



## UvA-DARE (Digital Academic Repository)

### Epistemic modelling and protocol dynamics

Wang, Y.

**Publication date**  
2010

[Link to publication](#)

#### **Citation for published version (APA):**

Wang, Y. (2010). *Epistemic modelling and protocol dynamics*. [Thesis, fully internal, Universiteit van Amsterdam]. Institute for Logic, Language and Computation.

#### **General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

#### **Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

## 6.1 Introduction

In Chapter 3, we argued that a formal definition of a protocol should come with a finite set of formulas which specify the initial setting in which the protocol is to be executed. The verification of the protocol should be performed against the models which satisfy these initial requirements. By making the initial requirements explicit in terms of logical formulas, we may narrow down the gap between the informal scenario and the formal initial model, which is supposed to be a mathematical abstraction of the former. This may help us to gain more grip on the “*model hacking*” which precedes model checking. A natural question to ask at this point is whether the initial assumptions induce a *unique* model? Moreover, if not, *how many* different models are there? The precise meaning of the above questions depends on the logical language we use and the notion of equivalence between models. Since we are interested in modal logics which are bisimulation invariant, we fix bisimulation as the equivalence notion.

First note there are formulas that have unique models modulo bisimulation. We say a formula *characterizes* a pointed model  $(\mathcal{M}, s)$  if any model of it is bisimilar to  $(\mathcal{M}, s)$ . It is shown in [BM96, vB98] that a modal logic equipped with iteration or fixed point operators can characterize arbitrary finite models up to bisimulation<sup>1</sup>. For example, given  $\mathbf{P} = \{p\}$  and  $\mathbf{I} = \{1, 2\}$ , interested readers can verify that the following formula has a unique S5 model modulo bisimulation<sup>2</sup>:

$$C_{1,2}((p \rightarrow (\widehat{K}_1 p \wedge \widehat{K}_1 \neg p \wedge K_2 p \wedge \widehat{K}_2 p)) \wedge (\neg p \rightarrow (\widehat{K}_1 p \wedge \widehat{K}_1 \neg p \wedge K_2 \neg p \wedge \widehat{K}_2 \neg p)))$$

Similar results in computer science with respect to branching time temporal logics can be found, e.g., in [BCG87]. In the context of dynamic epistemic modelling,

<sup>1</sup>It is not hard to see that a logic that enjoys the finite model property can only characterize finite models up to bisimulation.

<sup>2</sup>The model should represent a scenario where 1 does not know whether  $p$  but 2 knows, and this is common knowledge.

[vDvdHK03a, vD02] demonstrate that there are intuitive epistemic formulas (*descriptions*) that characterize the initial models in the case of the card games. However, in general, a set of formulas translated from an informal description of the scenario may not have a unique model.

In this chapter, we make the initial steps towards an answer for the question: how many non-bisimilar models are there for a given formula? To make our results general enough, we take the *Propositional Modal  $\mu$ -Calculus* ( $\mathbf{Mu}$ ) as the logical language for specifying the initial requirements.

Since its invention by Kozen [Koz83],  $\mathbf{Mu}$  has received great interest in computer science due to its neat syntax, strong expressive power and nice model-theoretical properties (see, e.g., [BS06] for an introduction). The language of  $\mathbf{Mu}$  includes general least (greatest) fixed-point formulas in the shape of  $\mu X.\phi$  ( $\nu X.\phi$ ), thus superseding the usual temporal logics in expressive power. For example, a PDL formula  $[a^*]p$  can be translated as  $\nu X.(p \wedge [a]X)$  in  $\mathbf{Mu}$ . It is shown in [vBI08] that adding product updates to  $\mathbf{Mu}$  does not increase its expressive power. Therefore,  $\mathbf{Mu}$  can actually be considered as a very powerful logic of communication, bearing in mind the potential dynamic epistemic applications. This is the rationale behind the choice of this very powerful logic as the logic in focus in this chapter.

Many theoretical issues of  $\mathbf{Mu}$  have been studied extensively during the past three decades (cf., e.g., [Wal00, JW95, BS06]). In particular, this chapter is inspired by [Niw91], in which Niwiński tackled the cardinality question of the tree languages recognized by Rabin tree automata. It is also suggested in [Niw91] that the results induce a method to evaluate the number of models of a given formula, due to the fact that formulas of temporal logics can be reduced to Rabin automata. However, in the original paper, the author focused on automata on binary trees and counted models modulo *isomorphism*, which limited the use of the results in the setting of modal logics which are invariant under bisimulation on Kripke models.

In this chapter, we want to pursue this line of research further by counting models modulo bisimulation. Note that we will not work with the  $\mathbf{Mu}$  formulas directly, but consider the corresponding *alternating tree automata* (*ATA*) ([JW95] shows the equivalence of formulas of  $\mathbf{Mu}$  and *ATA*). Like [Niw91], but with non-trivial complication of the proof, we show that an *ATA* recognizable set of image-finite models modulo bisimulation is uncountable iff it is of the cardinality continuum iff it contains a non-regular tree. We also give a normal form of the countable languages recognized by an *ATA*. These results constitute the first steps towards an algorithm to output the number of models modulo bisimulation.

**Related Work** In the field of epistemic logic, an important question to ask is: how many different *states* of knowledge of a given fact are there? This question has been addressed by [Aum89, Har96] and [PK92, Par03] in different settings. It is interesting to see that these discussions in the literature can be unified from a perspective of counting models. For example, though presented in the setting of the *information structures* of [Aum76], the set-theoretical counterpart of epistemic logic for economists, [Har96] essentially shows that there are continuum many  $S5$  Kripke

models such that any two models are separated by an epistemic formula based on a basic proposition  $p$ . If we only consider image-finite models, then bisimulation on S5 models coincides with the logical equivalence of epistemic logic, therefore the above result can be equally rephrased as: there are continuum many non-bisimilar S5 models of  $p$ .

On the other hand, the main result in [PK92, Par03] says that there are countably many different realizable “levels” of knowledge, where a level of knowledge realized by a certain pointed S5 Kripke model  $(\mathcal{M}, s)$  is a set:

$$\{(i_1, i_2, \dots, i_n) \in \mathbf{I}^* \mid \mathcal{M}, s \models K_{i_1} K_{i_2} \dots K_{i_n} p\}$$

In our setting, this amounts to counting models of  $p$  modulo an equivalence notion which is much weaker than bisimulation: it is not hard to see that two models have the same level of knowledge iff they have the same set of labelled paths to  $\neg p$  worlds<sup>3</sup>. Therefore, it is expected that there are less levels of knowledge than states of knowledge in the sense of [Har96].

A generalization of the results in [Niw91] is presented in [BKR09], where the focus is on the elimination of the *uncountability quantifiers* in the setting of the monadic second-order logic of order over image-finite trees.

**The structure of the chapter** In the next section, we will recall some standard definitions and results on alternating tree automata. Section 6.3 presents our main result on the cardinality of ATA recognizable languages. In Section 6.4 we give a normal form of the countable languages recognized by an ATA. In the last section, we discuss some interesting implications and further extensions of our result in an epistemic setting.

## 6.2 Preliminaries

We first define *trees*. For a non-empty set  $\Delta$ , a  $\Delta$  labelled-tree is a tuple:

$$\mathcal{T} = (W, \Delta, \Sigma, \rightarrow, L, r)$$

where  $W$  is a set of vertices with designated node  $r \in W$  as the *root*,  $L : W \rightarrow \Delta$  is a labelling function for the nodes and  $\{\rightarrow_a\}_{a \in \Sigma} \subset W \times W$  is a set of edges such that the root has no incoming edges, and there is a unique directed *finite* path from the root to every node. We also write  $W_{\mathcal{T}}$  for the domain of  $\mathcal{T}$ , similarly for  $L_{\mathcal{T}}$ ,  $\rightarrow_{\mathcal{T}}$  and  $r_{\mathcal{T}}$ . We call two nodes  $w$  and  $w'$  *siblings* if they are two successors of the same node. By abusing the notation, we sometimes write  $v \in \mathcal{T}$  for  $v \in W_{\mathcal{T}}$ .

**Notation** The reader should not be confused when seeing  $w, v$  as the vertices in a tree while they are reserved to denote sequences in other parts of the thesis. In fact,

<sup>3</sup>The case for common knowledge in [Par03], can be viewed in the same light with little adaptation.

a tree is sometimes represented as a prefix closed subset of  $\mathbb{N}^*$  and thus every vertex in a tree is essentially a prefix of a branch (a sequence).

