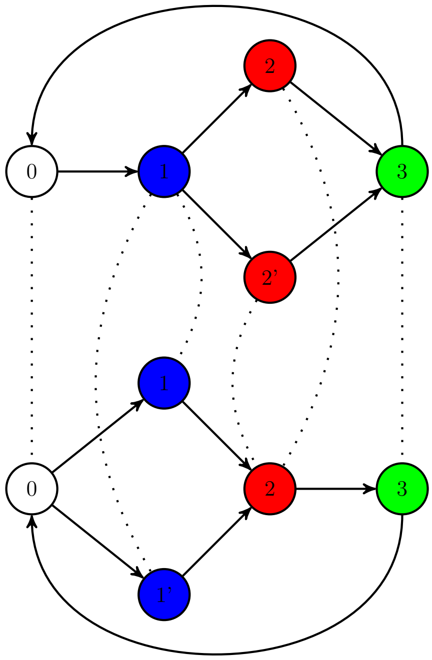
With these two examples in mind, we may now define the appropriate equivalence relation capturing our intuitions – a bisimulation relation.

Definition 2.6. Given two models $\mathcal{M}, \mathcal{M}'$ we call a (non-empty) relation $Z \subseteq U_{\mathcal{M}} \times U_{\mathcal{M}'}$ a *bisimulation* $\overset{df}{\iff}$ Z satisfies three conditions (so-called base, forth and back). Given p, p' s.t. pZp':

- (base) Related points p and p' satisfy the same atomic propositions, i.e. $c_{\mathcal{M}}(p) = c_{\mathcal{M}'}(p')$.
- (forth) For every successor of p, $q \in f_{\mathcal{M}}(p)$, there exists a successor of p', $q' \in f_{\mathcal{M}'}(p')$, s.t. qZq'.
- (back) For every successor of p', $q' \in f_{\mathcal{M}'}(p')$, there exists a successor of p, $q \in f_{\mathcal{M}}(p)$ s.t. qZq'.

Points $p \in \mathcal{M}, p' \in \mathcal{M}'$ are bisimilar (denoted $\mathcal{M}, p \leftrightarroweq \mathcal{M}', p'$) $\overset{df}{\iff}$ there exists a bisimulation Z between $\mathcal{M}$ and $\mathcal{M}'$ linking p and p' (the relation $\leftrightarroweq$ is called *bisimilarity*). $\lhd$

Just a note – bisimulation is defined to be a relation satisfying certain properties. The reader could wonder if it is possible to reformulate this definition in terms of functions from one model to another. In fact such a function – i.e. function whose graph is a bisimulation – is sometimes called a *bounded morphism* or a *p-morphism*. However, unlike in the case of isomorphism, it is not true that every bisimulation is a graph of some function – as it is easy to come up with an example of $\mathcal{M}$ and $\mathcal{M}'$ where we need to link two elements from $\mathcal{M}$ with one element of $\mathcal{M}'$ and the other way round *at the same time*. The following picture presents such a situation (dashed lines represent the only possible bisimulation):



There are many natural examples of properties that are invariant under bisimulation, e.g.: satisfying $a \in \Sigma$; having a son satisfying a and another one not satisfying a; well-foundedness. On the other hand, non-examples could be: having two different sons; having an even number of them (see Example 2.5) or infiniteness (Example 2.4).

Let us now recall some basic constructions and properties of bisimulation.

First, let us observe that the definition of a bisimulation does not assume that the considered models are actually different. Therefore it also makes sense to talk about bisimulation *inside* a single model. This allows us to take a quotient of a model by bisimulation, where equivalence classes satisfy atomic propositions iff any (equivalently: all) its members do and two classes are related iff they have any related members.

Definition 2.7. Given a model $\mathcal{M}$ and a bisimulation $Z \subseteq U_{\mathcal{M}} \times U_{\mathcal{M}}$, we define the quotient $\mathcal{M}_{/Z}$ to be the following model:

- a universe $U_{\mathcal{M}_{/Z}} \stackrel{\mathit{df}}{=} (U_{\mathcal{M}})_{/Z}$ consisting of the equivalence classes of Z;
- an accessibility function given as $f_{\mathcal{M}_{/Z}}([p]_{/Z}) \stackrel{\mathit{df}}{=} \{[q]_{/Z} \mid \exists_{r \in [p]_{/Z}} \exists_{s \in [q]_{/Z}}\ r R_{\mathcal{M}} s\}$;
- a valuation given by $c_{\mathcal{M}_{/Z}}([p]_{/Z}) \stackrel{\mathit{df}}{=} c_{\mathcal{M}}(p)$. $\lhd$

Note that the valuation is well defined, since any two points can only be bisimilar if they have equal colour. It is also not hard to see that $\mathcal{M}, p \leftrightarroweq \mathcal{M}_{/Z}, [p]_{/Z}$. Indeed, it is straightforward to check that the graph of the natural projection map $q \stackrel{\pi}{\mapsto} [q]_{/Z}$ (that is, the relation $\{(q, \pi(q)) \mid q \in \mathcal{M}\}$) is a bisimulation.

Bisimulations are closed under unions – as all the conditions are preserved. Therefore, given two bisimilar models $\mathcal{M}$ and $\mathcal{M}'$, one can always take *the greatest* bisimulation Z_{max} – i.e. the sum of all bisimulations between $\mathcal{M}$ and $\mathcal{M}'$ – which is the same as the *bisimilarity* relation $\leftrightarrows$ (i.e. $\mathcal{M}, p \leftrightarrows \mathcal{M}', p' \iff pZ_{max}p'$ for any $p \in \mathcal{M}$ and $p' \in \mathcal{M}'$). In particular, one can always consider the greatest bisimulation between a model and itself. If we consider only reachable models, then the quotient model is in fact minimal:

Proposition 2.8. *For any class $\mathbb{C}$ of pairwise bisimilar reachable pointed models there exists a model $\mathcal{M}_{\mathbb{C}}, p_{\mathbb{C}}$ which is minimal in the sense that for any other $(\mathcal{M}, p) \in \mathbb{C}$ there is a unique bounded morphism from $\mathcal{M}, p$ to $\mathcal{M}_{\mathbb{C}}, p_{\mathbb{C}}$.*

Another basic yet important property of bisimulations is that every model is bisimilar to a tree. This is witnessed by the *tree unravelling* of a model – i.e. the tree of all paths from the original model. Informally, we start from the root and for any point we replace its successors with their fresh copies. It can be formalised in the following way:

Definition 2.9. Given a pointed model $\mathcal{M}, p$ its *tree unravelling* is a model $\mathcal{M}', p'$ defined as follows:

We first take $\mathcal{M}'' \stackrel{df}{=} ((U_{\mathcal{M}})^+, R_{\mathcal{M}''}, V_{\mathcal{M}''})$ where

- $(U_{\mathcal{M}})^+$ is the set of all non-empty sequences over $U_{\mathcal{M}}$.
- $R_{\mathcal{M}''}$ is given[2] by $f_{\mathcal{M}''}(w \cdot q) \stackrel{df}{=} w \cdot q \cdot f_{\mathcal{M}}(q)$.
- $c_{\mathcal{M}''} \stackrel{df}{=} c_{\mathcal{M}} \circ \pi_{last}$ (here π_{last} is a function that returns the last element of a sequence).

Then we define $\mathcal{M}'$ to be the part of $\mathcal{M}''$ reachable from p: $\mathcal{M}', p' \stackrel{df}{=} <p>_{\mathcal{M}'', p}, p$. $\lhd$

Note that in $\mathcal{M}''$ there could not be two different points with a common successor, hence it is a forest. Therefore, as a reachable subset of a forest, $\mathcal{M}'$ is a tree.

On the other hand, π_{last}, the projection on the last element (or, more precisely, the relation $\{(w, \pi_{last}(w)) |\ w \in U_{\mathcal{M}}\}$) is a bisimulation. Thus:

Proposition 2.10. *Given a model $\mathcal{M}, p$ and its tree unravelling $\mathcal{M}', p'$, we have that $\mathcal{M}, p \leftrightarrows \mathcal{M}', p'$. In particular, every model is bisimilar to a tree.*

The two models defined in Example 2.4 could serve as an example, as the infinite model can be seen as the tree unravelling of the two-element one.

Another obvious construction is the disjoint union of models.

Definition 2.11. Given a family of models $\{\mathcal{M}_i \mid i \in I\}$ we define its disjoint union $\mathcal{M} \stackrel{df}{=} \bigsqcup_{i \in I} \mathcal{M}_i$ in a natural way:

- $U_{\mathcal{M}} \stackrel{df}{=} \bigcup_{i \in I} U_{\mathcal{M}_i} \times \{i\}$;

[2] Note that '$\cdot$' is our notation for concatenation – including concatenation of sequences with sets of sequences (see the Appendix for more details). Also note that w is an element of $(U_{\mathcal{M}})^*$ – in particular, it might be the empty sequence.

- $f_{\mathcal{M}}(q,i) \stackrel{df}{=} f_{\mathcal{M}_i}(q) \times \{i\}$;
- $c_{\mathcal{M}}(p,i) \stackrel{df}{=} c_{\mathcal{M}_i}(p)$.

As long as it does not cause ambiguity, we will abuse notation and simply refer to the elements of the disjoint union as to their projections to the first coordinate, i.e. instead of (q,i) we will just write q. We will also denote the disjoint union of two models $\mathcal{M}, \mathcal{N}$ by $\mathcal{M} \sqcup \mathcal{N}$ instead of $\bigsqcup\{\mathcal{M}, \mathcal{N}\}$. ◁

A reader familiar with category theory may check that this construction is just the coproduct in the category of modal models with bounded morphisms (i.e. functions whose graphs are bisimulations) as morphisms.

2.3. Modal logic

There are number of formalisms that can describe properties invariant under bisimulation (various examples of logics, automata or games). One natural, classical example is the *modal logic* – an extension of propositional logic by *modal operators* $\Diamond$ and its dual $\Box$.

Definition 2.12. *Modal formulae* (or *modal sentences*) are given by the following grammar:

$$S \to \neg S \mid S \vee S \mid \Diamond S \mid a \text{ for } a \in \Sigma.$$

Note that in order to be fully precise we should include parentheses in the definition. However, we skip it as it would only unnecessarily complicate the otherwise clear idea.

We will denote the set of all modal formulae by ML.

We can also define other connectives ('$\Rightarrow$' and '$\wedge$') in a standard way. A box '$\Box$' is a shorthand for '$\neg\Diamond\neg$'. ◁

Just a note for a reader not familiar with the concept of a grammar: ML is just the least set containing Σ and closed under prefixing a given word with '$\neg$' and '$\Diamond$'; and joining two words with '$\vee$' ('$\Diamond$' is a special (unary) modal operator that will be discussed in detail later).

Moreover, analogously to the concept of a quantifier depth, one can introduce a modal depth of a formula – the maximal number of nested modal operators:

Definition 2.13. *A modal depth* of a formula $md : \text{ML} \to \mathbb{N}$ is defined inductively as follows:

- $md(a) \stackrel{df}{=} 0$ for all $a \in \Sigma$;
- $md(\neg\varphi) \stackrel{df}{=} md(\varphi)$;
- $md(\varphi \vee \psi) \stackrel{df}{=} max(md(\varphi), md(\psi))$;
- $md(\Diamond\varphi) \stackrel{df}{=} md(\varphi) + 1$. ◁

We can now inductively define when a point in a model satisfies a modal formula. For atomic sentences it is straightforward – a point p satisfies $a \in \Sigma$ iff $p \in V(a)$. For connectives the definition is the same as for classical propositional logic – e.g. p satisfies $\varphi \vee \psi$ iff p satisfies either φ or ψ. In the case of diamonds (or boxes), our formula $\Diamond\varphi$ (resp. $\Box\varphi$) is satisfied at a point p iff φ holds at some son of p (resp. all p's sons). That is:

Definition 2.14. Given a pointed model $\mathcal{M}, p$, we say that a formula φ is *satisfied* (or *true*) at p (notation $\mathcal{M}, p \models \varphi$) $\overset{df}{\iff}$

- φ is an atomic proposition $a \in \Sigma$ and $p \in V(a)$; or
- φ is of the form $\neg\psi$ and ψ is not satisfied at the current point: $\mathcal{M}, p \not\models \psi$; or
- φ is of the form $\psi \vee \rho$ and one of the disjuncts is satisfied at the current point: $\mathcal{M}, p \models \psi$ or $\mathcal{M}, p \models \rho$; or
- φ is of the form $\Diamond\psi$ and there exists a successor of p where ψ holds: there is some $q \in \mathcal{M}$ s.t. pRq and $\mathcal{M}, q \models \psi$.

The notion of satisfaction can be easily lifted from formulae to sets of formulae, i.e. for $t \subseteq$ ML and a pointed model $\mathcal{M}, p$, we say that $\mathcal{M}, p$ satisfies t iff $\mathcal{M}, p$ satisfies all its members ($\mathcal{M}, p \models t \overset{df}{\iff} \forall_{\varphi \in t}\ \mathcal{M}, p \models \varphi$).

