

AN ALGORITHM TO QUANTIFY BEHAVIOURAL SIMILARITY  
BETWEEN PROBABILISTIC SYSTEMS

BABITA SHARMA

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES  
IN PARTIAL FULFILMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF

MASTER OF SCIENCE

GRADUATE PROGRAM IN COMPUTER SCIENCE  
YORK UNIVERSITY  
TORONTO, CANADA

DECEMBER 2006

**AN ALGORITHM TO QUANTIFY BEHAVIOURAL  
SIMILARITY BETWEEN PROBABILISTIC SYSTEMS**

by **Babita Sharma**

a thesis submitted to the Faculty of Graduate Studies of York University in partial fulfilment of the requirements for the degree of

**MASTER OF SCIENCE**

© 2006

Permission has been granted to: a) YORK UNIVERSITY LIBRARIES to lend or sell copies of this dissertation in paper, microform or electronic formats, and b) LIBRARY AND ARCHIVES CANADA to reproduce, lend, distribute, or sell copies of this thesis anywhere in the world in microform, paper or electronic formats *and* to authorise or procure the reproduction, loan, distribution or sale of copies of this thesis anywhere in the world in microform, paper or electronic formats.

The author reserves other publication rights, and neither the thesis nor extensive extracts for it may be printed or otherwise reproduced without the author's written permission.

**AN ALGORITHM TO QUANTIFY BEHAVIOURAL SIMILARITY BETWEEN  
PROBABILISTIC SYSTEMS**

by **Babita Sharma**

By virtue of submitting this document electronically, the author certifies that this is a true electronic equivalent of the copy of the thesis approved by York University for the award of the degree. No alteration of the content has occurred and if there are any minor variations in formatting, they are as a result of the conversion to Adobe Acrobat format (or similar software application).

Examination Committee Members:

1. Franck van Breugel
2. Jeff Edmonds
3. Parke Godfrey
4. Stephen Watson

# Abstract

This thesis presents research work in the area of concurrency theory and probabilistic systems. Behavioural pseudometrics allow us to provide a quantitative notion of equivalence for systems involving quantitative data. Desharnais, Gupta, Jagadeesan and Panangaden introduced a family of behavioural pseudometrics for probabilistic transition systems. These pseudometrics are a quantitative analogue of probabilistic bisimilarity. Distance zero captures probabilistic bisimilarity. Each pseudometric has a discount factor, a real number in the interval  $(0, 1]$ . The smaller the discount factor, the more the future is discounted. If the discount factor is one, then the future is not discounted at all. Desharnais et al. showed that the behavioural distances can be calculated up to any desired degree of accuracy if the discount factor is smaller than one. In this thesis, we show that the distances can also be approximated if the future is not discounted. We present the first algorithm to approximate distances in the undiscounted setting and thereby solve a problem which had been open since 1999. A key ingredient of our algorithm is Tarski's decision procedure for the first order theory over reals. By exploiting the Kantorovich-Rubinstein duality theorem we can restrict to the existential fragment of this theory. For this fragment, more efficient decision procedures exist.

*To my parents*

## Acknowledgements

I am extremely thankful to my supervisor Professor Franck van Breugel for his guidance, support, patience and constant encouragement. I thank him for introducing me to the area of concurrency theory and probabilistic systems. I learned many new concepts under his supervision. He was friendly and approachable and gave me time for discussions whenever needed. Our weekly meetings kept me motivated and helped in steady progress of my research work. I cannot thank him enough for his contributions to my work and help with thesis writing. It has been my privilege working with him.

I am grateful to Professor Jeff Edmonds, Professor Parke Godfrey and Professor Stephen Watson for agreeing to serve on my supervisory committee and providing very useful feedback on my thesis. I greatly appreciate Professor Parke Godfrey's efforts in reading my thesis in great depth and providing detailed comments. My heartfelt thanks to Professor Jeff Edmonds for not only providing technical feedback but also for being a great friend and support.

I am thankful to Professor Christel Baier for providing useful pointers to the literature. I wish to thank graduate director Professor Richard Wildes for his help in arranging my thesis defence at a short notice. I also thank graduate assistant Mel Poteck for her help with administrative matters. I am thankful to the Computer Science and Engineering Department and NSERC for providing financial support.

I thank my fellow graduate students and friends especially Hai Feng Huang, Sergey Kulikov, Dusty Philips, Amir Ali Ghadiri, Jing Yang, Niloufar Shafiei, Olena Borzenko, Shakil Khan, Anton Belov and Nelson Moniz for the wonderful company and cheerful environment in the department.