Let  $dep(\mathcal{T}, v)$  be the *depth* of  $v$  in  $\mathcal{T}$ , namely the length of the path from  $r_{\mathcal{T}}$  to  $v$ . We say  $v \rightarrow^* w$  if  $w = v$  or  $w$  is reachable from  $v$  by following the edges, and  $v \rightarrow^n w$  if it is reachable in  $n$  steps. Intuitively, the nodes that are of the same depth form a *level* of the tree. If  $\mathcal{T}$  is a tree then  $\mathcal{T}^v$  denotes its *subtree* rooted at  $v$ . Let  $Sub_{\leftrightarrow}(\mathcal{T})$  be the set of subtrees of  $\mathcal{T}$  modulo bisimulation:  $\{[\mathcal{T}^v]_{\leftrightarrow} \mid v \in \mathcal{T}\}$ . We say  $\mathcal{T}$  is *bisimulation-regular* (*B-regular* for short) if  $|Sub_{\leftrightarrow}(\mathcal{T})|$  is finite. For any  $u, s$  in  $\mathcal{T}$ , let  $\mathcal{T}[u \setminus s]$  be the tree constructed from  $\mathcal{T}$  by replacing the subtree  $\mathcal{T}^u$  with  $\mathcal{T}^s$ .

It is clear that every tree can be viewed as a pointed Kripke model with labelling function instead of valuations. On the other hand, for an arbitrary Kripke model  $\mathcal{M}, s_0$  we can associate its *tree unravelling*, the tree  $\mathcal{T}^{\mathcal{M}} = (W, 2^{\mathbf{P}}, \Sigma, \rightarrow, L, r)$  where  $W$  is the set of all possible finite paths  $s_0 a_0 s_1 a_1 \dots a_n s_n$  starting with  $s_0$  in  $\mathcal{M}$ , and  $w' \xrightarrow{b} w \iff w = w'bs$  in  $\mathcal{M}$ . It is not hard to see that each Kripke model is bisimilar to its tree unravelling.

Recall that given a Kripke model  $\mathcal{M} = (S, \Sigma, \rightarrow, V)$ , the bisimulation contraction of  $\mathcal{M}$  is the quotient model  $\mathcal{M}_{/\equiv_b}$  (see Definition 2.2.3). It follows that if a tree is B-regular, then its bisimulation contraction is a finite Kripke model.

In [JW95], general alternating tree automata (ATA) on Kripke models are defined. For technical convenience, we work with the  $\mu$ -automata as in [DN05], which can be viewed as an equivalent but intuitive exposition of ATA. These  $\mu$ -automata run on  $2^{\mathbf{P}}$ -labelled (infinite) trees, where  $\mathbf{P}$  is the set of basic propositions.

**6.2.1. DEFINITION. ( $\mu$ -automata [JW95, DN05])** A  $\mu$ -automaton  $\mathbf{A}$  on set of basic propositions  $\mathbf{P}$  and set of basic actions  $\Sigma$  is a tuple:

$$\mathbf{A} = (Q, B, q_0, \rightarrow_{\text{OR}}, \rightarrow_{\text{BR}}, L, \Omega)$$

where:

- $Q$  is a non-empty, finite set of OR (choice) states,
- $B$  is a finite set of BR (branching) states such that  $B \cap Q = \emptyset$ ,
- $q_0 \in Q$  is an start state,
- $\rightarrow_{\text{OR}} \subseteq Q \times B$  is an unlabelled choice relation from  $Q$  to  $B$ ,
- $\rightarrow_{\text{BR}} \subseteq B \times \Sigma \times Q$  is a labelled branching relation from  $B$  to  $Q$ ,
- $L : B \rightarrow 2^{\mathbf{P}}$  is a labelling function mapping each branching state to a set of basic propositions,
- $\Omega : Q \rightarrow \mathbb{N}$  is an indexing function.

Let  $\text{BR}(b, a) = \{q \mid (b, a, q) \in \rightarrow_{\text{BR}}\}$  and  $\text{OR}(q) = \{b \mid (q, b) \in \rightarrow_{\text{OR}}\}$ . A  $q$ -run  $\mathcal{R}$  of  $\mathbf{A} = (Q, B, q_0, \rightarrow_{\text{OR}}, \rightarrow_{\text{BR}}, L, \Omega)$  on a  $(2^{\mathbb{P}})$ -labelled tree  $\mathcal{T} = (W_{\mathcal{T}}, 2^{\mathbb{P}}, \Sigma, \rightarrow_{\mathcal{T}}, L_{\mathcal{T}}, r)$  is a  $(W_{\mathcal{T}} \times (Q \cup B))$ -labelled tree  $\mathcal{R} = (W_{\mathcal{R}}, W_{\mathcal{T}} \times (Q \cup B), \Sigma \cup \{\tau\}, \rightarrow_{\mathcal{R}}, L_{\mathcal{R}}, r')$  such that the following conditions are satisfied:

- $L_{\mathcal{R}}(r') = (r, q_0)$ .
- (OR) If  $L_{\mathcal{R}}(w) = (v, q)$ , where  $q \in Q$  then  $w$  has exactly one  $\tau$ -child  $w'$  such that  $L_{\mathcal{R}}(w') = (v, b)$  for some  $b \in \text{OR}(q)$ .
- (BR) If  $L_{\mathcal{R}}(w) = (v, b)$  where  $b \in B$  then:
  - $L_{\mathcal{T}}(v) = L(b)$ .
  - For every  $a$ -child  $v'$  of  $v$  in  $\mathcal{T}$ , there is an  $a$ -child  $w'$  of  $w$  in  $\mathcal{R}$  such that  $L_{\mathcal{R}}(w') = (v', q')$  for some  $q' \in \text{BR}(b, a)$ .
  - For every  $q' \in \text{BR}(b, a)$  there is an  $a$ -child  $w'$  of  $w$  in  $\mathcal{R}$  such that  $L_{\mathcal{R}}(w') = (v', q')$  for some  $a$ -child  $v'$  of  $v$  in  $\mathcal{T}$ .

For a path  $P$  of  $\mathcal{T}$ , we define  $Q^{\mathcal{R}}(P) = \{q \mid L_{\mathcal{R}}(w) = (v, q) \text{ for some } w \in \mathcal{R} \text{ and } v \text{ on } P\}$ . Note that  $Q^{\mathcal{R}}(v)$  is not always a singleton since one node in  $\mathcal{T}$  may correspond to more than one state in the automaton. For an infinite path  $P = v_0, v_1, \dots$  we define:

$$\text{Inf}(\mathcal{R}, P) = \{q \mid q \in Q^{\mathcal{R}}(v_i) \text{ for infinitely many } v_i\}$$

The acceptance of runs is defined by the *parity* condition: A  $q$ -run  $\mathcal{R}$  of  $\mathbf{A}$  on  $\mathcal{T}$  is *accepting* iff for every infinite path  $P$  in  $\mathcal{R}$ : the greatest value of  $\Omega^{\mathcal{R}}(q)$ , for  $q \in \text{Inf}(\mathcal{R}, P)$ , is even. We denote such greatest value as  $\Omega^{\mathcal{R}}(P)$ . A tree  $\mathcal{T}$  is accepted by  $\mathbf{A}$  if there is a  $q_0$ -accepting run on  $\mathcal{T}$ . Let  $\mathcal{L}(\mathbf{A})$  be the set of trees which are accepted by  $\mathbf{A}$ . A pointed Kripke model  $\mathcal{M}, s$  is accepted by  $\mathbf{A}$  iff its tree unravelling is accepted by  $\mathbf{A}$ .

Given a run  $\mathcal{R}$  of  $\mathbf{A}$  on  $\mathcal{T}$  and a state  $v \in \mathcal{T}$ , we let  $W^{\mathcal{R}}(v) = \{w \mid L_{\mathcal{R}}(w) = (v, q) \text{ for some } q \in Q_{\mathbf{A}}\}$ . By the definition of the run, we can verify that any two nodes in  $\mathcal{R}$  labelled by the same node in  $\mathcal{T}$  are at the same level in  $\mathcal{R}$ : for any  $w, w' \in W^{\mathcal{R}}(v)$ :  $\text{dep}(\mathcal{R}, w) = \text{dep}(\mathcal{R}, w')$ .

The following fundamental theorem relates the  $\mathbf{Mu}$  formulas with  $\mu$ -automata.