Given two models $\mathcal{M}$ and $\mathcal{N}$ we say that points $p \in \mathcal{M}$ and $q \in \mathcal{N}$ are *modally equivalent* or simply *equivalent* (notation: $\mathcal{M}, p \equiv \mathcal{N}, q$) iff they satisfy exactly the same modal formulae. We say they are *modally n-equivalent* or just *n-equivalent* iff they satisfy the same sentences of *modal depth up to n* (notation: $\mathcal{M}, p \equiv_n \mathcal{N}, q$). $\lhd$

As in the context of many other logics, we may define a modal *theory* or *type* – a maximal consistent set of modal formulae.

Definition 2.15. We call a set of modal formulae $t \subseteq$ ML a *modal theory* (or *modal type*) $\overset{df}{\iff}$ t is a maximal consistent[3] set of modal formulae, i.e. it is consistent and for every $\varphi \in$ML, either φ or $\neg\varphi$ belongs to t.

We denote the set of all modal types by $\mathbb{T}$.

Given a point in a model $p \in \mathcal{M}$ its *theory* or *type* $\mathrm{tp}_{\mathcal{M}}(p)$ is the set of modal sentences it satisfies: $\mathrm{tp}_{\mathcal{M}}(p) \overset{df}{=} \{\varphi \in \mathrm{ML} \mid \mathcal{M}, p \models \varphi\}$. Of course, every type is a type of some point in some model. As with the other notions, we will abuse notation and just write tp instead of $\mathrm{tp}_{\mathcal{M}}$ as long as it does not cause ambiguity.

It will be also useful to have a short notation for the set of all theories of sons of a point. Thus, let us define: $\mathbb{T}_{\mathcal{M},p} = \{\mathrm{tp}_{\mathcal{M}}(q) \in \mathbb{T} \mid q \in f_{\mathcal{M}}(p)\}$. As with the other definitions, we will usually write $\mathbb{T}_p$ skipping the model implicitly clear from the context.

[3] In some contexts, it makes sense to distinguish two notions - *syntactic* and *semantic* consistency - i.e. lack of a proof of contradiction and existence of a model, respectively. However, in our investigations we do not consider any notion of proof. Therefore, whenever we refer to *consistency*, semantic consistency is meant.

Note that all the above definitions can be restricted to modal formulae *of a fixed modal depth*. Hence modal n-type is a maximal consistent set of modal formulae *of modal depth at most* n. We will denote the set of all modal n-types by $\mathbb{T}^{(n)}$.

We will also use a convenient notation from the theory of formal languages and write ${}_{\Diamond\backslash}t$ for the left quotient of t with $\Diamond$ (that is: ${}_{\Diamond\backslash}t \stackrel{df}{=} \{\varphi \mid \Diamond\varphi \in t\}$ or, equivalently, ${}_{\Diamond\backslash}t = \bigcup \mathbb{T}_{\mathcal{M},p}$ where $\mathcal{M}, p$ is a model of t). ◁

Note that although the quotient ${}_{\Diamond\backslash}t$ of a theory $t \in \mathbb{T}$ does not have to be consistent, it is true that it is *complete*, i.e.:

Proposition 2.16. *For a quotient* ${}_{\Diamond\backslash}t$ *of a theory* $t \in \mathbb{T}$ *exactly one of the following holds:*

- *For any* $\varphi \in ML$, *either* φ *or* $\neg\varphi$ *belongs to* ${}_{\Diamond\backslash}t$.
- *The quotient* ${}_{\Diamond\backslash}t$ *is empty.*

Proof. By definition, t is a theory, so in particular it is consistent – i.e. it has a model $\mathcal{M}, p$. The first item corresponds to the case when p has at least one son, the second one – to the other case. □

The following fact is rather straightforward, but deserves being spelled out explicitly:

Proposition 2.17. *A modal* n*-type* $t \in \mathbb{T}^{(n)}$ *is uniquely determined by the following information:*

- $t \cap \Sigma$, *i.e. which atomic propositions are satisfied in the root;*
- ${}_{\Diamond\backslash}t$, *i.e. which sentences of modal depth* $\leq (n-1)$ *are satisfied by the root's sons.*

As a consequence, a modal theory $t \in \mathbb{T}$ *is uniquely determined by the information:*

- $t \cap \Sigma$, *i.e. which atomic propositions are satisfied in the root;*
- ${}_{\Diamond\backslash}t$, *i.e. which sentences are satisfied by the root's sons.*

Proof. It can be shown by straightforward induction on the complexity of formulae that for any $\varphi \in$ ML, $t \cap \Sigma = t' \cap \Sigma$ and ${}_{\Diamond\backslash}t = {}_{\Diamond\backslash}t'$ together imply $\varphi \in t \iff \varphi \in t'$ (which is the same as $t = t'$). □

Another observation is that if Σ is finite, then there are only finitely many n-types. This also implies that every n-type is equivalent to a single formula.

Proposition 2.18. *Assume* $|\Sigma| < \infty$. *Then* $\mathbb{T}^{(n)}$ *is finite. As a consequence, every* n*-type* $t \in \mathbb{T}^{(n)}$ *is equivalent to a single formula* ψ_t *of modal depth at most* n.

Proof. Given Proposition 2.17, the proof of finiteness of $\mathbb{T}^{(n)}$ is a straightforward induction on n.
To find a formula equivalent to an n-type $t \in \mathbb{T}^{(n)}$, observe that for any other n-type $t' \in \mathbb{T}^{(n)}$, there must be a formula $\varphi_{t,t'}$ of a modal depth at most n s.t. $\varphi_{t,t'} \in t$ but $\varphi_{t,t'} \notin t'$. Therefore, since $\mathbb{T}^{(n)}$ is finite, we can write down a formula $\psi_t \stackrel{df}{=} \bigwedge_{t' \in \mathbb{T}^{(n)}} \varphi_{t,t'}$ that is equivalent to t. □

Note that the two above propositions also easily imply the finite model property of ML:

Proposition 2.19. *If a formula $\varphi \in ML$ has a model, then it has a finite one.*

Proof. First, using the two previous propositions, we prove by induction on n that if $|\Sigma| < \infty$ then every n-type $t \in \mathbb{T}^{(n)}$ has a finite model. It is obvious for $n = 0$, as it suffices to take a model consisting of a root of appropriate colour $t \cap \Sigma$. For $n + 1$, we construct a model $\mathcal{M}, p$ consisting of (i) a root p (of appropriate colour $t \cap \Sigma$) linked to (ii) a model $\mathcal{N}_i, s_i \models t_i$ for every n-type $t_i \subseteq {}_{\Diamond\backslash}t$. By Proposition 2.17, $\mathcal{M}, p \models t$. On the other hand, by Proposition 2.18, there are only finitely many t_i's and each of them is equivalent to some formula ψ_{t_i} of modal depth at most n – and hence, by the induction hypothesis, we may assume that every $\mathcal{N}_i$ is finite.

Since every $\varphi \in$ ML uses only finitely many atomic propositions $a_1, ..., a_k \in \Sigma$, we may consider φ as a formula of modal logic ML$'$ over a restricted set of atomic sentences, $\Sigma' = \{a_1, ..., a_k\}$. Thus, since φ is an element of some n-type t, it has a finite model (in the sense of the restricted set Σ') – let us call it $\mathcal{M}$. Now, it suffices to extend $\mathcal{M}$ into a model in the sense of our initial Σ by simply defining all the atomic propositions $b \in \Sigma - \{a_1, ..., a_k\}$ to be false everywhere. $\square$

It is not hard to see – using straightforward induction on the complexity of a formula – that modal logic is invariant under bisimulation: any two bisimilar models satisfy exactly the same modal formulae.

On the other hand, modal logic can be easily encoded in first order logic FO with a signature σ consisting of a unary predicate P_a for every atomic proposition $a \in \Sigma$ and a single binary relational symbol for the accessibility relation.

Definition 2.20. Given a formula $\varphi \in$ ML, we recursively define its *standard translation* $\mathrm{ST}_x(\varphi)$ as follows:

- $\mathrm{ST}_x(a) \stackrel{df}{=} P_a(x)$;
- $\mathrm{ST}_x(\neg\varphi) \stackrel{df}{=} \neg\mathrm{ST}_x(\varphi)$;
- $\mathrm{ST}_x(\varphi \vee \psi) \stackrel{df}{=} \mathrm{ST}_x(\varphi) \vee \mathrm{ST}_x(\psi)$;
- $\mathrm{ST}_x(\Diamond\varphi) \stackrel{df}{=} \exists_y.xRy \wedge \mathrm{ST}_y(\varphi)$. $\lhd$

It follows directly from the definition of the semantics of ML that for any $\varphi \in$ ML, we have $\mathcal{M}, p \models \varphi \iff \mathrm{ST}_x(\varphi)$ is satisfied by p in $\mathcal{M}$ (here $\mathcal{M}$ is viewed as a model for FO over the signature σ in a natural way, i.e. $P_a(x)$ holds iff $x \in V(a)$ and $R_{\mathcal{M}}$ is the interpretation of the relational symbol R). Hence any modal sentence is equivalent to a bisimulation-invariant first order formula (i.e. an FO formula $\psi(x)$ with one free variable x s.t. $\mathcal{M}, p \leftrightarroweq \mathcal{M}', p'$ implies that $\psi(p)$ holds iff $\psi(p')$ holds). However, it turns out that the implication can be switched – due to van Benthem ([2]) we have a beautiful characterisation of the modal logic as precisely the bisimulation invariant fragment of first order logic (FO):

Theorem 2.21 (van Benthem)**.** *Modal logic ML is the bisimulation invariant fragment of first order logic, i.e. for any $\varphi \in FO$, φ is invariant under bisimulation iff φ is equivalent to $ST_x(\psi)$ for some $\psi \in ML$.*

The way we introduced semantics for modal logic is straightforward. However, it can be rephrased in a more algebraic fashion which is elegant and can easily lead to some natural extensions of modal logic.

Definition 2.22. Given a model $\mathcal{M}$, we inductively define its *semantics* $v_{\mathcal{M}} : \mathrm{ML} \to \mathcal{P}(U)$ as follows:

- $v_{\mathcal{M}}(a) \stackrel{df}{=} V(a)$ for $a \in \Sigma$;
- $v_{\mathcal{M}}(\neg\varphi) \stackrel{df}{=} U - v_{\mathcal{M}}(\varphi)$;
- $v_{\mathcal{M}}(\varphi \vee \psi) \stackrel{df}{=} v_{\mathcal{M}}(\varphi) \cup v_{\mathcal{M}}(\psi)$;
- $v_{\mathcal{M}}(\Diamond\varphi) \stackrel{df}{=} \{p \mid f_{\mathcal{M}}(p) \cap v_{\mathcal{M}}(\varphi) \neq \emptyset\}$.

It can be shown by a straightforward induction that the two definitions are in fact equivalent – that is, a formula φ is satisfied at a point $p \iff p$ belongs to the φ's semantics: $\mathcal{M}, p \models \varphi \iff p \in v_{\mathcal{M}}(\varphi)$.

The notion of semantics can be also lifted in a natural way to entire theories – semantics of a theory t consists of the points that satisfy it: $v_{\mathcal{M}}(t) \stackrel{df}{=} \bigcap_{\varphi \in t} v_{\mathcal{M}}(\varphi)$. ◁

Before we proceed further, let us introduce a few more concepts and facts that can be helpful in order to better understand the relation between modal logic and bisimilarity.

First, it is useful to introduce another approach to bisimilarity – the game approach. As it often happens in the context of equivalence relations, the bisimilarity can be described in terms of a game played between two players – Adam and Eve (also called Spoiler and Duplicator, Existential and Universal player, or simply $\forall$ and $\exists$) – the first one trying to show that considered structures differ, the second one – that they are the same.

Definition 2.23. Given two pointed models $\mathcal{M}_0, p_0$ and $\mathcal{M}_1, p_1$ we define the *bisimilarity game* $\mathcal{G}(\mathcal{M}_0, p_0, \mathcal{M}_1, p_1)$ played between two players, Adam and Eve. The game starts at the position (p_0, p_1) – the two initial points – and continues in three subsequent phases:

- First, atomic propositions are checked for consistency (that is, it is checked whether: $c_{\mathcal{M}_0}(p_0) = c_{\mathcal{M}_1}(p_1)$). If the points do not agree on the atomic propositions, Eve immediately looses.
- Second, Adam picks a successor of either p_0 or p_1 – i.e. some $p_i' \in f_{\mathcal{M}_i}(p_i)$.
- Finally, Eve has to respond with a successor of the other point $p_{1-i}' \in f_{\mathcal{M}_{1-i}}(p_{1-i})$.

After that, the game continues starting from the position (p_0', p_1'). If any of the players is stuck (i.e. has no moves), he/she looses immediately. In the case of an infinite play, Eve wins. ◁

It is straightforward to check that this game captures the notion of bisimilarity – as its definition reflects the three conditions from the definition of bisimilarity, with phase (1) corresponding to the base condition and subsequent phases (2) and (3) – to back and forth conditions (depending on the coordinate – first or second – on which Adam moves).