Finally, I would like to thank my parents and family members in India for their constant love. I thank my husband Jaisingh for the long technical discussions, help with proofs and most importantly for his love and understanding.

# Table of Contents

<b>Abstract</b>	<b>iv</b>
<b>Acknowledgements</b>	<b>vi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Problem Statement and Contributions . . . . .	3
1.2 Related Work . . . . .	4
1.3 Organization of Thesis . . . . .	6
<b>2 Probabilistic Systems and Equivalences</b>	<b>8</b>
2.1 Modelling Probabilistic Systems . . . . .	8
2.1.1 Reactive and Generative Models . . . . .	11
2.2 Equivalences . . . . .	12
2.2.1 Trace Equivalence . . . . .	13
2.2.2 Bisimulation Equivalence . . . . .	13
2.2.3 Simulation Preorder . . . . .	15
2.2.4 Probabilistic Bisimulation . . . . .	16
<b>3 Behavioural Pseudometrics</b>	<b>18</b>
3.1 Metric Spaces . . . . .	18
3.2 $\epsilon$ -Bisimilarity . . . . .	19
3.3 Logical Characterization of Equivalence . . . . .	21
3.4 Coalgebraic Approach . . . . .	24
<b>4 Our Work: A Fixpoint Characterization of <math>d_1</math></b>	<b>25</b>
4.1 Lattices and Fixpoint Theorems . . . . .	25
4.2 A Fixpoint Characterization of $d_1$ . . . . .	28
4.3 Reaching a Fixpoint by Iteration . . . . .	32
<b>5 The Algorithm: Reduction to First Order Theory over Reals</b>	<b>37</b>
5.1 Duality Theorem . . . . .	37
5.2 First Order Theory over Reals . . . . .	38
5.3 The Reduction . . . . .	39
5.4 Optimizations . . . . .	41
5.4.1 Classification of States . . . . .	41
5.4.2 State Space Minimization . . . . .	44
5.4.3 Simplifications . . . . .	46
5.4.4 An Example Formula . . . . .	47

<b>6</b>	<b>Extension of our Approach to other Models</b>	<b>48</b>
6.1	Partially-defined Probabilistic Systems . . . . .	48
6.2	Labelled Probabilistic Transition Systems . . . . .	51
6.3	Metric-labelled Transition Systems . . . . .	52
<b>7</b>	<b>Implementation</b>	<b>54</b>
7.1	Decision Procedures for the First Order Theory over Reals . . . . .	54
7.2	Implementation based on Mathematica . . . . .	55
<b>8</b>	<b>Conclusion</b>	<b>57</b>
	<b>Bibliography</b>	<b>59</b>
<b>A</b>	<b>Calculation of Termination Probabilities</b>	<b>65</b>
<b>B</b>	<b>An Example Formula</b>	<b>67</b>



# 1 Introduction

Digital systems are present everywhere in the modern world. We interact with digital systems such as elevators, tele-communication systems and access-control systems everyday. These systems interact with the environment and are used to control physical systems and devices. With the rapid increase in the use of such systems, there is also a need for reasoning about their behaviour. For example, we may be interested in determining if a system satisfies its requirements or in analyzing its performance. Models of systems are often used for analysis and reasoning. This is because models such as state-transition diagrams hide irrelevant details of the system and thereby provide an abstract view amenable to analysis.

There are systems such as modern distributed systems, network protocols, randomized algorithms and neural networks which exhibit probabilistic behaviour. Such systems can be modelled using probabilistic transition systems or probabilistic automata. Also, probability is an abstraction mechanism that is sometimes used to hide inessential or unknown details. For example, when off-the-shelf components are used for constructing systems, very little information is available about their internal design. In such cases, their behaviour can be approximated probabilistically by observing results (outputs/response) in multiple runs or executions of the system.

Equivalence of systems is a fundamental question in concurrency theory. By equivalence we mean similar (or identical) behaviour of two systems. When two systems are equivalent, one can be replaced by the other. We are specifically interested in behavioural equivalence of probabilistic systems. For nondeterministic systems the notion of *bisimulation* given by Milner [Mil83] and Park [Par81] is regarded as the standard notion of equivalence. Two systems are said to be *bisimilar* if they can match each other's moves and cannot be distinguished by an observer. If one system produces a step to some next state then the other system can produce a similar step resulting in a related next state. Larsen and Skou [LS91] extended this notion to probabilistic systems and defined *probabilistic bisimulation*. The idea is to match not only the transitions but also the probabilities with which they are taken. Bisimulation and probabilistic bisimulation are formally defined in the next chapter.