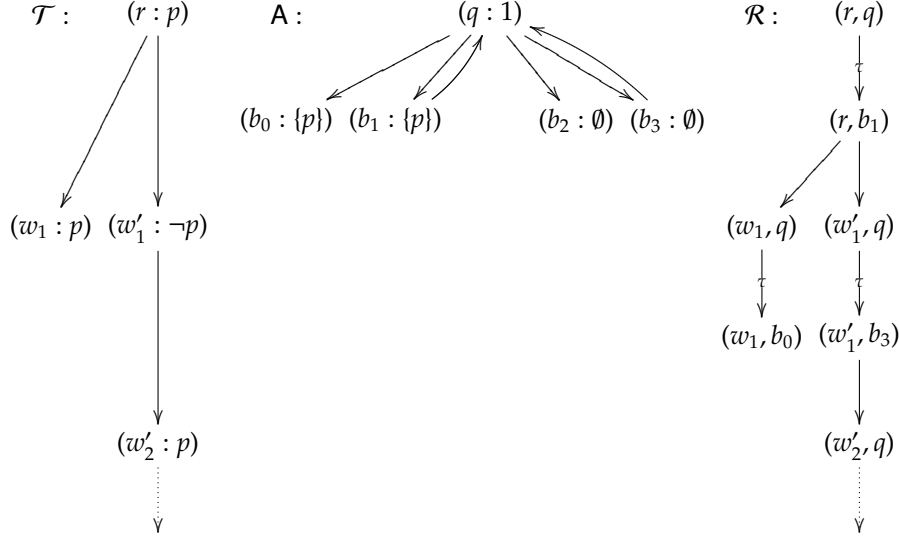
**6.2.2. THEOREM.** [JW95] *For each  $\mu$ -automaton there is an equivalent  $\mathbf{Mu}$ -formula. For each  $\mathbf{Mu}$ -formula there is an equivalent  $\mu$ -automaton.*

Since  $\mathbf{Mu}$  is invariant under bisimulation then we have:

**6.2.3. COROLLARY.** *If  $\mathcal{M}, s \Leftrightarrow \mathcal{N}, t$ , then  $\mathcal{M}, s$  is accepted by a  $\mu$ -automaton  $\mathbf{A}$  iff  $\mathcal{N}, t$  is accepted by  $\mathbf{A}$ .*

We end this section with an example to illustrate how the  $\mu$ -automata work:

## 6.2.4. EXAMPLE.



where we fix  $\mathbf{P} = \{p\}$ . In the tree  $\mathcal{T}$ ,  $(x : y)$  indicates that  $y$  holds at state  $x$ ; In  $A$ ,  $(q : 1)$  indicates that  $\Omega(q) = 1$  and  $(b_i : Z)$  means  $L_A(b_i) = Z$ ; In the run  $\mathcal{R}$ ,  $(x, y)$  are the labels associated to the nodes. It is not hard to see that  $\mathcal{R}$  is a run of  $A$  on  $\mathcal{T}$ , however, it is not accepting. Actually  $A$  is the  $\mu$ -automaton corresponding to the  $\mathbf{Mu}$  formula  $\mu X. \Box X$  which expresses well-foundedness: there is no infinite descending chain (cf., e.g., [GO06]).  $\Omega$

### 6.3 Cardinality of the Tree Languages

The collection of the models recognized by an arbitrary  $\mu$ -automaton is not always a set, even up to bisimulation. For instance, take the  $\mu$ -automaton  $A$  in Example 6.2.4 which is equivalent to the  $\mu$ -calculus formula  $\mu X. \Box X$  that expresses well-foundedness. Now for any ordinal  $\alpha$ , consider the Kripke structure  $\mathcal{M}_\alpha = \{\{\beta \mid \beta \leq \alpha\}, \rightarrow, V\}$  where  $\beta \rightarrow \beta' \iff \beta' < \beta \leq \alpha$  (the inverse of the order relation). It is clear that for all  $\alpha$ ,  $\mathcal{M}_\alpha$  is well-founded thus recognized by the automaton  $A$ . By induction, one can prove that  $\mathcal{M}_\alpha$  is not bisimilar to  $\mathcal{M}_\beta$  if  $\alpha \neq \beta$ . Therefore there are  $\mu$ -automata which recognize class-size tree languages.

In this chapter, we concentrate on image-finite models in which each state has only finitely many successors w.r.t the same label in  $\Sigma$ . It is clear that the tree unravelling of an image-finite model is again an image-finite model. Note that the class of image-finite models has the Hennessy-Milner property for  $\mathbf{Mu}$ : for any image-finite models  $\mathcal{M}, s_0$  and  $\mathcal{N}, t_0$ :  $\mathcal{M}, s_0 \simeq \mathcal{N}, t_0 \iff \mathcal{M}, s_0$  and  $\mathcal{N}, t_0$  satisfy the same  $\mathbf{Mu}$  formulas (cf., e.g., [BdRV02]). Therefore, if a  $\mathbf{Mu}$ -formula  $\phi$  has  $n$  non-bisimilar image-finite models, then all these different models can be told apart by  $\mathbf{Mu}$ -formulas.

The rest of this section is devoted to our main result (Theorem 6.3.5): a  $\mu$ -automata recognizable set of image-finite models modulo bisimulation is uncountable iff it has the cardinality of the continuum iff it contains a non-B-regular tree. This can be seen as an analogy of the main result in [Niw91], with Rabin tree automata on ranked infinite-trees replaced by  $\mu$ -automata on Kripke models, and isomorphism<sup>4</sup> replaced by bisimulation. The new setting significantly complicates the proof of the main theorem by two reasons:

1. A path in a tree that is accepted by a  $\mu$ -automaton, may correspond to a *tree* in the accepting run, not always a *path* as in the case of Rabin automata on ranked trees.
2. Bisimulation requires much more care than isomorphism, as we will see in the proof of Theorem 6.3.5, where we need infinitely many non-bisimilar trees with complicated construction, while two non-isomorphic trees are enough for the corresponding proof in [Niw91].

To simplify the discussion, we focus on trees without action labels. Therefore we omit  $\Sigma$  in the definition of Kripke models and trees. To generalize the result to the case with finitely many action labels is a standard exercise. In the context of unlabelled trees, it is conventional to say a tree is *image-finite* if it is finitely branching. We also fix a label set  $2^P$  for the trees and omit the label set  $W_{\mathcal{T}} \times (Q_A \cup B_A)$  in runs of  $A$  on a tree  $\mathcal{T}$  when the context is clear.

The following two propositions and the later lemma intend to deal with the first complication we mentioned earlier. First, we show that for image-finite trees the accepting runs (if they exist) can also be image-finite.

**6.3.1. PROPOSITION.** *If an image-finite tree  $\mathcal{T}$  is accepted by  $A$  then there is an image-finite accepting run  $\mathcal{R}'$  of  $A$  on  $\mathcal{T}$ , such that there are no two sibling nodes  $w$  and  $w'$  in  $\mathcal{R}'$  satisfying  $L_{\mathcal{R}'}(w) = L_{\mathcal{R}'}(w')$ .*

**PROOF** Suppose there is an accepting run  $\mathcal{R}$  of  $A$  on  $\mathcal{T}$  and there is a node  $w$  in  $\mathcal{R}$  which has infinitely many successors. It is clear that  $w$  must be labelled by a BR state in  $A$ , thus  $L_{\mathcal{R}}(w) = (v, b)$  for some  $v$  in  $\mathcal{T}$  and  $b \in B_A$ . We define an equivalence relation  $\sim$  among the successors of  $w$  such that  $w' \sim w'' \iff L_{\mathcal{R}}(w') = L_{\mathcal{R}}(w'')$ . With the presence of the Axiom of Choice, we pick one  $w'$  as a representative in each of the equivalence classes w.r.t  $\sim$ . Since  $\mathcal{T}$  is image-finite and  $Q_A$  is finite, there are only finitely many such equivalence classes. A moment's reflection should confirm that we can then prune the successors other than the representatives safely, such that the resulting run  $\mathcal{R}'$  is still an accepting run of  $A$  on  $\mathcal{T}$ .  $\times$

Given a  $\mu$ -automaton  $A$ , we call a run  $\mathcal{R}$  of  $A$  on  $\mathcal{T}$  *non-parallel* if for any  $v \in \mathcal{T}$ :  $Q^{\mathcal{R}}(v)$  is a singleton. We sometimes write  $Q^{\mathcal{R}}(v) = q$  instead of  $Q^{\mathcal{R}}(v) = \{q\}$ .

<sup>4</sup>more precisely, some notion of tree identity.

**6.3.2. PROPOSITION.** *If a  $\mu$ -automaton  $\mathbf{A}$  accepts an image-finite tree  $\mathcal{T}$ , then  $\mathbf{A}$  accepts an image-finite tree  $\mathcal{T}'$  such that there is a non-parallel accepting run of  $\mathbf{A}$  on  $\mathcal{T}'$  and  $\mathcal{T} \simeq \mathcal{T}'$ .*

**PROOF** Suppose  $\mathcal{R} = \{W_{\mathcal{R}}, \rightarrow_{\mathcal{R}}, L_{\mathcal{R}}, r_{\mathcal{R}}\}$  is a successful run of  $\mathcal{T} = \{W_{\mathcal{T}}, \rightarrow_{\mathcal{T}}, L_{\mathcal{T}}, r_{\mathcal{T}}\}$  on  $\mathbf{A}$ . Let  $\mathcal{T}' = \{W_{\mathcal{T}'}, \rightarrow_{\mathcal{T}'}, L_{\mathcal{T}'}, r_{\mathcal{T}'}\}$  where:

- $W_{\mathcal{T}'} = \{w \in W_{\mathcal{R}} \mid L_{\mathcal{R}}(w) = (v, q) \text{ for some } v \in \mathcal{T}, q \in Q_{\mathbf{A}}\}$ .
- $v \rightarrow_{\mathcal{T}'} v' \iff v \xrightarrow{\tau}_{\mathcal{R}} w \rightarrow_{\mathcal{R}} v' \text{ for some } w \in \mathcal{R}$ .
- $L_{\mathcal{T}'}(w) = L_{\mathcal{T}}(v)$  for  $v \in \mathcal{T}$  such that  $L_{\mathcal{R}}(w) = (v, q)$ .
- $r_{\mathcal{T}'} = r_{\mathcal{R}}$ .