Proposition 2.24. *Given two pointed models $\mathcal{M}, p$ and $\mathcal{M}', p'$, the following are equivalent:*

- $\mathcal{M}, p \leftrightarroweq \mathcal{M}', p'$;
- *Eve has a winning strategy in $\mathcal{G}(\mathcal{M}, p, \mathcal{M}', p')$.*

Another useful notion is a partial bisimulation. The notion of a bisimulation is somewhat strong and can be seen as a modal counterpart of isomorphism. However, one could consider a weaker relation (or – to be more precise – a family of relations) – so-called n-bisimulation – a bisimulation restricted to a fixed depth – which is an approximation of the full bisimulation relation. This relation can be defined inductively with n playing a role of a parameter.

Definition 2.25. Given two pointed models $\mathcal{M}, r$ and $\mathcal{M}', r'$, we say they are *n-bisimilar* (denoted $\mathcal{M}, r \leftrightarroweq_n \mathcal{M}', r'$) $\overset{df}{\iff}$ there exists a family of relations $Z_1, ..., Z_n \subseteq U_{\mathcal{M}} \times U_{\mathcal{M}'}$ s.t. (i) rZ_nr' and (ii) for every Z_k and points p, p' s.t. pZ_kp':

- (base) Related points satisfy the same atomic propositions, i.e. $c_{\mathcal{M}}(p) = c_{\mathcal{M}'}(p')$.
- (forth) If $0 < k$, then for every successor of p, $q \in f_{\mathcal{M}}(p)$ there exists a successor of p', $q' \in f_{\mathcal{M}'}(p')$ s.t. $qZ_{k-1}q'$.
- (back) If $0 < k$, then for every successor of p', $q' \in f_{\mathcal{M}'}(p')$ there exists a successor of p, $q \in f_{\mathcal{M}}(p)$ s.t. $qZ_{k-1}q'$. $\lhd$

As well as the full bisimilarity, the n-bisimilarity has a nice definition in terms of games. One can show by a simple inductive argument that n-bisimilarity corresponds to the n-round bisimilarity game – the bisimilarity game restricted to n rounds. On the other hand, under some additional conditions, partial bisimilarity can be related to logic – even more directly than the full one. The n-bisimilarity plays a role analogous to the role of partial isomorphism for the first order logic where the notion of Ehrenfeucht-Fraïsse game captures partial FO types. Indeed, this resemblance turns out to be deep as the correspondence between finite games and partial equivalence transfers to the world of bisimulation-invariant properties.

Proposition 2.26. *Given pointed models $\mathcal{M}, p$ and $\mathcal{M}', p'$, the following are equivalent:*

- $\mathcal{M}, p \leftrightarroweq_n \mathcal{M}', p'$;
- *Eve has a winning strategy in $\mathcal{G}_n(\mathcal{M}, p, \mathcal{M}', p')$, the bisimilarity game restricted to n steps after which Eve automatically wins.*[4]

Moreover, if additionally $|\Sigma| < \infty$, then the two above items are both equivalent to the third one (if Σ is infinite, then it is strictly weaker):

- $\mathcal{M}, p \equiv_n \mathcal{M}', p'$.

Proof. Equivalence of the first two items, as well as the implication from any of them to the third one can be shown by straightforward induction on n.

[4] More precisely, the game consists of n-rounds of the bisimilarity game (Def 2.24), after which consistency of atomic propositions is checked once again, so that in her last move Eve cannot cheat by responding with a point of a different colour than the one picked by Adam.

The place where finiteness of Σ comes into play is the remaining implication from the last item to one of the previous two. The assumption that $|\Sigma| < \infty$ allows us to prove by induction on n that $\mathcal{M}, p \equiv_n \mathcal{M}', p'$ implies $\mathcal{M}, p \leftrightarrows_n \mathcal{M}', p'$.
For n=0, it is immediate.
For $n+1$, let us take $(n+1)$-equivalent points q_0 and q_1 and any son r_i of q_i – we will show that q_{1-i} has to have a son r_{1-i} that is n-bisimilar to r_i.
Let us consider the n-type $t \in \mathbb{T}^{(n)}$ of r_i. Since Σ is finite, t is (by Proposition 2.18) equivalent to a single formula φ_t of modal depth at most n. This means that q_i satisfies $\Diamond\varphi_t$ – and by $(n+1)$-equivalence of q_0 and q_1 – also q_{1-i} has to satisfy it. However, this means that some son r_{1-i} of q_{1-i} satisfies φ_t and therefore is n-equivalent to r_i. But by the induction hypothesis, n-equivalence implies n-bisimilarity, which finishes the proof. $\square$

Chapter 3

Bisimulational categoricity

In the context of first order logic there is a well-studied notion of categoricity – the property of having a unique model up to isomorphism. Due to the classic result of Skolem and Löwenheim ([18]), this notion is nontrivial only in a relativised form.

Theorem 3.1 (Skolem–Löwenheim)**.** *If a first order theory t over a signature σ has an infinite model, then t has a model of every infinite cardinality $\kappa \geq |\sigma|$.*

Since, by definition, an isomorphism is a bijection satisfying certain properties, the only way for a first order theory to have a unique model is when it has a unique finite model. In fact, it is easy to check that if the signature σ is finite, then the complete theory of a finite model is *always* categorical.

Because of these limitations, the question about model uniqueness is relativised to a fixed cardinality – instead of asking when a theory t has a unique model *in general*, we ask when it has a model that is unique *among models of cardinality κ.*

Since modal logic ignores some structural aspects of the described model, it does not make sense to ask when a *modal* theory has a unique model up to isomorphism – as it is never the case. A modal theory t has always a number of non-isomorphic models, e.g. taking a reachable model of t of size λ and the disjoint union of its two copies we get two models of t s.t. only the first one is reachable (here λ is the reader's favourite cardinal with the restriction that $\lambda \geq \max(\omega, |\Sigma|)$).

However, as we argued in the introduction, in the study of modal models it often makes sense to look at the models up to *bisimilarity*, not up to *isomorphism.* This gives rise to the question when a modal theory has a unique model *up to bisimulation.* In fact, in the modal world this question is even more natural, as unlike isomorphism, bisimulation can relate models of different cardinalities. This is because our relaxed notion of equivalence does not assume that the structures of models are exactly the same. Instead, as explained in the introduction, we only look at the important part of the information – the information about model's behaviour. If we ignore the redundant part of data, it makes perfect sense to ask when models technically having *different* structure are actually *equivalent.* And, as it will turn out, this is not a mere hopeful hypothesis, but an actual situation – we are able to study bisimulational categoricity as a well-behaved autonomous notion independent of the full structural categoricity.

A notion that appears to be the closest to our investigations is the one of a *Hennessy-Milner class of models* – introduced by Golblatt in [7]. We say that a class of modal models has the Hennessy–Milner property iff within that class modal equivalence implies bisimilarity. However, it turns out that although both underlying tools and notions had been developed to some extent, our questions were not posed before. Hence, before we move further, let us name our key notion, which was already described – the *bisimulational categoricity*:

Definition 3.2. A set of formulae t is *bisimulationally categorical* $\overset{df}{\iff}$ t has a unique model up to bisimulation (i.e. all its models are bisimilar). $\lhd$

The notion of bisimulational categoricity has at least two advantages over its possible rephrasing in terms of the mentioned Hennessy–Milner property. First, it is more loosely connected to modal logic – and even to logic at all. It is natural to ask about modal categoricity for other bisimulation-invariant logics (both theories and single sentences) and formalisms recognising models. Second, it resembles analogous notion of categoricity for non-bisimulation-invariant logic (e.g. FO) – and as such also easily leads to further extensions, where we investigate when a formalism invariant under a chosen equivalence relation $\sim$ can characterise a model uniquely *up to that equivalence relation* $\sim$. We will say something about instances of these two problems at the end of this thesis.

Before we move on to the harder question about categoricity of entire *theories*, let us start with a warm-up and say something about categoricity of *single modal formulae.* If Σ is infinite, then there is no hope for any kind of categoricity here, as a single formula can only capture a finite number of atomic propositions. However, if we assume finiteness of Σ, we can easily determine which models can be characterised by a single formula.

Proposition 3.3. *If $|\Sigma| < \infty$, then given a pointed model $\mathcal{M}, p$, the following are equivalent:*

- *$\mathcal{M}, p$ can be characterised by a single modal sentence (i.e. there is $\varphi \in ML$ s.t. $\mathcal{M}, p \models \varphi$ and $\{\varphi\}$ is bisimulationally categorical).*
- *$\mathcal{M}, p$ is bisimilar to a finite tree.*

Proof. If $\mathcal{M}, p$ is a finite tree, then there is a bound n on the maximal depth of a path in it. However, this means that n-step bisimilarity also implies the full one, as no play longer than n is possible in the game $\mathcal{G}_n(\mathcal{M}, p, \mathcal{N}, q)$ for any $\mathcal{N}, q$. However, by Proposition 2.26, $\leftrightarroweq_n$ is the same as $\equiv_n$, and thus the n-type t of p characterises $\mathcal{M}, p$. But by Proposition 2.18, t is equivalent to a single formula, which finishes the proof.

For the other direction it suffices to prove that any consistent modal formula has a model that is a finite tree. Recall that any modal formula has a finite model (Proposition 2.19), so taking its tree unravelling we get a finitely branching tree satisfying φ. On the other hand, given the game-theoretic characterisation of n-bisimilarity, it is easy to see that a tree restricted to the points at depth at most n is n-bisimilar to the original model (as, again, no play longer that n is possible anyway in the bisimilarity game). But a finitely branching tree of a fixed depth is of course finite. $\square$

Let us now return to our main question: when a modal theory is bisimulationally categorical? A partial answer is given as a reformulation of the Hennessy–Milner theorem which states that finitely branching models have the Hennessy–Milner property – they are bisimilar iff they are modally equivalent ([9]).

Theorem 3.4 (Hennessy-Milner)**.** *Given models $\mathcal{M}, p$ and $\mathcal{M}', p'$ that are finitely branching (i.e. $f_{\mathcal{M}}(q)$ and $f_{\mathcal{M}'}(q')$ are finite for all q, q'), we have that if $\mathcal{M}, p \equiv \mathcal{M}', p'$ then also $\mathcal{M}, p \leftrightarrows \mathcal{M}', p'$.*

A strengthened version of the Hennessy–Milner theorem, which appears to be a folklore result, tells us a bit more: whenever two models are modally equivalent and *at least one of them* is finitely branching, they are bisimilar. Both theorems can be easily proved in an elementary way. However, the framework we develop here will allow us to give an elegant, high-level argument for both – the basic and the strengthened version of the theorem.

On the other hand, there exist non-bisimilar but modally equivalent models. A simple set-theoretic argument for this uses the fact that there are $2^{|\Sigma|}$ many modal theories but the collection of all models (up to bisimulation) is a proper class.

Proposition 3.5. *The collection of all the pointed models – counted up to bisimulation – cannot be represented as a set.*

This fact is witnessed by the following example:

Example 3.6. We define a family of pairwise non-bisimilar models by enriching ordinals with a model structure.

Given an ordinal κ, we define a pointed model $\mathcal{M}_\kappa, p_\kappa$ as follows:

- $U_{\mathcal{M}_\kappa} \stackrel{df}{=} \kappa + 1$;
- $R_{\mathcal{M}_\kappa} \stackrel{df}{=} >$ (i.e. $\alpha R \beta \iff \alpha > \beta$ with the standard ordinals ordering);
- $V_{\mathcal{M}_\kappa} \stackrel{df}{=} \lambda a.\emptyset$.

The root p_κ is just κ.

It is not hard to see that $\mathcal{M}_\alpha, p_\alpha \not\leftrightarrows \mathcal{M}_\beta, p_\beta$ for $\alpha \neq \beta$. Indeed, it can be shown by straightforward transfinite induction that if $\alpha < \beta$, then p_β has a son that cannot be bisimilar to any p_α's sons.

What is worth pointing out is that the family of models defined above is well-founded as no ordinal contains an infinite descending chain. It also does not use any atomic propositions, as they are all false everywhere. Therefore any restriction on these properties cannot be sufficient for categoricity.
Although the above construction is simple, the argument based on cardinality is not informative at all. However, it is not hard to find actual examples of infinitely branching models that are modally equivalent but not bisimilar. Below we present our canonical example which is – as it will follow from our further investigations – as simple as possible (for it has a trivial colouring and contains only a single point with infinitely many sons).

Example 3.7. Consider a model $\mathcal{H}$ – a 'hedgehog' consisting of a root and branches of any length rooted in it (with a trivial valuation):

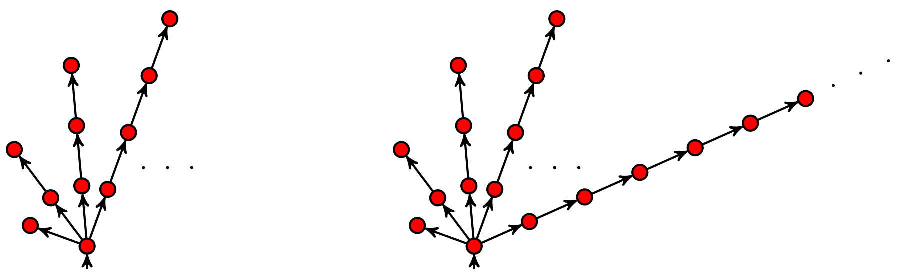
- $U_{\mathcal{H}} \overset{df}{=} \{(x,y) \in \omega \times \omega \mid y \leq x\} \cup \{root\}$;
- $f_{\mathcal{H}}(root) \overset{df}{=} \omega \times \{0\}$ and $f_{\mathcal{H}}(x,y) \overset{df}{=} \begin{cases} \{(x,y+1)\} & \text{if } y < x \\ \emptyset & \text{otherwise} \end{cases}$;
- $V_{\mathcal{H}} \overset{df}{=} \lambda a.\emptyset$.

An extension of the original hedgehog, a 'hedgehog with a horn' $\mathcal{H}'$, can be constructed by adding a 'limit spike' – an infinite branch – to the original hedgehog:

- $U_{\mathcal{H}'} \overset{df}{=} \{(x,y) \in (\omega+1) \times \omega \mid y \leq x\} \cup \{root'\}$;
- $f_{\mathcal{H}'}(root') \overset{df}{=} (\omega+1) \times \{0\}$ and $f_{\mathcal{H}'}(x,y) \overset{df}{=} \begin{cases} \{(x,y+1)\} & \text{if } y < x \\ \emptyset & \text{otherwise} \end{cases}$;
- $V_{\mathcal{H}'} \overset{df}{=} \lambda a.\emptyset$.

Of course, in these definitions we assume that both *root* and *root*$'$ are fresh elements distinct from $(\omega+1) \times \omega$. ◁

A picture of $\mathcal{H}, root$ and $\mathcal{H}', root'$:



On the one hand, the two above models are modally equivalent, i.e. $\mathcal{H}, root \equiv \mathcal{H}', root'$. To see this, by Proposition 2.26, it suffices to provide a strategy that guarantees Eve a victory in an n-round bisimilarity game for every n. Each such a strategy will imply n-equivalence, and hence, altogether, they will imply full equivalence.
Note that the only non-trivial case in the bisimulational game is when Adam picks the infinite spike in his first move – as if he picks a spike of length k, Eve can respond with a spike of the same length in the other model, which is not only bisimilar, but isomorphic to the chosen one. However, the duration of the game – n – is known in advance. Therefore, if Adam chooses the infinite spike, it suffices for Eve to respond with a spike that is long enough, i.e. longer than n. In fact, one can show – e.g. using Ehrenfeucht-Fraïsse games – that the two models cannot be distinguished even if we use the full expressive power of first order logic.[1]

[1] It is interesting that, in a very loose sense, a linear search is the best that modal logic can do – whereas the limits of first order logic are given by a binary search.

On the other hand, the two models cannot be bisimilar, as Adam can win the full, infinite bisimilarity game by just picking the infinite spike and then following the only possible path until Eve is stuck.

This example is an instance of an important phenomenon. It turns out that the situation of the hedgehog is not an accident, but an instance of a general rule. Namely, modal logic cannot capture the existence of something that is a *limit* of other things already present in the model. The intuition underlying this observation will be made precise and lead us to a nice characterisation of bisimulational categoricity.
We will formulate and prove a theorem that can be seen as a converse and completion of the Hennessy–Milner theorem: a modal theory is bisimulationally categorical iff all its models are finitely branching (up to bisimulation) iff it has at least one finitely branching model.

Let us point out a subtle hack regarding the notion of finite-branching *up to bisimulation*. Although a natural definition of finite branching could be as simple as saying that no point has infinitely many pairwise non-bisimilar sons, there are technical reasons for which we will use an alternative definition – which essentially tells us the same, but technically is not equivalent. We will say that a model is finitely branching up to bisimulation if it is bisimilar to a finitely branching one.
If we only take reachable models into account – i.e. models, where every point is accessible from the initial one – then the two definitions coincide: being bisimilar to a finite branching model is the same as not having a point with infinitely many pairwise non-bisimilar sons (because every two bisimilar points always have the same number of pairwise non-bisimilar sons). However, if we allow for non-accessible points, it could happen that some models are finitely branching in the first, but not in the second sense.

An approach alternative to the one we choose could be to consider only reachable models. This does not seem to be a serious restriction, as every model is bisimilar to its reachable part and hence the non-reachable part is entirely redundant from the point of view of any bisimulation invariant semantics. The fact that we allow for non-reachable points in a model is rather a technical trick than something that captures any deep ideas. However, there are two reasons for our choice: (i) in the literature it is standard to allow for non-reachable points, (ii) it makes some definitions simpler.

Definition 3.8. A model $\mathcal{M}$ is *finitely branching* $\overset{df}{\iff}$ every point has finitely many sons. It is κ-branching $\overset{df}{\iff}$ every point in $\mathcal{M}$ has at most κ many sons.
A pointed model $\mathcal{M}, p$ is finitely branching (or κ-branching) *up to bisimulation* $\overset{df}{\iff}$ it is bisimilar to a finitely branching (resp. κ-branching) model. It is infinitely branching *up to bisimulation* $\overset{df}{\iff}$ it is not bisimilar to a finitely branching model. $\triangleleft$

Note that the last item – i.e. the property of being infinitely branching up to bisimulation – is not just being bisimilar to an infinitely branching model – as trivially any model has that property (recall Example 2.5). Instead, we are interested in the property complementary to being finitely branching up to bisimulation.

Now we are ready to formulate our main result.

Theorem 3.9 (The main theorem)**.** *Let $t \in ML$ be a complete modal theory. The following are equivalent:*

1. *The theory t is bisimulationally categorical.*
2. *Every model of t is finitely branching up to bisimulation.*
3. *There exists a finitely branching model of t.*

The entirely new implication is the one from (1) to (2). The opposite one, i.e. the one from (2) to (1) is the classical result of Hennessy and Milner (however, the proof we present differs from the original one) and the implication from (3) to (1) is its strengthened folklore version.

Chapter 4

Proof of the main theorem

4.1. Proof-specific definitions and facts

Before we proceed with the proof of the main theorem 3.9, let us introduce a few more definitions and facts that will simplify our investigations. Not all of them are completely new: the notion of modal saturation (sometimes called 'm-saturation') has been introduced in [6], the Canonical Model seems to appear for the first time in [13] and the idea of equipping the space of logical types with a topological structure is a standard concept in the context of FO. However, these notions are less popular than the very basic notions like bisimulation and exceed the elementary basics of the theory of modal logic.

Let us start with a rather technical notion of *non-redundancy* which can be seen as a kind of normal form for models.

Definition 4.1. A model $\mathcal{M}$ is called *non-redundant* $\overset{df}{\iff}$ no point $p \in \mathcal{M}$ has two bisimilar sons. $\triangleleft$

It is easy to see that every model is bisimilar to some non-redundant tree.

Proposition 4.2. *Every pointed model is bisimilar to a non-redundant tree.*

Proof. Given a model $\mathcal{M}, p$ it suffices to first take its quotient by bisimilarity $\mathcal{M}_{/\leftrightarrows}, [p]_{/\leftrightarrows}$ (Definition 2.7) and then unravel it to a tree $\mathcal{M}', p'$ (Definition 2.9). By definition , in $\mathcal{M}_{/\leftrightarrows}$ no two points are bisimilar. On the other hand, it is easy to see that the construction for model unravelling that we gave does preserve non-redundancy. Indeed, $w \cdot q \in \mathcal{M}'$ (i.e. a sequence with q as the last element) is bisimilar to $q \in \mathcal{M}_{/\leftrightarrows}$ and therefore, if a sequence $w \cdot q$ had two bisimilar sons $\mathcal{M}', w \cdot q \cdot r \leftrightarrows \mathcal{M}', w \cdot q \cdot r'$, it would imply that in $\mathcal{M}_{/\leftrightarrows}$ two different sons of q – namely r and r' – are bisimilar, which contradicts the definition of a quotient. $\square$

For the sake of completeness let us define formally what we mean by a substitution of models.

Definition 4.3. Given two models $\mathcal{M}$ and $\mathcal{N}$ with two points $p \in \mathcal{M}$ and $q \in \mathcal{N}$ we define the *substitution* $\mathcal{M}[p \leftarrow \mathcal{N}, q]$ to be the result of replacing p with $(\mathcal{N}, q)$ in $\mathcal{M}$. Formally, if we denote $\mathcal{M}' \stackrel{df}{=} \mathcal{M}[p \leftarrow \mathcal{N}, q]$, then:

- $U_{\mathcal{M}'} \stackrel{df}{=} U_{\mathcal{M} \sqcup \mathcal{N}} - \{p\}$;
- $f_{\mathcal{M}'}(s) \stackrel{df}{=} \begin{cases} (f_{\mathcal{M} \sqcup \mathcal{N}}(s) - \{p\}) \cup \{q\} & \text{if } f_{\mathcal{M} \sqcup \mathcal{N}}(s) \text{ contains } p \\ f_{\mathcal{M} \sqcup \mathcal{N}}(s) & \text{otherwise} \end{cases}$;
- $V_{\mathcal{M}'} \stackrel{df}{=} V_{\mathcal{M} \sqcup \mathcal{N}}$. ◁

In words: we simply take the disjoint union of $\mathcal{M}$ and $\mathcal{N}$ (excluding the point p) and then change the accessibility relation in a way such that any point previously related to p is now related to q instead (note that p's successors do not became q's successors though).

It is not hard to see that both bisimilarity and modal equivalence are congruences for substitutions, i.e. substituting bisimilar (resp. modally equivalent) models yields bisimilar (resp. modally equivalent) results:

Proposition 4.4. *Bisimilarity and modal equivalence are both congruences of substitutions of pointed models. That is:*

- *Given pointed models* $\mathcal{N}_0, q_0 \leftrightarrows \mathcal{N}_1, q_1$ *and* $\mathcal{M}$ *together with a point* $p \in \mathcal{M}$ *we have:* $\mathcal{M}[p \leftarrow \mathcal{N}_0, q_0], r \leftrightarrows \mathcal{M}[p \leftarrow \mathcal{N}_1, q_1], r$ *for any* $r \in \mathcal{M} - \{p\}$.
- *Given pointed models* $\mathcal{N}_0, q_0 \equiv \mathcal{N}_1, q_1$ *and* $\mathcal{M}$ *together with a point* $p \in \mathcal{M}$ *we have:* $\mathcal{M}[p \leftarrow \mathcal{N}_0, q_0], r \equiv \mathcal{M}[p \leftarrow \mathcal{N}_1, q_1], r$ *for any* $r \in \mathcal{M} - \{p\}$.

Proof. First, for bisimilarity: given the game-theoretic characterisation of bisimilarity (Proposition 2.24) it suffices to provide a winning strategy for Eve. Indeed, she can win by just responding with the same node as played by Adam as long as the play stays in $\mathcal{M}$ and once (if ever) q_i is reached, she responds with q_{1-i} and continues with the winning strategy whose existence follows from the assumption that $\mathcal{N}_0, q_0 \leftrightarrows \mathcal{N}_1, q_1$.

For modal equivalence we will use induction on n to show that for any n and any $r \in \mathcal{M} - \{p\}$, we have $\mathcal{M}[p \leftarrow \mathcal{N}_0, q_0], r \equiv_n \mathcal{M}[p \leftarrow \mathcal{N}_1, q_1], r$.
The induction basis (i.e. $n = 0$) is immediate, as the valuation of r is the same as in $\mathcal{M}$. For the inductive step $(n+1)$, it suffices to prove that in both models r satisfies the same atomic propositions and that its sons satisfy the same sentences of modal depth n (Proposition 2.17). The first is again immediate. For the second, consider two cases: (i) q_i is not a son of r – then the proposition follows directly from the induction hypothesis; (ii) q_i is r's son – then observe that the full equivalence implies n-equivalence, which also allows us to apply the induction hypothesis to finish the proof. □

As in the context of other logics, ML can be seen from an elegant, topological perspective that allows us to avoid dirty details in the proofs. Let us introduce a topology on the set of all modal types $\mathbb{T}$ (which is essentially the same as the standard topology on first order types).

Definition 4.5. We define a topology on $\mathbb{T}$: the basic open sets are of the form $\mathcal{U}_\varphi \stackrel{df}{=} \{t \in \mathbb{T} \mid \varphi \in t\}$ for all $\varphi \in$ ML. ◁

Proposition 4.6. $\mathbb{T}$ *is a compact topological space.*

Proof. It is straightforward to check that $\mathbb{T}$ satisfies all the axioms of a topology. Closure under unions follows from the definition of a basis, closure under finite intersections follows from a simple observation that $\mathcal{U}_\varphi \cap \mathcal{U}_\psi = \mathcal{U}_{\varphi\wedge\psi}$. Of course, $\emptyset = \mathcal{U}_{\varphi\wedge\neg\varphi}$ and $\mathbb{T} = \mathcal{U}_{\varphi\vee\neg\varphi}$.

The non-trivial part is compactness – which is usually described in a slightly different manner by stating that if any finite fragment of a set of sentences is consistent, then the entire set is consistent as well. Let us call the two kinds of compactness *topological* and *logical*, respectively.