It is not hard to see that  $\mathcal{T}'$  is accepted by  $\mathbf{A}$  through a successful run

$$\mathcal{R}' = \{W_{\mathcal{R}'}, \rightarrow_{\mathcal{R}'}, L', r_{\mathcal{R}'}\}$$

where:

$$L'(w) = \begin{cases} (w, q) & \text{if } L_{\mathcal{R}}(w) = (v, q) \text{ for some } v \in \mathcal{T} \\ (w', b) & \text{if } L_{\mathcal{R}}(w) = (v, b) \text{ for some } v \in \mathcal{T} \text{ and } w' \xrightarrow{\tau} w \text{ in } \mathcal{R} \end{cases}$$

Given  $u \in \mathcal{T}'$ , suppose  $L'(w) = (u, q)$  and  $L'(w') = (u, q')$  for some  $w, w' \in \mathcal{R}'$ . By the definition of  $L'$ , we have  $w = u = w'$  thus  $q = q'$ . Therefore, for all  $u \in \mathcal{T}' : Q^{\mathcal{R}'}(u)$  is a singleton.

We now define binary relations  $\sim$  on  $W_{\mathcal{T}} \times W_{\mathcal{T}'}$  as follows:

$$v \sim u \iff L_{\mathcal{R}}(u) = (v, q) \text{ for some } q \in Q_{\mathbf{A}}.$$

We claim  $\sim$  is a bisimulation between  $\mathcal{T}$  and  $\mathcal{T}'$ . For every pair  $(v, u) \in \sim$  with  $L_{\mathcal{R}}(u) = (v, q)$  the three conditions of bisimulation hold:

1.  $L_{\mathcal{T}}(v) = L_{\mathcal{T}'}(u)$ : by definition.
2. Suppose  $v \rightarrow_{\mathcal{T}} v'$ . Then by the definition of the run  $\mathcal{R}$ , there must be a node  $u' \in \mathcal{R}$  such that  $L_{\mathcal{R}}(u') = (v', q')$  for some  $q'$  and  $u \xrightarrow{\tau}_{\mathcal{R}} u'$ . It is clear that  $u \rightarrow_{\mathcal{T}'} u'$  in  $\mathcal{T}'$  and  $v' \sim u'$ .
3. Suppose  $u \rightarrow_{\mathcal{T}'} u'$  in  $\mathcal{T}'$ . Then by the definition of  $\mathcal{T}'$ ,  $u \xrightarrow{\tau}_{\mathcal{R}} u'$  in  $\mathcal{R}$ . By the definition of  $\mathcal{R}$ , there is some  $v' \in \mathcal{T}$  such that  $L_{\mathcal{R}}(u') = (v', q')$  and  $v \rightarrow_{\mathcal{T}} v'$ .

✱

We call a run  $\mathcal{R}$  of  $\mathbf{A}$  on a tree  $\mathcal{T}$  *simple* if  $\mathcal{R}$  is image-finite and  $W^{\mathcal{R}}(v)$  is a singleton for any  $v \in T$ . Now based on Propositions 6.3.1 and 6.3.2, we can prove the following lemma:

**6.3.3. LEMMA.** *If a  $\mu$ -automaton  $\mathbf{A}$  accepts an image-finite tree  $\mathcal{T}$ , then  $\mathbf{A}$  accepts an image-finite tree  $\mathcal{T}'$  such that  $\mathcal{T} \Leftrightarrow \mathcal{T}'$  and there is an accepting simple run  $\mathcal{R}$  of  $\mathbf{A}$  on  $\mathcal{T}'$ . Therefore, a path in  $\mathcal{T}'$  has a unique corresponding path in  $\mathcal{R}$ .*

**PROOF** Given  $\mathcal{T}$ , we first build an image-finite accepting run as in the proof of Proposition 6.3.1. Then we can convert this run into a tree  $\mathcal{T}'$  which is bisimilar to  $\mathcal{T}$  according to Proposition 6.3.2, with a non-parallel run  $\mathcal{R}$ . Now we show that  $W^{\mathcal{R}}(v)$  is a singleton for any  $v \in \mathcal{T}'$ . Suppose not, then there is a node  $v \in \mathcal{T}'$  such that there are  $w$  and  $w' \in \mathcal{R}$  such that  $L_{\mathcal{R}}(w) = (v, q)$  and  $L_{\mathcal{R}}(w') = (v, q)$ . By the definition of the run, it is clear that  $w$  and  $w'$  are at the same level of  $\mathcal{R}$ . According to Proposition 6.3.1,  $w$  and  $w'$  can not be siblings. Therefore there must be a *departure node*  $w_0$  with  $L_{\mathcal{R}}(w_0) = (v', b)$  for some  $v' \in \mathcal{T}'$  and  $b \in B_{\mathbf{A}}$ , such that  $w_0 \rightarrow w_1 \rightarrow^n w$  and  $w_0 \rightarrow w_2 \rightarrow^n w'$  for some  $w_1 \neq w_2$  and a natural number  $n > 0$ . Suppose  $L(w_1) = (v_1, q_1)$  and  $L(w_2) = (v_2, q_2)$ , then it is not hard to see that  $v_1 \neq v_2$ , since  $w_1$  and  $w_2$  are siblings. It is clear that  $v_1$  and  $v_2$  are also siblings in  $\mathcal{T}'$ . However, it is impossible to reach the same point  $v$  from two sibling nodes in  $\mathcal{T}'$ , according to the definition of the trees. Contradiction.

For any  $v \in \mathcal{T}'$ , since  $W^{\mathcal{R}}(v)$  is a singleton, we let  $\mathcal{R}(v)$  be the unique element in  $W^{\mathcal{R}}(v)$ . Now given an infinite path  $P : v_0, v_1, v_2 \dots$  in  $\mathcal{T}'$  we can find the corresponding unique path in  $\mathcal{R} : \mathcal{R}(v_0) \xrightarrow{\tau} w_0 \rightarrow \mathcal{R}(v_1) \xrightarrow{\tau} w_1 \rightarrow \mathcal{R}(v_2) \xrightarrow{\tau} \dots$ , where  $w_i$  are unique successors of  $\mathcal{R}(v_i)$  in  $\mathcal{R}$ .

✕

Before we go to the main theorem, we need the following lemma which helps to provide a source tree to be pumped in the later proof.

**6.3.4. LEMMA.** *If an image-finite tree  $\mathcal{T}$  is not B-regular then there is an infinite path  $P = v_0, v_1, \dots$  such that:*

$$\text{for any } k \in \mathbb{N}, \text{ any } v \text{ such that } \text{dep}(\mathcal{T}, v) < \text{dep}(\mathcal{T}, v_k): \mathcal{T}^{v_k} \not\Leftarrow \mathcal{T}^v.$$

*Namely, for each  $v_k$ ,  $\mathcal{T}^{v_k}$  is a “new” subtree which does not appear up to bisimulation at earlier levels of the tree.*

**PROOF** Let  $V = \{v \in \mathcal{T} \mid \forall w \in \mathcal{T} : \mathcal{T}^v \Leftrightarrow \mathcal{T}^w \implies \text{dep}(\mathcal{T}, v) \leq \text{dep}(\mathcal{T}, w)\}$ . Intuitively,  $V$  is the set of nodes where a “new” tree appears, in top-bottom order. Let  $V' = \{v : v \rightarrow^* w \text{ in } \mathcal{T}, \text{ for some } w \in V\}$ , then  $\mathcal{T}' = \{V', \rightarrow|_{V' \times V'}, L|_{V'}\}$  is a tree. Due to the fact that  $\mathcal{T}$  is image-finite and non-B-regular,  $\mathcal{T}'$  is an image-finite infinite tree. With the presence of Axiom of Choice, we recall König's Lemma on unordered trees:

An image-finite infinite tree has an infinite path.