To see that the two notions are in fact equivalent, first suppose that there is an infinite open cover of $\mathbb{T}$ with no finite subcover. W.l.o.g. it consists of basic open sets – a family $\mathcal{B} \stackrel{df}{=} \{\mathcal{U}_{\varphi_i} \mid i \in I\}$. Since $\mathcal{B}$ has no finite subcover, no finite $\mathcal{B}_{I_0} \stackrel{df}{=} \{\mathcal{U}_{\varphi_i} \mid i \in I_0 \subseteq_{fin} I\} \subseteq_{fin} \mathcal{B}$ exhausts all the types and therefore $\{\neg\varphi_i \mid i \in I_0\}$ is consistent. Now – by logical compactness – we know that actually the entire $\{\neg\varphi_i \mid i \in I\}$ has to be consistent as well. However, this means that there is a type t that cannot belong to any $\mathcal{U}_{\varphi_i}$ – which contradicts the assumption that $\mathcal{B}$ is a cover of $\mathbb{T}$.

For the opposite direction, suppose towards contradiction that there is a set $t \subseteq \mathrm{ML}$ s.t. every its finite subset is consistent, but the entire t is not. We claim that the family $\mathcal{B} \stackrel{df}{=} \{\mathcal{U}_{\neg\varphi} \mid \varphi \in t\}$ is an open cover of $\mathbb{T}$ with no finite subcover.
Indeed, $\mathcal{B}$ covers the entire $\mathbb{T}$: if there were a type $t' \notin \bigcup\mathcal{B}$, it would not be contained in $\mathcal{U}_{\neg\varphi}$ for any $\varphi \in t$. But since t' is a type, $\varphi \notin t'$ is the same as $\neg\varphi \in t'$ and hence $t \subseteq t'$ – which cannot happen as t is inconsistent.
On the other hand, no finite subcover $\mathcal{B}_0 = \{\mathcal{U}_{\neg\varphi_1}, ..., \mathcal{U}_{\neg\varphi_k}\}$ of $\mathcal{B}$ could exist, as $\{\varphi_1, ..., \varphi_k\} \subseteq_{fin} t$ is consistent and thus there is a type $t \notin \bigcup\mathcal{B}_0$.

A simple high-level argument for the logical compactness is that ML is a fragment of FO which is compact and as such has to be compact itself.[1] □

Note that the space $\mathbb{T}$ is homeomorphic – via identification of a type t with its characteristic function χ_t – to a subset of the product space 2^{ML} (where 2 stands for the space $\{0,1\}$ with discrete topology). In particular, if Σ is at most countable, we can take an enumeration of ML and define a metric $d(t,t') \stackrel{df}{=} \frac{1}{k}$ where k is the first coordinate on which t and t' disagree. Although the metric itself depends on the order of ML, the resulting topology is always the same and it is just the one we defined.

Since we want to prove our theorem in full generality, we do not want to put any restrictions on Σ. However, it may help the reader actually think of the proofs we give as concerning this metric space.

[1] This argument can be presented in an even more elegant way, in strictly topological terms. We skip the details of this alternative presentation – however, it is worth to give a sketch.
Let us denote the set of all FO formulae with one free variable by $\mathrm{FO}(x)$. The space of all FO 1-types (i.e. maximal consistent subsets of $\mathrm{FO}(x)$, see [14] for more details) is known to be compact. On the other hand, every modal formula can be seen (via the standard translation – see Definition 2.20) as an FO formula with one free variable. Therefore, the set of all FO 1-types (viewed as a subset of $2^{\mathrm{FO}(x)}$, the product space of a two-point discrete space) can be projected on $\mathbb{T}$ (viewed as a subset of the product space 2^{ML}) in a natural way. It is easy to see that this map is continuous, and hence $\mathbb{T}$ is a continuous image of a compact space, which means that it is compact itself.

Let us also separate a simple topological fact which will be crucial in the proof of the main theorem – namely the existence of a *limit point* under certain conditions.

Definition 4.7. Given a subset X of a topological space Y we call a point $q \in Y$ a *limit point of* X $\overset{df}{\iff}$ $\overline{X - \{q\}} = \overline{X} = \overline{X \cup \{q\}}$. $\lhd$

Proposition 4.8. *Every infinite subset X of a compact topological space Y has a limit point. In particular, any infinite set of types has a limit point.*

Proof. Suppose towards contradiction that every point $p \in Y$ has a neighbourhood $\mathcal{U}_p$ disjoint with $X - \{p\}$. Then $\mathcal{B} \overset{df}{=} \{\mathcal{U}_p \mid p \in Y\}$ forms an open cover of Y with no finite subcover. It exhausts the entire Y as it contains a neighbourhood of every point, but on the other hand each $\mathcal{U}_p$ contains at most one point from X – and thus, by infiniteness of X, no finite subset of $\mathcal{B}$ covers Y. Thus, there must be a point q s.t. its every neighbourhood intersects X and so it is our desired limit point. □

We will use a notion of *saturation* that adapts similar ideas from the context of first order logic. The idea of saturation of a first order theory is not new (see [14] for more details). Also, its modal counterpart has been introduced earlier (for instance [16] uses it in an alternative proof of the van Benthem characterisation theorem 2.21). Note that the usual convention is to use the name 'modal saturation' or just 'm-saturation' in order to avoid confusion with saturation in the sense of first order logic (as for example in [8]). However, since the latter is not considered in this thesis, we simply use the name 'saturation' as it never causes ambiguity. It turns out that this notion suits our purposes very well and helps us develop a better understanding of the phenomena we investigate.

Definition 4.9. We call a point in a model $q \in \mathcal{M}$ *saturated* $\overset{df}{\iff}$ for every modal type $t \in \mathbb{T}$, if for every finite subset $t_{fin} \subseteq_{fin} t$, some son of q satisfies it, then some son of q satisfies the entire t.

We call a model saturated if all its points are saturated. $\lhd$

The hedgehog (Example 3.7) is an example of a model that is not saturated. More precisely, the root is the place where saturation fails: although every sentence of the form $\Diamond ... \Diamond \top$ is satisfied in some son of the root, there is no limit spike that could satisfy all these sentences at once (here '$\top$' stands for any tautology, e.g.: $a \vee \neg a$).
On the other hand, the hedgehog with a horn is an example of a saturated model – as it is the result of adding a son satisfying the only missing type to the original hedgehog.

Note that the definition of saturation only mentions a *complete* type t, but it could be equivalently replaced with an arbitrary set of modal formulae.

Proposition 4.10. *Given a model $\mathcal{M}$ and a point $p \in \mathcal{M}$, the following are equivalent:*

- *The point p is saturated.*
- *For any nonempty set of modal formulae $t \subseteq ML$, if for every finite subset $t_{fin} \subseteq_{fin} t$ some son of p satisfies t_{fin}, then some son of p satisfies the entire t.*

Proof. The bottom-up implication is obvious, as a modal theory is a special case of a set of formulae. For the other direction, observe that the requirement of satisfaction of finite fragments of t can be equivalently expressed by stating that t is closed under conjunctions and $t \subseteq \bigcup \mathbb{T}_p$.[2] Therefore, it suffices to show that $t \subseteq \bigcup \mathbb{T}_p$ can be extended to a complete theory $t' \subseteq \bigcup \mathbb{T}_p$. Given that, we can use saturation of p to find its son satisfying t' – and hence also t.

Note that the set of all consistent subsets of $\bigcup \mathbb{T}_p$ having t as a subset, that is $S \stackrel{df}{=} \{l \subseteq \bigcup \mathbb{T}_p \mid l \text{ is consistent and } t \subseteq l\}$, can be partially ordered by the inclusion relation '$\subseteq$'. Then, any ascending chain $(t_i)_{i \in I}$ has an upper bound in S. Indeed, a natural candidate for such a bound, the sum $\bigcup_{i \in I} t_i$, is a subset of $\bigcup \mathbb{T}_p$ and it is consistent by compactness of ML. Thus, $(S, \subseteq)$ satisfies the assumptions of the Kuratowski–Zorn lemma and hence it has a *maximal* element $t' \in S$.
By definition of S, we have that t' is consistent and $t \subseteq t'$, so to show that t' is in fact a theory it suffices to prove that it is complete. Let us take any $\varphi \in$ ML. Then, by consistency of t', either φ or $\neg\varphi$ is consistent with it. On the other hand, since $\bigcup \mathbb{T}_p = {}_{\diamond\backslash}\mathrm{tp}_{\mathcal{M}}(p)$, by Proposition 2.16 $\bigcup \mathbb{T}_p$ is complete (for it has a non-empty subset t and hence cannot be empty) – and thus either $t' \cup \{\varphi\}$ or $t' \cup \{\neg\varphi\}$ is an element of S. However, by maximality of t', this actually means that either $\varphi \in t'$ or $\neg\varphi \in t'$. □

Also, although the notion of saturation – as defined above – only concerns immediate successors, it can be easily shown that it actually extends to any finite number of steps from the current point. That is:

Definition 4.11. For a point $p_0 \in \mathcal{M}$, we call points accessible from p_0 in exactly n steps (i.e. points p_n s.t. there exists a path of length n in $\mathcal{M}$: $p_0 R p_1 R ... R p_{n-1} R p_n$) its n-descendants. We say that a point p is n-step saturated $\stackrel{df}{\iff}$ given a set of sentences $t \subseteq$ ML, if any finite $t_{\text{fin}} \subseteq_{\mathit{fin}} t$ is satisfied in some n-descendant of p, then p has an n-descendant that satisfies the entire t. Likewise, we call a model n-step saturated $\stackrel{df}{\iff}$ all its points are n-step saturated. ◁

It eventually turns out that the two notions of model saturation – standard and n-step saturation – are equivalent:

Proposition 4.12. *Given a model $\mathcal{M}$, the following are equivalent:*

- *For every n, $\mathcal{M}$ is n-step saturated.*
- *$\mathcal{M}$ is saturated.*

Proof. The top-down implication is obvious, as saturation is a special case of n-step saturation (namely: 1-step).

We will prove the bottom-up one by induction on n. For $n = 1$ it is obvious, as 1-step saturation is just the standard one (because by the previous Proposition 4.10, saturation with respect to *sets of sentences* is the same as saturation with respect to *theories*).

For $n+1$, assume that any finite subset $s \subseteq_{\mathit{fin}} t$ is satisfied in some $(n+1)$-descendant of p – call it q_s. Then, q_s's predecessors – in particular some p's n-descendant q'_s – satisfy $\diamond \bigwedge s$. Consider the set $l \stackrel{df}{=} \{\diamond \bigwedge s \mid s \subseteq_{\mathit{fin}} t\}$. Any finite subset $k \stackrel{df}{=} \{\diamond \bigwedge s_1, ..., \diamond \bigwedge s_j\} \subseteq_{\mathit{fin}} l$ is satisfied

[2] Recall that $\mathbb{T}_p$ is the set of all theories of sons of p (Definition 2.15).

in some n-descendant of p, for $\Diamond(\bigwedge_i \varphi_i)$ implies $\bigwedge_i \Diamond\varphi_i$ and thus $k = \{\Diamond \bigwedge s_1, ..., \Diamond \bigwedge s_j\}$ is satisfied in $q'_{\bigcup\{s_1,...,s_j\}}$. However, we may now use the induction hypothesis and conclude from the n-step saturation of $\mathcal{M}$ that some n-descendant of p satisfies the entire l – let us call that point r. Now, since r satisfies l, we get that every finite subset of our original theory t is satisfied in some son of r. But this finishes the proof, as by saturation some son of r (which means – $(n+1)$-descendant of p) satisfies the entire t. □

The next fact is an important link between the two notions we just introduced – topology on $\mathbb{T}$ and saturation.

Proposition 4.13. *Given a model and a point $p \in \mathcal{M}$, the two conditions are equivalent:*

- $\mathbb{T}_p$ *– the set of types of all p's sons – is closed in a topological sense (i.e. $\mathbb{T}_p = \overline{\mathbb{T}}_p$).*
- *The point p is saturated.*

Proof. For the top-down implication, suppose that $\mathbb{T}_p$ is closed and take a type $t \notin \mathbb{T}_p$. Then, by definition of a closure, there must be an open neighbourhood of t disjoint with $\mathbb{T}_p$, w.l.o.g. it is a basic open set – say $\mathcal{U}_\varphi$. But then no type in $\mathbb{T}_p$ contains φ and thus no son of p satisfies $\{\varphi\} \subseteq_{fin} t$.

For the other direction, suppose that p is saturated and take $t \in \overline{\mathbb{T}}_p$. For any finite $s \subseteq_{fin} t$, $\mathcal{U}_{\bigwedge s}$ is an open neighbourhood of t and thus has a non-empty intersection with $\mathbb{T}_p$. However, this means that some son of p satisfies s and therefore, by saturation of p, $t \in \mathbb{T}_p$. □

Let us now give an important example of a saturated model that is universal in a certain sense – the canonical model. It will be useful in our further investigations. It is a model whose universe is the set of all modal types. We will equip it with an accessibility relation and a valuation so that the model will be saturated.