Thus there is an infinite path  $P = v_0, v_1, v_2, \dots$  in  $\mathcal{T}'$ . Clearly  $P$  is also an infinite path in  $\mathcal{T}$ . Now suppose towards contradiction that there are  $v_k \in P$  and  $v \in \mathcal{T}$  such that  $\text{dep}(\mathcal{T}, v) < \text{dep}(\mathcal{T}, v_k)$  and  $\mathcal{T}^{v_k} \Leftrightarrow \mathcal{T}^v$ . By the definition of  $V'$ , there is a  $w$  such that  $v_k \rightarrow^n w$  and  $w \in V$  for some  $n$ . Since  $\mathcal{T}^v \Leftrightarrow \mathcal{T}^{v_k}$  then there is a  $w' : v \rightarrow^n w'$  in  $\mathcal{T}$

such that  $\mathcal{T}^w \leftrightarrow \mathcal{T}^{w'}$ . However, it is clear that  $\text{dep}(\mathcal{T}, w') < \text{dep}(\mathcal{T}, w)$ , thus  $w \notin V$ , contradiction.  $\spadesuit$

Now we come to our main theorem. Note that we only consider image-finite models in the sequel.

**6.3.5. THEOREM.** *Let  $\mathbf{A}$  be a  $\mu$ -automaton. Then the following are equivalent:*

1.  $|\mathcal{L}(\mathbf{A})|_{/\leftrightarrow} = 2^{\aleph_0}$ ,
2.  $|\mathcal{L}(\mathbf{A})|_{/\leftrightarrow} > \aleph_0$ ,
3.  $\mathcal{L}(\mathbf{A})$  contains a non-B-regular tree.

**PROOF** (1)  $\implies$  (2) is straightforward.

(2)  $\implies$  (3) : Suppose  $\mathcal{L}(\mathbf{A})$  only contains B-regular trees. Then by the definition of B-regular trees, each tree in  $\mathcal{L}(\mathbf{A})$  is bisimilar to its bisimulation contraction, which is finite. However, there are only countably many such finite Kripke models, given a finite set  $\mathbf{P}$  of basic propositions.

(3)  $\implies$  (1) : Observe that each  $\omega$ -branching tree<sup>5</sup> can be viewed as a downward closed subset of  $\mathbb{N}^*$ . Since  $|\mathbb{N}^*| = \aleph_0$ , clearly  $|\mathcal{L}(\mathbf{A})|_{/\leftrightarrow} \leq 2^{\aleph_0}$ . We will prove (3)  $\implies |\mathcal{L}(\mathbf{A})|_{/\leftrightarrow} \geq 2^{\aleph_0}$ . The idea of the proof is to *pump*  $2^{\aleph_0}$  many non-bisimilar acceptable trees out of a non-B-regular tree.

Suppose  $\mathcal{L}(\mathbf{A})$  contains an image-finite non-B-regular tree  $\mathcal{T}^o$ . By Lemma 6.3.3,  $\mathcal{L}(\mathbf{A})$  contains an image-finite tree  $\mathcal{T} = (W, \rightarrow, L, r)$  bisimilar to  $\mathcal{T}^o$ , such that there is an accepting run  $\mathcal{R}$  of  $\mathbf{A}$  on  $\mathcal{T}$  and  $W^{\mathcal{R}}(v)$  is a singleton for each  $v \in \mathcal{T}$ . Clearly,  $\mathcal{T}$  is also non-B-regular.

By Lemma 6.3.4, there is an infinite path  $P : v_0 = r, v_1, v_2, \dots$  such that for any  $k \in \mathbb{N}$  and any  $v \in \mathcal{T} : \text{dep}(\mathcal{T}, v) < \text{dep}(\mathcal{T}, v_k) \implies \mathcal{T}^{v_k} \not\leftrightarrow \mathcal{T}^v$ . It is obvious that for any  $j \neq i : \mathcal{T}^{v_i} \not\leftrightarrow \mathcal{T}^{v_j}$ . Since  $\mathcal{R}$  is non-parallel,  $Q_{\mathcal{R}}(v_i)$  is a singleton for any  $v_i \in P$ .

We now pick a *distant* node  $v_m \in P$  such that  $Q^{\mathcal{R}}(v_m) = \{q\} \subseteq \text{Inf}(\mathcal{R}, P)$  for some  $q \in Q_{\mathbf{A}}$  and  $Q^{\mathcal{R}}(P^{v_m}) = \text{Inf}(\mathcal{R}, P)$ , where  $P^{v_m}$  is the suffix of  $P$  starting at  $v_m$ . Intuitively, we pick  $v_m$  in such a way that all the points in  $P$  after  $v_m$  are only matched with the states in the automaton that appear infinitely often according to the labelling of  $P$  in  $\mathcal{R}$ . We then find an infinite subsequence  $P' : v'_0, v'_1, v'_2, \dots$  of  $P$  such that  $v'_0 = v_m$  and for each  $k : Q^{\mathcal{R}}(v'_{k+1}) = \{q\}$  and  $Q^{\mathcal{R}}(P^{v'_k}) = \text{Inf}(\mathcal{R}, P)$ , where  $P^{v'_k}$  is the segment of  $P$  between  $v'_k$  and  $v'_{k+1}$ . Note that such  $P'$  exists since  $q \in \text{Inf}(\mathcal{R}, P)$ .

Now we are ready to pump the tree  $\mathcal{T}$  into  $2^{\aleph_0}$  many non-bisimilar trees. For each infinite sequence  $\alpha$  of 0s and 1s, we construct a tree  $\mathcal{T}_{\alpha}$  which is accepted by  $\mathbf{A}$ . We do this by building the sequence of triples  $(\mathcal{T}_{\alpha_n}, u_{\alpha_n}, s_{\alpha_n})$  where  $\alpha_n$  is a prefix of  $\alpha$  of the length  $n$  and  $u_{\alpha_n}, s_{\alpha_n} \in P'$ . Intuitively  $u_{\alpha_n}$  is the “replacing point” and  $\mathcal{T}^{s_{\alpha_n}}$  is the “substitution”.

<sup>5</sup> $\omega$ -branching trees are the trees that have at most  $\aleph_0$  many successors at each node.

Before defining the pumped trees formally, recall Fact 2.2.5: if two pointed models  $(\mathcal{M}, s_0)$  and  $(\mathcal{N}, t_0)$  are not bisimilar then Spoiler has a winning strategy in some  $n$ -round bisimulation game  $\mathcal{G}_n((\mathcal{M}, s), (\mathcal{N}, t))$ . We let  $g_{\not\sim}((\mathcal{M}, s), (\mathcal{N}, t))$  be the minimal number  $n_0$  such that the spoiler has a winning strategy for the  $n_0$ -round bisimulation game  $\mathcal{G}_{n_0}((\mathcal{M}, s), (\mathcal{N}, t))$ .

We start from  $\mathcal{T}_\epsilon = \mathcal{T}$ ,  $u_\epsilon = v'_0$ ,  $s_\epsilon = v'_1$ . For any finite binary sequence  $\beta$ ,  $\mathcal{T}_{\beta \cdot 0}$  and  $\mathcal{T}_{\beta \cdot 1}$  are constructed as follows:

- $\mathcal{T}_{\beta \cdot 1} = \mathcal{T}_\beta[u_\beta \setminus s_\beta]$ ,  $u_{\beta \cdot 1} = \text{next}(s_\beta, D(\mathcal{T}_\beta, u_\beta, s_\beta))$  and  $s_{\beta \cdot 1} = \text{next}(u_{\beta \cdot 1}, 0)$ .
- $\mathcal{T}_{\beta \cdot 0} = \mathcal{T}_\beta$ ,  $u_{\beta \cdot 0} = u_\beta$  and  $s_{\beta \cdot 0} = s_\beta$ .

where

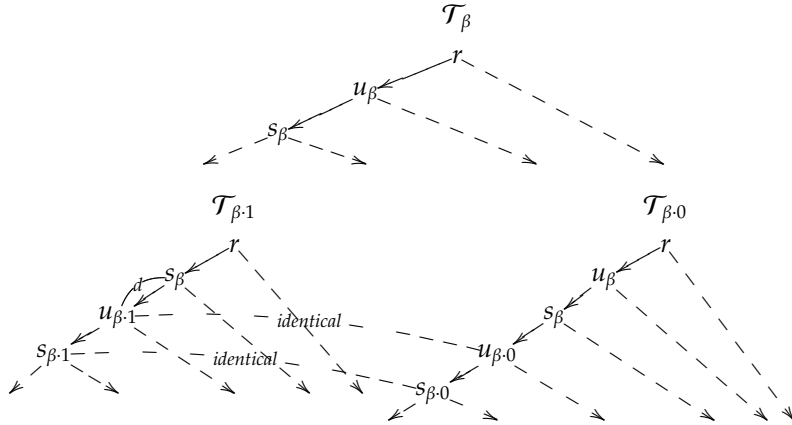
- for  $u, s$  in  $P'$  such that  $\text{dep}(\mathcal{T}, u) < \text{dep}(\mathcal{T}, s)$ ,  $D(\mathcal{T}, u, s)$  is the depth of the zone where the non-bisimilarity between  $\mathcal{T}^s$  and any other subtree at the same level of  $\mathcal{T}^u$  can be detected (by bisimulation games). Formally:

$$D(\mathcal{T}, u, s) = \max\{g_{\not\sim}(\mathcal{T}^{u'}, \mathcal{T}^s) \mid \text{dep}(\mathcal{T}, u') = \text{dep}(\mathcal{T}, u)\}.$$

- for  $s$  in  $P'$ ,  $n \in \mathbb{N}$ :  $\text{next}(s, n)$  is defined as  $v'_k$  in  $P'$  such that  $k$  is the minimal index satisfying  $\text{dep}(\mathcal{T}, v'_k) - \text{dep}(\mathcal{T}, s) > n$ . Intuitively,  $\text{next}(s, n)$  is the next point in  $P'$  after  $s$ , such that the zone with depth  $n$  can be preserved.

Note that  $g_{\not\sim}(\mathcal{T}^{u'}, \mathcal{T}^s)$  is well-defined since for any  $s$  on  $P'$ ,  $\text{dep}(\mathcal{T}, u) < \text{dep}(\mathcal{T}, s)$  implies  $\mathcal{T}^u \not\sim \mathcal{T}^s$ . Moreover  $\{g_{\not\sim}(\mathcal{T}^{u'}, \mathcal{T}^s) \mid \text{dep}(\mathcal{T}, u') = \text{dep}(\mathcal{T}, u)\}$  is a finite set since  $\mathcal{T}$  is image-finite, thus the maximal element of this set exists.

Let  $d = D(\mathcal{T}_\beta, u_\beta, s_\beta)$ , the intuition behind the above construction is illustrated as follows:



We build  $\mathcal{T}_{\beta \cdot 1}$  by placing the substitution  $\mathcal{T}^{s_\beta}$  at the replacing point  $u_\beta$ . Then we let the next replacing point  $u_{\beta \cdot 1}$  be far away enough from the previous substitution point  $s_\beta$ , such that the non-bisimilarity of  $\mathcal{T}^{s_\beta}$  and its new neighbour subtrees at the same

level in  $\mathcal{T}_\beta[u_\beta \setminus s_\beta]$  can be detected by a bisimulation game before reaching the new replacing point. Finally we let the new substitution be the “next” suitable point after  $s_{\beta-1}$  in  $P'$ . For  $\mathcal{T}_{\beta-0}$  we simply do not execute the substitution but change the  $s_{\beta-0}$  and  $u_{\beta-0}$  as in the case of  $\mathcal{T}_{\beta-1}$ .

Given an infinite binary sequence  $\alpha$ , we can now build  $\mathcal{T}_{\alpha_n}$  for each  $n$ . To build  $\mathcal{T}_\alpha$ , we define the *stable* part of  $\mathcal{T}_{\alpha_n}$  as follows:

- *stable domain*:  $sdom(\mathcal{T}_{\alpha_n}) = W_{\mathcal{T}_{\alpha_n}} - \{v \mid u_{\alpha_n} \rightarrow^* v \text{ in } \mathcal{T}_{\alpha_n}\}$
- *stable edges*:  $sedg(\mathcal{T}_{\alpha_n}) = \rightarrow_{\mathcal{T}_{\alpha_n}} \upharpoonright_{sdom(\mathcal{T}_{\alpha_n}) \times sdom(\mathcal{T}_{\alpha_n})}$ .
- *stable label*:  $slabel(\mathcal{T}_{\alpha_n}) = L_{\mathcal{T}_{\alpha_n}} \upharpoonright_{sdom(\mathcal{T}_{\alpha_n})}$ .

It is not hard to see that the above defined stable part of  $\mathcal{T}_{\alpha_n}$  does not get altered in  $\mathcal{T}_{\alpha_m}$  for  $m > n$ . Due to such monotonicity, we can now build the limit tree:

$$\mathcal{T}_\alpha = \left( \bigcup_{n < \omega} sdom(\mathcal{T}_{\alpha_n}), \bigcup_{n < \omega} sedg(\mathcal{T}_{\alpha_n}), \bigcup_{n < \omega} slabel(\mathcal{T}_{\alpha_n}) \right)$$

Since  $\mathcal{R}$  is an accepting run of  $\mathbf{A}$  on  $\mathcal{T}$  such that  $W^{\mathcal{R}}(v)$  is a singleton for each  $v \in \mathcal{T}$ , every node in  $\mathcal{T}$  corresponds to a two-node path  $w \xrightarrow{\tau} w'$  in  $\mathcal{R}$  such that  $L_{\mathcal{R}}(w) = (v, q)$  and  $L_{\mathcal{R}}(w') = (v, b)$  for some  $b \in B_{\mathbf{A}}$  and  $\{q\} = Q^{\mathcal{R}}(w)$ . Then it is easy to see that we can build each  $\mathcal{R}_{\alpha_n}$  based on  $\mathcal{R}_{\alpha_{n-1}}$ , by the corresponding substitution as in the construction for  $\mathcal{T}_{\alpha_n}$ . Such  $\mathcal{R}_{\alpha_n}$  is indeed a run of  $\mathcal{T}_{\alpha_n}$  since the replacing point and the substitution point in  $\mathcal{T}_{\alpha_{n-1}}$  are all labelled with the same  $q \in Q_{\mathbf{A}}$ . Similarly, we can define the stable parts of each  $\mathcal{R}_{\alpha_n}$  and let:

$$\mathcal{R}_\alpha = \left( \bigcup_{n < \omega} sdom(\mathcal{R}_{\alpha_n}), \bigcup_{n < \omega} sedg(\mathcal{R}_{\alpha_n}), \bigcup_{n < \omega} slabel(\mathcal{R}_{\alpha_n}) \right)$$

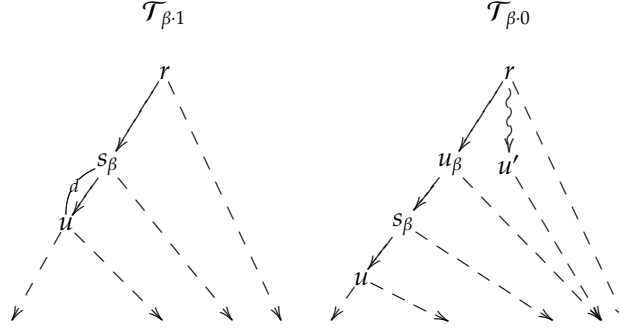
To see the limit run  $\mathcal{R}_\alpha$  is also an accepting run of  $\mathbf{A}$  on  $\mathcal{T}_\alpha$ , we need to check the parity condition for every infinite path in  $\mathcal{R}_\alpha$ .

Note that if an infinite path  $P_1$  in  $\mathcal{R}_\alpha$  is contained in a stable part of  $\mathcal{R}_{\alpha_n}$  for some  $n$ , then we can find a path  $P'_1$  in  $\mathcal{R}$  such that  $P_1$  and  $P'_1$  share an infinite suffix. Then it is clear that  $P_1$  and  $P'_1$  only differ in their finite prefixes, thus  $\Omega^{\mathcal{R}_\alpha}(P_1) = \Omega^{\mathcal{R}_{\alpha_n}}(P_1) = \Omega^{\mathcal{R}}(P'_1)$  which is even. The non-trivial case is the *limit path*  $P_\alpha$  which is not contained in the stable part of any  $\mathcal{T}_{\alpha_n}$ , thus goes through infinitely many substitution points  $r, u_{\alpha_0}, u_{\alpha_1}, u_{\alpha_2} \dots$ . Since  $Q_{\mathcal{R}}(P^{v_0}) = Inf(\mathcal{R}, P)$  and  $Q_{\mathcal{R}}(P^{v'_k}) = Inf(\mathcal{R}, P)$  for any  $k < \omega$ , then the construction of the limit run can neither make any  $q' \notin Inf(\mathcal{R}, P)$  occur infinitely often nor make any  $q' \in Inf(\mathcal{R}, P)$  occur only finitely often. Therefore  $\Omega^{\mathcal{R}_\alpha}(P_\alpha) = \Omega^{\mathcal{R}}(P)$ . In sum,  $\mathcal{R}_\alpha$  is indeed an accepting run of  $\mathbf{A}$  on  $\mathcal{T}_\alpha$ .