Definition 4.14 (Canonical model)**.** The canonical model $\mathcal{M}_\mathbb{T}$ is defined as follows:

- $U_{\mathcal{M}_\mathbb{T}} \stackrel{df}{=} \mathbb{T}$;
- $f_{\mathcal{M}_\mathbb{T}}(t) \stackrel{df}{=} \{s \in \mathbb{T} \mid s \subseteq {}_\Diamond\backslash t\}$;[3]
- $V_{\mathcal{M}_\mathbb{T}}(a) \stackrel{df}{=} \{t \in \mathbb{T} \mid a \in t\}$. ◁

The relation between points of $\mathcal{M}_\mathbb{T}$ – which are themselves modal theories – and the theories they *satisfy* is as simple as one could imagine:

Proposition 4.15. *In $\mathcal{M}_\mathbb{T}$ every point (viewed as a theory) is equal to the set of sentences it satisfies, i.e. $tp_{\mathcal{M}_\mathbb{T}}(t) = t$. Moreover, $\mathcal{M}_\mathbb{T}$ is saturated. In particular, every theory has a saturated model.*

Note that since t is complete – contains any formula or its negation – the first part of this proposition can be equivalently expressed by saying that for any formula $\varphi \in$ ML, $\mathcal{M}_\mathbb{T}, t \models \varphi \iff \varphi \in t$.

[3] Recall that ${}_\Diamond\backslash t$ is the left quotient of t with $\Diamond$, i.e. ${}_\Diamond\backslash t = \{\varphi \mid \Diamond\varphi \in t\}$)

Proof. For the first part, we use induction on the complexity of formulae to prove that $\mathcal{M}_{\mathbb{T}}, t \models \varphi \iff \varphi \in t$. For atomic propositions it follows directly from the definition of $V_{\mathcal{M}_{\mathbb{T}}}$. For connectives $\vee$ and $\neg$ it follows from the fact that we consider *complete* theories – *maximal* consistent sets of formulae.

The only remaining case is $\Diamond$. Given a point $t \in \mathcal{M}_{\mathbb{T}}$ and a sentence $\Diamond\varphi$ we show that:

$$\Diamond\varphi \in t \iff t \text{ has a son satisfying } \varphi.$$

The right to left implication is easy. If t has a son – call it s – satisfying φ, then by induction hypothesis $\varphi \in s$. On the other hand by the definition of the accessibility relation $s \subseteq_{\Diamond\backslash} t$ and hence $\Diamond\varphi \in t$.
For the other direction, suppose $\Diamond\varphi \in t$. Since t is consistent, it has a model $\mathcal{M}, p$ where some son of p – call it q – satisfies φ. Then the type t' of q (i.e. $t' \stackrel{df}{=} \mathrm{tp}_{\mathcal{M}}(q)$) obviously contains φ – and hence, by the induction hypothesis, $\mathcal{M}_{\mathbb{T}}, t' \models \varphi$. On the other hand, since q is a son of p (in $\mathcal{M}$), it follows that $t' \subseteq_{\Diamond\backslash} t$, and thus (in $\mathcal{M}_{\mathbb{T}}$) t' is a son of t – which finishes the proof.

The only remaining thing is that the canonical model is saturated – which is rather straightforward. Suppose every finite subset $s_{fin} \subseteq s$ of a type $s \in \mathbb{T}$ is satisfied in some son of t. This means that $s \subseteq_{\Diamond\backslash} t$ and hence s is a son of t. On the other hand, by what we just proved, s (viewed as a point in $\mathcal{M}_{\mathbb{T}}$) satisfies s (viewed as a theory). □

4.2. Proof of the main theorem

Let us recall the main theorem (3.9) we are going to prove:

Theorem. *Let t be a complete modal theory. The following are equivalent:*

1. *The theory t is bisimulationally categorical.*

2. *Every model of t is finitely branching up to bisimulation.*

3. *There exists a finitely branching model of t.*

Note that we actually prove more implications than necessary to complete the proof of equivalence of the three items. However, this is the most natural way to prove it as the second item seems closer to the first and the third one than they are to each other.

4.2.1. (1) ⇒ (2)

We split the proof of this implication into two lemmata. The first one can be seen as a partial converse of Proposition 4.4.

Lemma 4.16. *Given a non-redundant reachable model $\mathcal{M}, r$ any point $p \in \mathcal{M} - \{r\}$ and $\mathcal{N}, q$ s.t. $\mathcal{N}, q \not\leftrightarrows <p>, p$ (or, equivalently, just $\mathcal{N}, q \not\leftrightarrows \mathcal{M}, p$), the substitution $\mathcal{M}' \stackrel{df}{=} \mathcal{M}[p \leftarrow \mathcal{N}, q]$, is not bisimilar to the original model: $\mathcal{M}, r \not\leftrightarrows \mathcal{M}', r$.* ◁

Proof. Suppose towards contradiction that $\mathcal{M}, r \leftrightarrows \mathcal{M}', r$. Since $\mathcal{M}, r$ is reachable, there is a path[4] from the root r to p. Thus, let us take such a path π that has the *minimal length.* Then, looking at the situation from the game-theoretic point of view, we have that the path $\pi = (r, ..., p) \in (U_{\mathcal{M}})^+$ has to be bisimilar (point-wise) to some path in $\mathcal{M}'$, $\pi' = (r, ..., p') \in (U_{\mathcal{M}'})^+$, consisting of Eve's responses to the moves of Adam picking consecutive points from π.

We will show that π' belongs not only to $(U_{\mathcal{M}'})^+$, but also to $(U_{\mathcal{M}})^+$. Note that by definition of a substitution (Definition 4.3), in $\mathcal{M}' = \mathcal{M}[p \leftarrow \mathcal{N}, q]$ no point of $\mathcal{N}$ is reachable from r without passing through q. Therefore, every path in $\mathcal{M}'$ that starts at r and does not contain q is actually also a path in $\mathcal{M}$, so it suffices to show that π' does not contain q.
Thus, let us suppose towards contradiction that π' *does* contain q. Then, π' is of the form $\sigma \cdot q \cdot \tau$ for some $\sigma, \tau \in (U_{\mathcal{M}'})^*$ s.t. q does not occur in σ.
Note that since $(r, ..., p') = \pi' = \sigma \cdot q \cdot \tau$, it follows that τ must be nonempty for otherwise we would have that $p' = q$, which is impossible as $\mathcal{M}', p' \leftrightarrows \mathcal{M}, p \not\leftrightarrows \mathcal{N}, q \leftrightarrows \mathcal{M}', q$ implies $p' \neq q$. Likewise, since $r \neq q$ (because $r \in \mathcal{M}$ while $q \in \mathcal{N}$), we have that σ must be nonempty as well.
By our choice of σ, it does not contain q so it belongs not only to $(U_{\mathcal{M}'})^+$, but also to $(U_{\mathcal{M}})^+$. However, $\sigma \cdot q$ is a path in $\mathcal{M}'$. Therefore, the last element of σ must be linked to q by $R_{\mathcal{M}'}$, and hence it must be linked to p by $R_{\mathcal{M}}$. Thus, $\sigma \cdot p$ is a path in $\mathcal{M}$.
On the other hand, π and π' have the same length and nonemptiness of τ implies that $\sigma \cdot q$ is a proper prefix of π', so $\sigma \cdot p$ is a path in $\mathcal{M}$ leading from r to p shorter than π, which contradicts the fact that π has a minimal length.

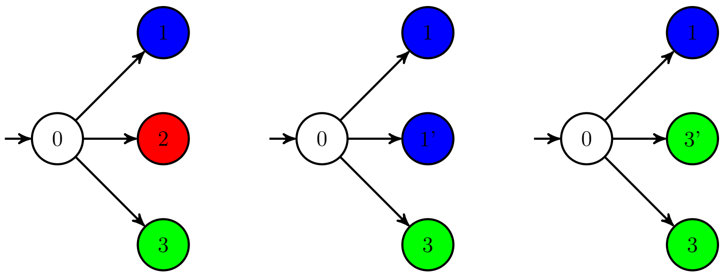
Having that $\pi' \in (U_{\mathcal{M}})^+$, the proof is rather straightforward. Both paths $\pi, \pi' \in (U_{\mathcal{M}})^+$ have a common root r, so we may consider the last point up to which they agree – call it s. Of course, $p \neq p'$, because $\pi' \in (U_{\mathcal{M}'})^+$ but $p \notin \mathcal{M}'$ – and thus $s \neq p$. However, this contradicts the assumption that our model was non-redundant, as s has two bisimilar sons – one belonging to π and the other to π'. □

Note that one could be tempted to try to actually prove a stronger statement and consider two substitutions *at once*, i.e. $\mathcal{M}[p \leftarrow \mathcal{N}_0, q_0]$ and $\mathcal{M}[p \leftarrow \mathcal{N}_1, q_1]$ for $\mathcal{N}_0, q_0 \not\leftrightarrows \mathcal{N}_1, q_1$. However, substitutions do not have to preserve non-redundancy (i.e. $\mathcal{M}[p \leftarrow \mathcal{N}, q]$ need not to be non-redundant even if both $\mathcal{M}$ and $\mathcal{N}$ are), and therefore it could happen that substituting non-bisimilar models to the same point gives us bisimilar results.

Example 4.17. Consider a model consisting of a root and its three pairwise non-bisimilar sons. The result of replacing the first one with the second one is bisimilar to the result of replacing it with the third one.

A picture presenting three models – an original one and two results of substituting non-bisimilar (blue and green, respectively) one-point models for the red node:

[4] In this context a path in a model $\mathcal{K}$ is just a *finite* sequence of its elements (i.e. an element of $(U_{\mathcal{K}})^*$ – the set of all finite sequences over $U_{\mathcal{K}}$, or $(U_{\mathcal{K}})^+$ – if we consider only *non-empty* paths) s.t. consecutive elements are related by $R_{\mathcal{K}}$. Therefore, we treat paths as usual sequences and denote their concatenation by '$\cdot$' (with the obvious restriction that for any $\sigma, \rho \in (U_{\mathcal{K}})^+$, $\sigma \cdot \rho$ is a path only if $R_{\mathcal{K}}$ relates the last element of σ to the first element of ρ).



Formally, $\mathcal{M} \overset{df}{=} (\{0,1,2,3\}, R, V)$ s.t. $f_{\mathcal{M}}(0) \overset{df}{=} \{1,2,3\}$ and $f_{\mathcal{M}}(i) \overset{df}{=} \emptyset$ for $i \neq 0$ and V is any valuation that distinguishes all the points so that they have different colours and hence cannot be bisimilar. Then, of course, $<1>, 1 \not\leftrightarrows <3>, 3$, but $\mathcal{M}[2 \leftarrow <1>, 1], 0 \leftrightarrows \mathcal{M}[2 \leftarrow <3>, 3], 0$, as they are both bisimilar to $\mathcal{M}$ restricted to $\{0,1,3\}$.

Fortunately, we do not need the strengthened (and false) version of the lemma, so the above example is not a problem.

The second lemma, which is the heart of the theorem, tells us that whenever a point has infinitely many sons, they have a limit that can be added or removed without changing the theory of the model.

Lemma 4.18. *For a given non-redundant model $\mathcal{M}, p$ s.t. p has infinitely many sons, there exists $\mathcal{M}', p'$ s.t. $\mathcal{M}, p \equiv \mathcal{M}', p'$ but $\mathcal{M}, p \not\leftrightarrows \mathcal{M}', p'$.* ◁

Proof. We look at the types of the sons of p: $\mathbb{T}_p$. W.l.o.g. $\mathbb{T}_p$ is infinite, for otherwise – by the pigeonhole principle – some theory $t \in \mathbb{T}_p$ would be realised in two non-bisimilar sons of p and then we could substitute one for another to obtain the desired $\mathcal{M}', p$ – which would not be bisimilar to the original model by previous Lemma (4.16), but would be equivalent by Proposition 4.4.

Since $\mathbb{T}_p$ is infinite, by Proposition 4.8, there must be a limit point t of $\mathbb{T}_p$.

We may define $\mathcal{M}', p'$ to be a model that only differs from $\mathcal{M}, p$ by presence/absence of p's son realising t. Let us take an arbitrary model of t – e.g. the canonical one $\mathcal{M}_{\mathbb{T}}, t$. We may now define $\mathcal{M}', p'$ as follows:

- $U_{\mathcal{M}'} \overset{df}{=} U_{\mathcal{M} \sqcup \mathcal{M}_{\mathbb{T}}}$;
- $f_{\mathcal{M}'}(s) \overset{df}{=} f_{\mathcal{M} \sqcup \mathcal{M}_{\mathbb{T}}}(s)$ for all $s \neq p$ and $f_{\mathcal{M}'}(p) \overset{df}{=} \begin{cases} f_{\mathcal{M} \sqcup \mathcal{M}_{\mathbb{T}}}(p) \cup \{t\} & \text{if } t \notin \mathbb{T}_{\mathcal{M},p} \\ f_{\mathcal{M} \sqcup \mathcal{M}_{\mathbb{T}}}(p) - v_{\mathcal{M} \sqcup \mathcal{M}_{\mathbb{T}}}(t) & \text{otherwise} \end{cases}$;
- $V_{\mathcal{M}'} \overset{df}{=} V_{\mathcal{M} \sqcup \mathcal{M}_{\mathbb{T}}}$;
- $p' \overset{df}{=} p$.

So we simply add/remove sons of p satisfying the limit type to the original model. Recall that $v_{\mathcal{N}}(s) \subseteq \mathcal{N}$ is the *semantics* of the theory $s \in \mathbb{T}$ in a model $\mathcal{N}$, i.e. the set of all the points that satisfy s (see Definition 2.22).

We will show that $\mathcal{M}, p \equiv \mathcal{M}', p'$. By Proposition 2.17 it suffices to show that (i) the colours of the roots p and p' are the same and (ii) the sentences beginning with the diamond agree (i.e. ${}_{\diamond\backslash}\mathrm{tp}_{\mathcal{M}}(p) = {}_{\diamond\backslash}\mathrm{tp}_{\mathcal{M}'}(p')$). Of course, the atomic propositions satisfied in p and p' are the same, so the first part holds.
For (ii), first observe that for any model $\mathcal{K}$ and a point $q \in \mathcal{K}$:

$$\bigcup \mathbb{T}_q = \bigcup \overline{\mathbb{T}}_q. \tag{*}$$

This equation holds as otherwise there must be a formula φ s.t. φ does not belong to any type $t \in \mathbb{T}_q$, but it does belong to some type $t' \in \overline{\mathbb{T}}_q - \mathbb{T}_q$. However, this is impossible as then $\mathcal{U}_\varphi$ would be an open neighbourhood of t' disjoint with $\mathbb{T}_q$ and so $t' \notin \overline{\mathbb{T}}_q$.
Since t is a limit point of $\mathbb{T}_{\mathcal{M},p}$ (Definition 4.7):

$$\overline{\mathbb{T}_{\mathcal{M},p} - \{t\}} = \overline{\mathbb{T}_{\mathcal{M},p}} = \overline{\mathbb{T}_{\mathcal{M},p} \cup \{t\}}. \tag{**}$$

On the other hand, by definition of $\mathcal{M}'$, we have that either $\mathbb{T}_{\mathcal{M}',p'} = \mathbb{T}_{\mathcal{M},p} - \{t\}$ or $\mathbb{T}_{\mathcal{M}',p'} = \mathbb{T}_{\mathcal{M},p} \cup \{t\}$ and hence:

$$\text{either } \overline{\mathbb{T}_{\mathcal{M}',p'}} = \overline{\mathbb{T}_{\mathcal{M},p} - \{t\}} \text{ or } \overline{\mathbb{T}_{\mathcal{M}',p'}} = \overline{\mathbb{T}_{\mathcal{M},p} \cup \{t\}}. \tag{***}$$

Combining (**) and (***) we get that $\overline{\mathbb{T}_{\mathcal{M},p}} = \overline{\mathbb{T}_{\mathcal{M}',p'}}$ and hence $\bigcup \overline{\mathbb{T}_{\mathcal{M},p}} = \bigcup \overline{\mathbb{T}_{\mathcal{M}',p'}}$. By (*), this implies $\bigcup \mathbb{T}_{\mathcal{M},p} = \bigcup \mathbb{T}_{\mathcal{M}',p'}$. However, this completes the argument for modal equivalence of the two models, as for any $\mathcal{N}, q$, we have that ${}_{\diamond\backslash}\mathrm{tp}_{\mathcal{N}}(q) = \bigcup \mathbb{T}_{\mathcal{N},q}$ and therefore ${}_{\diamond\backslash}\mathrm{tp}_{\mathcal{M}}(p) = \bigcup \mathbb{T}_{\mathcal{M},p} = \bigcup \mathbb{T}_{\mathcal{M}',p'} = {}_{\diamond\backslash}\mathrm{tp}_{\mathcal{M}'}(p')$.

Although equivalent, $\mathcal{M}, p$ and $\mathcal{M}', p'$ are not bisimilar. Only one of them has a point satisfying t as a son of the root p (resp. p'), and thus $\mathcal{M}, p \not\Leftrightarrow \mathcal{M}', p'$ since bisimilarity implies modal equivalence and so in one of the models the root has a son that cannot be bisimilar to any son of the root in the other model. □

Combining the two lemmata, we prove the implication (1) ⇒ (2) – i.e. given a pointed model with a reachable point that has infinitely many sons, we show existence of a model that is equivalent but not bisimilar to it. Indeed, given a model $\mathcal{M}, r$ – w.l.o.g. a non-redundant reachable one – and a point $p \in \mathcal{M}$ having infinitely many sons, we may use Lemma 4.18 to find a pointed model $\mathcal{N}, q$ that is equivalent but not bisimilar to $<p>_{\mathcal{M}}, p$. If $p = r$, we are done. Otherwise, the substitution $\mathcal{M}[p \leftarrow \mathcal{N}, q], r$ is modally equivalent (by Proposition 4.4) but not bisimilar (by Lemma 4.16) to $\mathcal{M}, r$.

4.2.2. (2) ⇒ (1)

It is just a reformulation of the Hennessy–Milner theorem which states that for finitely branching models modal equivalence implies bisimilarity. However, the tools we introduce in this thesis allow us to give a new, topological proof. Moreover, the facts from which this implication follows will be needed later to prove the last implication anyway, so for the sake of completeness we prove it here.

Lemma 4.19. *Given a pointed model $\mathcal{M}, p$, if $\mathbb{T}_{\mathcal{M},p}$ is finite, then p is saturated. In particular, points with finitely many sons are saturated. As a consequence, finitely branching models are saturated.*

Proof. Observe that in $\mathbb{T}$ every finite set is closed. First, singletons are closed, as for any $t \in \mathbb{T}$, the family $\bigcup_{\varphi \in t} \mathcal{U}_{\neg\varphi}$ is an open cover of the complement of $\{t\}$. Thus, since closed sets are closed under finite unions, every finite set is closed, as it is a finite union of singletons. Therefore, if $\mathbb{T}_p$ is finite, then by Proposition 4.13 it means that p is saturated. □

The next lemma tells us that in fact saturated models are unique (up to bisimulation).

Lemma 4.20. *Modally equivalent saturated pointed models are bisimilar.*

Proof. Take two saturated models $\mathcal{M}_0$ and $\mathcal{M}_1$. We will show that the relation of modal equivalence $\equiv$ between $\mathcal{M}_0$ and $\mathcal{M}_1$ (i.e. the relation $Z \subseteq U_{\mathcal{M}_0} \times U_{\mathcal{M}_1}$ s.t. $p_0 Z p_1 \overset{df}{\iff} \mathcal{M}_0, p_0 \equiv \mathcal{M}_1, p_1$) is itself a bisimulation.
Equivalent points satisfy the same atomic sentences. For the back and forth conditions take any p_0, p_1 s.t. $\mathcal{M}_0, p_0 \equiv \mathcal{M}_1, p_1$ and a son q_i of p_i. Consider the type of q_i, i.e. $t \overset{df}{=} \mathrm{tp}_{\mathcal{M}_i}(q_i)$. For any its finite subset $t_{fin} \subseteq_{fin} t$, we know that p_i – and by equivalence also p_{1-i} – satisfy $\diamond \bigwedge t_{fin}$. But this just means that some son of p_{i-i} satisfies t_{fin}. Therefore, by saturation, some son q_{1-i} of p_{1-i} satisfies the entire t and thus it is modally equivalent to q_i. □

The implication (2) ⇒ (1) follows immediately from the two above lemmata.

4.2.3. (2) ⇒ (3)

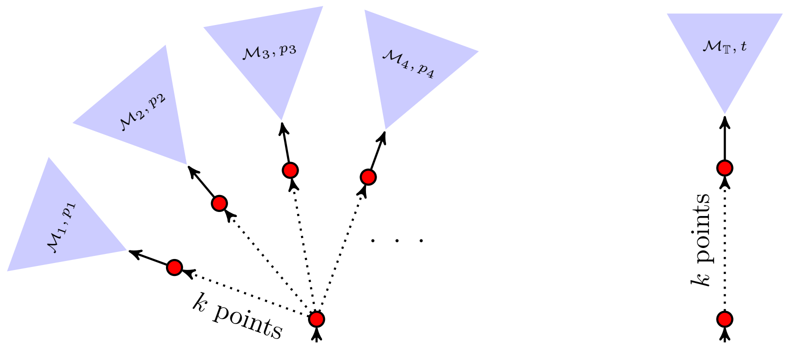
It is obvious.

4.2.4. (3) ⇒ (2)

This implication is the strengthened version of the Hennessy–Milner theorem (3.4) and appears to be a folklore result. Although it may be proved directly with elementary means, it is more insightful to give a proof involving the tools we developed, as it also sheds some light on the phenomenon of infinite branching up to bisimulation and its relation to modal logic.

Let us start with a warning. Ideally, we would like to prove that if a point $p \in \mathcal{M} \models t$ has infinitely many non-bisimilar sons, then any other point satisfying its theory $\mathcal{M}', p' \models t$ has infinitely many sons as well. Unfortunately, this is *not* true. The following example illustrates that the mere *existence* of a point with infinitely many sons is the best we can try to show:

Example 4.21. Recall that by Proposition 3.5, there are arbitrarily many pairwise non-bisimilar models. Thus, since $\mathbb{T}$ has a fixed cardinality, by the pigeonhole principle there are arbitrarily many pairwise non-bisimilar, *equivalent* models. Let us take an infinite family of such models $(\mathcal{M}_1, p_1), (\mathcal{M}_2, p_2), \ldots$ s.t. $\mathcal{M}_i, p_i \not\leftrightarrow \mathcal{M}_j, p_j$ for $i \neq j$, but they all have the same theory t, i.e. $\mathcal{M}_i, p_i \models t$ for all i.
We define a model $\mathcal{N}_k, q_k$ to be a model consisting of a root together with infinitely many identical paths of length k s.t. the i-th path ends with $\mathcal{M}_i, p_i$. This model is equivalent to a model $\mathcal{N}'_k, q'_k$ consisting of a single path of length k ending with the saturated model of t, i.e. $\mathcal{M}_{\mathbb{T}}, t$. The following picture illustrates both $\mathcal{N}_k, q_k$ and $\mathcal{N}'_k, q'_k$:



It is easy to see that the two models are modally equivalent, as $\mathcal{M}_{\mathbb{T}}, t$ and all $\mathcal{M}_i, p_i$ satisfy the same theory t. On the other hand, w.l.o.g. every $\mathcal{M}_i$ is non-redundant which implies that $\mathcal{N}_k$ is non-redundant as well. Hence, in $\mathcal{N}_k$ the root has infinitely many non-bisimilar sons, while in $\mathcal{N}'_k$ every point at depth lower than k has only one son.
We skip the formal description and the details of the proof of correctness of the above construction.

This also illustrates that *extending* models is not always as natural as one could expect. It is not hard to check that every model $\mathcal{M}$ can be extended to a saturated model $\mathcal{M}'$ by simply adding appropriate points satisfying all the limit theories. However, this operation could in fact *shrink* our initial model *up to bisimulation* – as all the points that were *equivalent* in $\mathcal{M}$ became *bisimilar* in $\mathcal{M}'$. Indeed, in the above example w.l.o.g. we may assume that $\mathcal{M}_1, p_1 = \mathcal{M}_{\mathbb{T}}, t$. Then, the extension of $\mathcal{N}_k, q_k$ to a saturated model would be bisimilar to its submodel $\mathcal{N}'_k, q'_k$.

Fortunately, it turns out that although modal equivalence does not preserve the property of having infinitely many sons, with a little bit more care we may prove a weaker claim that is still sufficient for our purposes.

Lemma 4.22. *If a theory t has a model $\mathcal{M}, p$ that is infinitely branching up to bisimulation, then every its model must be infinitely branching.*

Proof. Take a model $\mathcal{M}, p \models t$ that is infinitely branching up to bisimulation. As it easily follows from the Definition 3.8, being infinitely branching up to bisimulation is the same as having a reachable point with infinitely many pairwise non-bisimilar sons.[5] Thus, some reachable point $r \in \mathcal{M}$ has infinitely many pairwise non-bisimilar sons.

First, we will show that for some reachable point $q \in \mathcal{M}$, $\mathbb{T}_{\mathcal{M},q}$ is infinite. Suppose towards contradiction that it is not true. Then, by Lemma 4.19, every reachable point is saturated which means just the same as saturation of $<p>, p$. Therefore, $\mathcal{M}, p$ is bisimilar to a saturated model. However, every saturated model is bisimilar to the canonical one $\mathcal{M}_{\mathbb{T}}$, i.e. $\mathcal{M}, p \leftrightarrows \mathcal{M}_{\mathbb{T}}, t$ (Lemma 4.20), in which for any q, $\mathrm{tp}_{\mathcal{M}_{\mathbb{T}}}(q) = q$ (Proposition 4.15) and hence $f_{\mathcal{M}_{\mathbb{T}}}(q) = \mathbb{T}_{\mathcal{M}_{\mathbb{T}},q}$. Therefore, since by our assumption r is a point that has infinitely many non-bisimilar sons, its bisimilar counterpart $r' \in \mathcal{M}_{\mathbb{T}}$ has to have infinitely many sons as well and $\mathbb{T}_{\mathcal{M}_{\mathbb{T}},r'}$ is infinite. But by bisimilarity of r and r' also $\mathbb{T}_{\mathcal{M},r}$ has to be infinite, since every son q of r' has to have a bisimilar counterpart among r's sons satisfying q's theory.