To complete the above proof, we need to show that for any  $\alpha \neq \alpha' \in 2^\omega$ ,  $\mathcal{T}_\alpha \not\sim \mathcal{T}_{\alpha'}$ , which is proved by the following lemma. ✱

**6.3.6. LEMMA.** *If  $\alpha \neq \alpha' \in 2^\omega$ , then Spoiler can win the bisimulation game  $G_n(\mathcal{T}_\alpha, \mathcal{T}_{\alpha'})$  for some  $n < \omega$ .*

PROOF Since  $\alpha \neq \alpha'$ , there is a sequence  $\beta \in 2^*$  such that  $\alpha_m = \beta \cdot 0$  and  $\alpha'_m = \beta \cdot 1$  for some  $m$ . We recall the construction of  $\mathcal{T}_{\beta \cdot 1}$  and  $\mathcal{T}_{\beta \cdot 0}$  as follows:



where  $u = u_{\beta \cdot 0} = u_{\beta \cdot 1}$ . Note that in both trees, the parts which are not reachable from  $u$  will be preserved in  $\mathcal{T}_\alpha$  and  $\mathcal{T}_{\alpha'}$ , according to the construction of the limit tree. We claim that there is a winning strategy for spoiler in the game  $\mathcal{G}_n(\mathcal{T}_\alpha, \mathcal{T}_{\alpha'})$  where  $n = \text{dep}(\mathcal{T}_{\beta \cdot 1}, u) = \text{dep}(\mathcal{T}_\beta, u_\beta) + d$ .

It is clear that Spoiler has a strategy to reach the node  $s_\beta$  in  $\mathcal{T}'_\alpha$  by following a path in  $\mathcal{T}'_\alpha$  from  $r$ . To match Spoiler's moves, the verifier, if it has not lost already, will reach another point  $u'$  in  $\mathcal{T}_\alpha$  with the same depth. Since  $d > \max\{g_{\frac{d}{2}}(\mathcal{T}^v, \mathcal{T}^{s_\beta}) \mid \text{dep}(\mathcal{T}, v) = \text{dep}(\mathcal{T}, u_\beta)\}$  as in the construction of  $\mathcal{T}_{\beta \cdot 1}$ , Spoiler can win the bisimulation game  $\mathcal{G}_d(\mathcal{T}^{u'}, \mathcal{T}^{s_\beta})$ . Therefore Spoiler has a strategy to show the difference between  $\mathcal{T}^{u'}$  and  $\mathcal{T}^{s_\beta}$  in  $d$  steps, namely within the stable parts of  $\mathcal{T}_{\beta \cdot 1}$  and  $\mathcal{T}_{\beta \cdot 0}$ . Thus Spoiler can win the bisimulation game  $\mathcal{G}_d(\mathcal{T}'_\alpha, \mathcal{T}^{s_\beta})$ . In sum, Spoiler has a winning strategy for the game  $\mathcal{G}_n(\mathcal{T}_\alpha, \mathcal{T}_{\alpha'})$ .  $\star$

## 6.4 Normal Form of the Countable Languages

Following [Niw91], we call a tree  $\mathcal{T}$  *alive* if  $\mathcal{T}_{\downarrow \Leftrightarrow} \in \text{Sub}_{\Leftrightarrow}(\mathcal{T})$ , namely  $\mathcal{T}$  can *regenerate* itself up to bisimulation. Given a  $\mu$ -automaton  $\mathbf{A}$ , let  $\text{alive}(\mathcal{L}(\mathbf{A}))$  be the collection of all the alive subtrees of the trees accepted by  $\mathbf{A}$ :

$$\text{alive}(\mathcal{L}(\mathbf{A})) = \{\mathcal{T} \mid \mathcal{T} \text{ is an alive subtree of some } \mathcal{T}' \in \mathcal{L}(\mathbf{A})\}.$$

It is easy to see that a tree is alive if and only if there is a cycle in its bisimulation contraction starting from  $[r]_{\downarrow \Leftrightarrow}$ . An infinite path  $P$  is called a *regenerating path* of an alive tree  $\mathcal{T}$ , if  $P$  goes through infinitely many nodes:  $v_0, v_1 \dots$  such that each  $\mathcal{T}^{v_i} \Leftrightarrow \mathcal{T}$ . We call those  $v_i$  *regenerating points* of  $P$ . It is clear that every alive tree has at least one regenerating path starting at the root. We say an alive tree  $\mathcal{T}$  is *q-lively-accepted* by  $\mathbf{A}$ , if there is an accepting simple  $q$ -run  $\mathcal{R}$  of  $\mathbf{A}$  on  $\mathcal{T}$ , such that there are infinitely many regenerating points  $v_0, v_1 \dots$  in a regenerating path st.  $Q^{\mathcal{R}}(v_i) = \{q\}$  for each  $i \in \mathbb{N}$ .

**6.4.1. PROPOSITION.** *For every tree  $\mathcal{T}_0 \in \text{alive}(\mathcal{L}(\mathbf{A}))$ , there is a tree  $\mathcal{T}_1 \Leftrightarrow \mathcal{T}_0$  such that  $\mathcal{T}_1$  is  $q$ -lively-accepted by  $\mathbf{A}$  for some state  $q \in Q_{\mathbf{A}}$  which is reachable from the start state  $q_0$  of  $\mathbf{A}$ .*

**PROOF** Given a tree  $\mathcal{T}_0 \in \text{alive}(\mathcal{L}(\mathbf{A}))$ , let  $\mathcal{T} \in \mathcal{L}(\mathbf{A})$  be a tree such that  $\mathcal{T}_0$  is a subtree of  $\mathcal{T}$ . By Lemma 6.3.3, there is a tree  $\mathcal{T}' \Leftrightarrow \mathcal{T}$  such that there is a  $q_0$ -run  $\mathcal{R}'$  which is accepting and  $W^{\mathcal{R}'}(v)$  is a singleton for each  $v \in \mathcal{T}'$ . It is clear that  $\mathcal{T}'$  has an alive subtree  $\mathcal{T}'_0$  which is bisimilar to  $\mathcal{T}_0$ . Then there is an infinite regenerating path  $P$  starting at the root of  $\mathcal{T}'_0$ . Consider the corresponding path  $P'$  in  $\mathcal{R}'$ , it is easy to see that there are infinitely many nodes corresponding to the regenerating points in  $P$  which are labelled with the same state  $q \in Q_{\mathbf{A}}$ . ✱

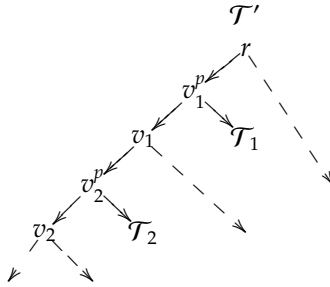
We now show a countable language only involves finitely many alive trees up to bisimulation. Differing from the case of Rabin Automata on ranked trees in [Niw91], our  $\mu$ -automata work on unranked trees, thus we can now construct new acceptable trees by *inserting* branches to an existing acceptable tree, as shown in the proof of the following lemma.

**6.4.2. LEMMA.** *Given a  $\mu$ -automaton  $\mathbf{A}$ , if  $\mathcal{L}(\mathbf{A})$  is countable up to bisimulation then  $|\text{alive}(\mathcal{L}(\mathbf{A}))|_{\Leftrightarrow}$  is finite.*

**PROOF** Suppose towards contradiction that  $\mathcal{L}(\mathbf{A})$  is countable but  $|\text{alive}(\mathcal{L}(\mathbf{A}))|_{\Leftrightarrow}$  is infinite. By Proposition 6.4.1 and the pigeon hole argument, there are infinitely many non-bisimilar alive trees  $\mathcal{T}_0, \mathcal{T}_1, \mathcal{T}_2, \dots$  that are  $q$ -lively accepted by  $\mathbf{A}$  for some  $q \in Q_{\mathbf{A}}$  which is reachable from the start state of  $\mathbf{A}$ .

By Proposition 6.4.1 let  $\mathcal{R}$  be a legal simple run of  $\mathbf{A}$  on a tree  $\mathcal{T}$  with some node  $v \in \mathcal{T}$  such that  $Q^{\mathcal{R}}(v) = \{q\}$ . Then  $\mathcal{T}[v \setminus \mathcal{T}_0]$  is accepted by the simple run  $\mathcal{R}[W^{\mathcal{R}}(v) \setminus \mathcal{R}_0]$ , where  $\mathcal{R}_0$  is the  $q$ -alive-accepting run of  $\mathbf{A}$  on  $\mathcal{T}_0$ .

Since  $\mathcal{T}_0$  is alive, we can find an infinite path in  $\mathcal{T}[v \setminus \mathcal{T}_0]$  containing an infinite subsequence of regenerating points  $P : v_1, v_2, \dots$  where  $\mathcal{T}^{v_i} \Leftrightarrow \mathcal{T}_0$  for each  $i \in \mathbb{N}$ . Based on  $\mathcal{T}_0$ , we now build a non-B-regular tree  $\mathcal{T}'$  by "inserting"  $\mathcal{T}_i$  as a child of the parent node of  $v_i$  for each  $i > 0$ :



where  $v_i^p$  is the parent node of  $v_i$ . Take the simple run  $\mathcal{R}[W^{\mathcal{R}}(v) \setminus \mathcal{R}_0]$  of  $\mathbf{A}$  on  $\mathcal{T}[v \setminus \mathcal{T}_0]$ . Since  $\mathcal{T}_i$  are all  $q$ -accepted, we can build a run  $\mathcal{R}'$  for  $\mathcal{T}'$  by inserting the corresponding

$q$ -run  $\mathcal{R}_i$  of  $\mathcal{T}_i$  as a child of the unique  $w_i$  in  $\mathcal{R}'$  such that  $L_{\mathcal{R}}(w_i) = (v_i^p, b)$  for some  $b$  in  $B_A$ .