[5] If there is no such a point, then in $<p>, p$ every point has finitely many pairwise non-bisimilar sons and its quotient by bisimilarity $<p>_{/\leftrightarrows}, [p]_{/\leftrightarrows}$ is a finitely branching model bisimilar to $\mathcal{M}, p$.

Having a reachable point q with an infinite $\mathbb{T}_q$, we show that no model of t is finitely branching. By definition of reachability, there must be a path in $\mathcal{M}$ – say of length n – from the root p to q. We will show that for any k, any model of t must contain at least k points at depth $n+1$ (and as a consequence, it contains infinitely many points at that depth). Indeed, taking distinct types $t_1, ...t_k \in \mathbb{T}_q$ we can distinguish any two different t_i, t_j with some formula $\varphi_{i,j}$ belonging to exactly one of them. This allows us to construct k sentences $\psi'_l \stackrel{df}{=} \bigwedge(t_l \cap \{\varphi_{i,j}, \neg\varphi_{i,j} \mid i,j \leq k\})$ and $\psi_l \stackrel{df}{=} \underbrace{\Diamond ... \Diamond}_{\text{n+1 times}} \psi'_l$. Since $\mathcal{M}, p$ satisfies every ψ_l, they all belong to t. On the other hand, no point can satisfy two different ψ'_i, ψ'_j – which means that any model of t has at least k points at depth $n+1$.

On the other hand, a finitely branching model can only have a finite number of points at depth $n+1$. Therefore, every model of t must be infinitely branching. □

This completes the proof of the main theorem. Q.E.D.

Chapter 5

Further investigations

5.1. Beyond modal logic

Let us start with a disclaimer. Due to limited space of this thesis and the fact that the further questions remain mostly unanswered, in this chapter we only present a sketch of insights and examples we have (most of which are actually counterexamples for natural, promising hypotheses).

Motivated by the main theorem of this thesis, one could wonder if there are any nice characterisations for other bisimulation-invariant logics analogous to the one we presented. For instance, it is natural to ask whether it is possible to characterise in a simple way the bisimulation categoricity of a complete μ-calculus theory (for more information on the μ-calculus μ-ML see [20]). However, this question appears to be much harder than the one we answered. The key problem is that, unlike ML, μ-ML is not compact. It turns out that at least some of the most natural hypotheses for μ-ML fail and we still lack a good understanding of the phenomena we have to deal with when fixpoint operators are introduced.

Given that, it seems reasonable to first focus on a restricted parts of μ-ML in order to develop some intuitions and tools before approaching the question about the full μ-ML. One possible direction of research is the *transitive* modal logic ML^+, where the diamond $\Diamond$ corresponds to the transitive closure R^+ of the accessibility relation R instead of just R. A side question that is interesting in its own right is about the categoricity of ML^+ up to a relaxed bisimulation relation – a *transitive bisimulation* $\leftrightarroweq^+$ (see [4], where ML^+ and $\leftrightarroweq^+$ on forests is discussed under the names 'logic EF' and 'EF-bisimulation', respectively).

5.2. Towards μ-ML

Given the nice and simple result we present in this thesis, it is tempting to ask if this characterisation is true for μ-ML. However, the proof we presented is no longer valid in the context of μ-ML, as this logic is not compact. In fact, it is not hard to find a counterexample for a generalisation of our main result to μ-ML: the full μ-ML theory of the Hedgehog (Example 3.7) *has* a unique model up to bisimulation – as it suffices to enrich its ML theory with a single μ-ML formula expressing well-foundedness[1]. At the same time the set of all μ-ML formulae has a fixed cardinality, so there are non-categorical μ-ML theories, as the class of all models counted up to bisimulation cannot be represented as a set (Proposition 3.5).

[1] Namely, $\neg\nu x.\Diamond x$ – 'It is not true that there exists an infinite path from the current point'.

On the other hand, although the Skolem-Löwenheim theorem cannot be generalised to the full MSO (e.g. there are MSO theories whose all models have cardinality continuum), it is true for its bisimulation-invariant part – the μ-ML – that if a μ-ML theory has an infinite model, then it has a model of size at most $\kappa \stackrel{df}{=} \max(\omega, |\Sigma|)$ (see [12]). This gives a bound on which theories can be categorical, as no theory with a model that is more than κ-branching up to bisimulation can be categorical.

A natural guess now could be that a μ-ML theory is categorical iff all its models are κ-branching up to bisimulation (for κ defined above). Unfortunately, this hypothesis also fails. One can find a μ-ML theory that has, up to bisimulation, exactly two models – both of which are countably branching up to bisimulation. Moreover, both models are well-founded.

5.3. Transitive logic

Since the question about categoricity of a complete μ-ML theory appears to be much more difficult (as second-order quantification, even restricted, seems harder to deal with), it is reasonable to investigate an easier case first. One possibility would be to consider a *transitive modal logic* ML^+ – which is the same as ML apart from the interpretation of the diamond $\Diamond$, where $\mathcal{M}, p \models \Diamond\varphi$ means 'There exists a point q *reachable in a finite number of steps* from p s.t. $\mathcal{M}, q \models \varphi$' – so we just use the transitive closure R^+ of the accessibility relation R instead of R. It is not hard to see that this logic is a part of μ-ML – any formula $\varphi \in \mathrm{ML}^+$ can be inductively translated to an equivalent μ-ML formula $\mathrm{tr}(\varphi) \in \mu$-ML, where the only non-trivial case is the diamond, whose translation can be defined by: $\mathrm{tr}(\Diamond\varphi) \stackrel{df}{=} \mu x.\Diamond(\mathrm{tr}(\varphi) \vee x)$.
While ML^+ is (strictly) weaker than μ-ML, it is *incomparable* with ML. There are contexts where it can express more than ML, ~~e.g. one can check that the ML^+ theory of the Hedgehog is categorical~~ [this statement is wrong, but there are other examples of properties expressible in ML^+, but not ML]. However, there are contexts where ML can express more than ML^+. Any two models where all the occurring colours are *dense* (i.e. for any colour c that is realised at least once in a model and any point p, p has a descendant q satisfying c) are indistinguishable from the perspective of ML^+. Therefore, even a unary branching model (i.e. a sequence) with only two colours does not have to have a bisimulationally categorical theory. For instance, the sequence (viewed as a model) $(a, b, a, b, ...)$ is indistinguishable from $(a, b, b, a, b, b, ...)$ as they satisfy the same ML^+ sentences. Generalisation of this example yields a relaxed notion of bisimulation: if we replace the word 'successor' with 'descendant' in the definition of a bisimulation (Definition 2.6), we get a notion of a *transitive bisimulation* (let us denote the transitive bisimilarity by $\leftrightarroweq^+$) that is slightly weaker than the full bisimilarity, but still suffices for ML^+, as it is invariant under $\leftrightarroweq^+$ defined that way.
A side question that seems interesting in itself is the problem of categoricity of ML^+ up to $\leftrightarroweq^+$. So far the only result we have here is that there are models whose ML^+ theory is not categorical with respect to $\leftrightarroweq^+$. To see this, first observe that for models with transitive accessibility relation, $\leftrightarroweq^+$ is the same as just $\leftrightarroweq$. Thus, the collection of all models counted up to transitive bisimulation $\leftrightarroweq^+$ cannot be represented as a set, for the construction of pairwise non-bisimilar models corresponding to each ordinal we gave (Example 3.6) actually provides transitive models.

Chapter 6

Appendix and notation

Here we provide a description of the notation we use.

- We often use pictures to illustrate models. The (hopefully intuitive) convention is as follows: nodes represent points of a model, arrows represent accessibility relation (that is, an arrow from a node representing point p to a node representing point q means that pRq holds). A tiny arrow 'from nowhere' distinguishes the initial point. Different colours of the nodes represent different colours of points (recall that a 'colour' of a node is the set of atomic proposition it satisfies – an element of $\mathcal{P}(\Sigma)$).

- We use the lambda notation to easily and concisely define functions. However, no knowledge of the lambda calculus is assumed except for the convention that $\lambda x.y$ denotes a function taking x as an argument and returning y as a value, e.g. $\lambda n.(n+1)$ is a function that takes a number n and returns this number plus 1.
  This example also illustrates the fact that one has to be careful while defining functions that way, as a lambda term alone does not bring any information about the intended *type* of a function. For instance, the mentioned lambda term $\lambda n.(n+1)$ could represent a function on naturals, integers, rationals, etc.

- Given two sequences w, v we denote its concatenation by $w \cdot v$. Moreover, we abuse notation by writing $w \cdot L$ for a concatenation of a sequence w with a *set* of sequences L, i.e. $w \cdot L \stackrel{df}{=} \{w \cdot v \mid v \in L\}$.

- We use the standard language-theoretic notation, so X^+ stands for the collection of all finite non-empty sequences over X (i.e. $X^+ = \bigcup_{n\in\omega} X^{n+1}$ – the least set containing X and closed under concatenation).

- Apart from the standard notation $X \subseteq Y$ denoting the fact that the set X is a subset of Y, we also write $X \subseteq_{fin} Y$ to highlight that X is a finite subset of Y.

- We use a bigger and smaller versions of the same symbol as a prefix and infix (respectively) to denote an operation on a set or a pair. For instance, the union of a family of sets $\mathcal{A}$ is denoted $\bigcup \mathcal{A}$, while $A \cup B$ stands for the union of two sets A and B. Another example could be the symbol for conjunction: $\varphi \wedge \psi$ is the conjunction of φ and ψ, and $\bigwedge \Phi$ is the conjunction of all the elements of Φ (note that in this case we have to assume finiteness of Φ).

Bibliography

[1] A. Baltag, H. van Ditmarsch, L.S. Moss, *Epistemic Logic and Information Update*, in Handbook on the Philosophy of Information, pp. 361-456, 2008

[2] J. van Benthem, *Modal Correspondence Theory*, PhD dissertation, Mathematisch Instituut & Instituut voor Grondslagenonderzoek, University of Amsterdam, 1976

[3] P. Blackburn, M. de Rijke, Y. Venema, *Modal Logic*, in Cambridge Tracts in Theoretical Computer Science, Vol. 53, 2001

[4] M. Bojańczyk and T. Idziaszek, *Algebra for Infinite Forests with an Application to the Temporal Logic EF*, in Proceedings of CONCUR, 2009

[5] V. Ciancia and Y. Venema, *Stream Automata Are Coalgebras*, in The Eleventh International Workshop of CMCS 2012, LNCS 7399, pp. 90-108, 2012

[6] K. Fine, *Some Connections Between Modal and Elementary Logic*, in Proceedings of The Third Scandinavian Logic Symposium, pp. 13-31, 1975

[7] R. Goldblatt, *Saturation and the Hennessy - Milner Property*, in Modal Logic and Process Algebra. A Bisimulation Perspective, CSLI Lecture Notes, Vol. 53, pp. 189-216, 1995

[8] V. Goranko and M. Otto, *Model Theory of Modal Logic* in Handbook of Modal Logic, pp. 255-325, 2006

[9] M. Hennessy and R. Milner, *Algebraic Laws for Indeterminism and Concurrency*, in Journal of the ACM, Vol. 32, pp. 137-162, 1985

[10] D. Janin and I. Walukiewicz, *On the Expressive Completeness of the Propositional μ-calculus w.r.t. Monadic Second-Order Logic*, in Proceedings of CONCUR, pp. 263-277, 1996

[11] D. Kozen, *A Finite Model Theorem for the Propositional μ-calculus*, in Studia Logica, Vol. 47, pp. 233-241, 1988

[12] D. Kozen and R. Parikh, *A Decision Procedure for the Propositional μ-calculus*, in Proceedings of the Workshop on Logics of Programs 1983, LNCS, pp. 313-325, 1983

[13] D.C. Makinson, *On Some Completeness Theorems in Modal Logic*, in Zeitschrift für mathemathische Logik und Grundlagen der Mathematik, Vol. 12, pp. 379-84, 1966

[14] D. Marker, *Model Theory: An Introduction*, in Graduate Texts in Mathematics, Vol. 217, 2002

[15] D. Park, *Concurrency and Automata on Infinite Sequences*, in Proceedings of The Fifth GI Conference, pp. 167-183, 1981

[16] P.H. Rodenburg, *Intuitionistic Correspondence Theory*, PhD thesis, University of Amsterdam, 1986.

[17] J.J.M.M. Rutten, *Automata and Coinduction (An Axercise in Coalgebra)*, in Proceedings of CONCUR, LNCS Vol. 1466, pp. 194-218, 1998.

[18] T. Skolem, *Logico-Combinatorical Investigations in the Satisfiability or Provabilitiy of Mathematical Propositions: A Simplified Proof of a Theorem by L. Löwenheim and Generalizations of the Theorem*, in From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931 (3rd ed.), Massachusetts: Harvard University Press, pp. 252-263, 1997

[19] A.Visser, *An Overview of Interpretability Logic*, in Advances in Modal Logic, Vol. 1, pp. 307-359, 1998

[20] Y. Venema, *Lectures on the Modal μ-calculus*, ILLC, University of Amsterdam, 2012