It is not hard to see that  $\mathcal{R}'$  is legal run and all the infinite paths satisfy the parity condition thus  $\mathcal{R}'$  is accepting. Thus  $\mathcal{L}(A)$  contains a non-B-regular tree. But then by Theorem 6.3.5,  $\mathcal{L}(A)$  is uncountable. Contradiction.  $\star$

Let  $\mathfrak{F}_n$  be the set of finite trees with some leaves possibly labelled by variables  $x_1, x_2, \dots, x_n$ . Given a  $\mathcal{T}_f \in \mathfrak{F}_n$ , let  $\mathcal{T}_f[x_1 \setminus \mathcal{T}_1, \dots, x_1 \setminus \mathcal{T}_n]$  be the tree obtained by replacing each  $x_i$ -labelled node in  $\mathcal{T}_f$  by the tree  $\mathcal{T}_i$ . We can show that each regular tree can be turned into a normal form:

**6.4.3. LEMMA.** *If  $\mathcal{T}$  is an image-finite B-regular tree, then there exist alive trees  $\mathcal{T}_1, \dots, \mathcal{T}_n$  and a  $\mathcal{T}_f \in \mathfrak{F}_n$  for some  $n \in \mathbb{N}$  such that:  $\mathcal{T} \Leftrightarrow \mathcal{T}_f[x_1 \setminus \mathcal{T}_1, \dots, x_1 \setminus \mathcal{T}_n]$ .*

**PROOF** Given a B-regular tree  $\mathcal{T}$ , suppose there are  $n$  different subtrees modulo bisimulation (call them  $\mathcal{T}_1, \dots, \mathcal{T}_n$ ). We now turn  $\mathcal{T}$  into the normal form by “rewriting” the first  $n$  levels of  $\mathcal{T}$ , in the top-bottom fashion starting from the root: if we reach a node  $v$  such that  $\mathcal{T}^v$  is bisimilar to some alive  $\mathcal{T}_i$ , then we turn  $v$  into  $x_i$  and discard all the nodes reachable from  $v$ . Since  $\mathcal{T}$  is image-finite, we only need to rewrite finitely many nodes. We claim that if we reach some node at level  $n$  which has not been discarded in an earlier state of the rewriting, then this node is a leaf in the tree  $\mathcal{T}$ . Suppose not, then there is a child  $w$  of this node. Since  $\text{dep}(\mathcal{T}, w) > n$ , along the path from the root to  $n$  there must be a  $w'$  such that  $\mathcal{T}^w \Leftrightarrow \mathcal{T}^{w'}$ , due to the fact that there are only  $n$  different subtrees modulo bisimulation. However, then  $\mathcal{T}^{w'}$  is an alive tree and thus the nodes reachable from  $w'$  should be discarded. Contradiction.

Let  $\mathcal{T}_f$  be the resulting tree. It is clear that  $\mathcal{T} \Leftrightarrow \mathcal{T}_f[x_1 \setminus \mathcal{T}_1, \dots, x_1 \setminus \mathcal{T}_n]$ .  $\star$

Given  $F_n \subseteq \mathfrak{F}_n$ , we let  $F_n[x_1 \setminus \mathcal{T}_1, \dots, x_1 \setminus \mathcal{T}_n] = \{\mathcal{T}[x_1 \setminus \mathcal{T}_1, \dots, x_1 \setminus \mathcal{T}_n] \mid \mathcal{T} \in F_n\}$ . We can now show the normal form theorem as follows:

**6.4.4. THEOREM.** *Given a  $\mu$ -automaton  $A$ , if  $\mathcal{L}(A)$  is countable up to bisimulation, then it can be represented by*

$$F_n[x_1 \setminus \mathcal{T}_1, \dots, x_1 \setminus \mathcal{T}_n]$$

for some  $n < \omega$ ,  $\{\mathcal{T}_1, \dots, \mathcal{T}_n\} \subseteq \text{alive}(\mathcal{L}(A))$ , and some  $F_n \subseteq \mathfrak{F}_n$  which is recognizable by an finite automaton  $B$  on finite trees in  $\mathfrak{F}_n$ .

**PROOF** (Sketch) From Lemma 6.4.2, we can list the finitely many different representatives of the alive trees in  $\mathcal{L}(A)$  as  $\mathcal{T}_1, \dots, \mathcal{T}_n$  for some  $n$ . We now build the finite automaton  $B$  on finite unranked trees based on  $A$ . Let  $B = (Q_A, B_A \cup \{x_1, \dots, x_n\}, q_0, \rightarrow_{\text{OR}}^A \cup \rightarrow'_{\text{OR}}, \rightarrow_{\text{BR}}^A, L')$  where:

$$q \rightarrow'_{\text{OR}} x_i \text{ iff there is an accepting } q\text{-run of } A \text{ on } \mathcal{T}_i$$

and

$$L'(b) = \begin{cases} L_A(b) & \text{if } b \in B_A \\ x_i & \text{if } b = x_i \end{cases}$$

We say a finite tree  $\mathcal{T}_f$  with variables is accepted by  $\mathbf{B}$ , if there is a legal run of  $\mathbf{B}$  as a  $\mu$ -automaton on  $\mathcal{T}_f$ . Let  $\mathcal{L}_f(\mathbf{B})$  be the set of trees in  $\mathfrak{F}_n$  that are accepted by  $\mathbf{B}$ . It is not hard to verify that  $\mathcal{L}(\mathbf{A})$  is equivalent (up to bisimulation) to  $\mathcal{L}(\mathbf{B})[x_1 \setminus \mathcal{T}_1, \dots, x_n \setminus \mathcal{T}_n]$ .  $\star$

## 6.5 Discussion and Future Work

This chapter extends a result by Niwiński [Niw91] on the cardinality of tree languages recognized by automata. We showed that a  $\mu$ -automata recognizable set of image-finite models modulo bisimulation is uncountable iff it is of the cardinality continuum iff it contains a non-B-regular tree. As in [Niw91], we give a normal form of the countable languages modulo bisimulation.

A straightforward consequence of Theorem 6.3.5 is that a  $\mu$ -formula has countably many image-finite bisimulation contracted models iff all these models are finite. Another interesting consequence in the case of countable languages is implied by Lemma 6.4.2: there are only finitely many non-bisimilar *S5* or *KD45* models of a  $\mu$  formula, if it has only finite bisimulation contracted models. To see this, first note that the *S5* models are reflexive, thus their unravellings are alive trees themselves, therefore there are only finitely many of them due to Lemma 6.4.2. For the case of *KD45* models, observe that the only state that is not reflexive for any labelled relations in a connected *KD45* pointed model can only be the designated state. Therefore, any *KD45* model of the given formula can be generated (modulo bisimulation) by linking an start state to some of the alive trees. If the set of labels and the set of basic propositions are finite, then we only have finitely many such *KD45* models.

However, it should be noted that our main theorem does not imply the following: a  $\mu$ -automata recognizable set of image-finite *S5* models modulo bisimulation is uncountable iff it is of the cardinality continuum iff it contains a non-B-regular tree. Our pumping construction in the proof of Theorem 6.3.5 does not preserve *S5* conditions, i.e., we may pump a tree which is not bisimilar to any *S5* models from an unravelling of a *S5* model. Interested readers may try to verify that transitivity is not preserved by our pumping construction in this sense. This calls for a closer look at finer classes of models which may require more sophisticated pumping constructions. In fact, the proof of Lemma 6.4.2 already suggests that a pumping construction which *adds* trees instead of *substituting* trees may also work to prove the main theorem in a way that may preserve the properties. We leave this for future work.

On the other hand, we may also look at finer classes of automata (or, say, classes of modal  $\mu$ -calculus formulas) to see whether we can have a better understanding for specific fragments of  $\mu$ . An interesting question is to characterize the maximal fragment of  $\mu$  in which each formula has a unique model up to bisimulation.

Note that in the proof of our Lemma 6.4.2, we did not give a bound on the number of alive trees as in [Niw91]. It is not yet clear how we can tighten the proof in order to obtain a finite bound. The difficulty actually lies in the proof of our main theorem,

where using infinitely many non-bisimilar trees is essential, while in [Niw91] two “different” trees are enough. This may be an obstacle to an algorithm, as an analogy to the algorithm given in [Niw91], to output the number of non-bisimilar models of a given modal  $\mu$ -calculus formula. We also leave this for future work.