Probabilistic bisimulation is regarded as the strongest form of behavioural equivalence. It is, however, a very rigid notion because two systems are either bisimilar or not. This is often restrictive in the presence of quantitative data such as probabilities because they are mostly estimates or averages and do not represent exact information. Two systems can be classified as *non-bisimilar* even if there is a slight difference in probabilities. For example consider the probabilistic transition systems shown in Figure 1.1. When  $\epsilon$  equals zero the systems are probabilistic bisimilar. However, even for very small non-zero value of  $\epsilon$  they become non-bisimilar.

A more robust notion of equivalence is therefore needed. Given three systems  $P_1, P_2$  and  $P_3$ , an analysis which reports that  $P_1$  is more similar to  $P_2$  than  $P_3$  is more useful than an analysis which reports them to be non-bisimilar to each other. Therefore, a notion of *approximate* equality instead of *exact* equivalence is more useful in systems involving quantitative data. Motivated by this reasoning, Giacalone, Jou and Smolka [GJS90] suggested weakening the notion of probabilistic

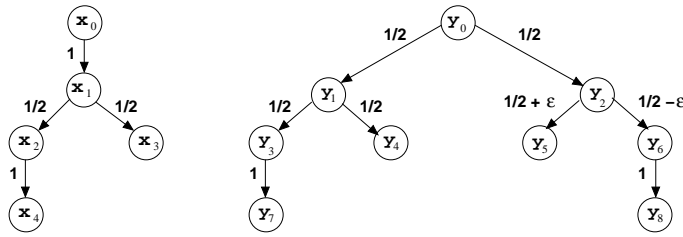


Figure 1.1: Probabilistic transition systems

bisimilarity of Larsen and Skou. They introduced the notion of *nearness* or *distance* between probabilistic systems via a *pseudometric*<sup>1</sup>. A pseudometric is a function which maps two given systems to a real number in the interval  $[0,1]$ . Therefore, in this approach, instead of giving a boolean answer to logical questions such as “Are two given systems equivalent or not?”, the answer is a real number called *distance*. The distance between two systems is a measure of similarity of their behaviours. The smaller the distance, the more alike the behaviour of the systems. In particular, distance zero implies behavioural equivalence.

Several *behavioural pseudometrics* (and closely related notions) were introduced thereafter (see, for example, [BW05, Bro98, CB02, AHM03, DCP06, DGJP04, DPHW03, GP05a, MRST06, Yin02]) for systems that contain quantitative information, like, for example, probabilities, time and costs.

There are a number of areas where analysis based on behavioural pseudometrics can be useful. For example, quantitative analysis based on pseudometrics allows us to answer questions like “If we have a large system, can we replace a component with another (cheaper/simpler) similar component?” There are also applications in the area of performance evaluation where one builds the model of the desired (expected) behaviour of a system. This model is often probabilistic (for example, modelling probability of circuit failure and probability of transmission errors). The model can then be compared using pseudometrics against the model built using experimental (actual) observations of the system. The quantification of behavioural similarity can also be used in virus detection and classification of viruses into virus-families [SKL06]. Often new viruses are developed by slight modifications to known representative viruses and therefore can be detected by computing *distances* between the control-flow graphs of their programs/executable code. Model fusion [NSC<sup>+</sup>] is another area of application where behavioural similarity is used as the basis for merging different models (views) of the same system to obtain a unified model.

Behavioural pseudometrics can also be applied on states of a single system. In this scenario, the distance between each pair of states is computed. The distance captures the behavioural similarity of the states. This can be applied to reduce the state space of systems by partitioning the state space such that states in the same partition are at most  $\epsilon$  distance apart, for some small fixed  $\epsilon$ . These reductions can be utilized in applications such as model checking, performance evaluation and stochastic planning (see, for example, [BDM01, GH02, GDG03]).

There are two types of behavioural pseudometrics: those that *discount the future* and those that do not. The *discount factor* denoted by  $\delta$ , is a real number in the interval  $(0,1]$ . In the discounted setting, when  $\delta$  is less than one, the differences in behaviour in the farther future have less impact on the distance. This means that the probabilities of transitions which are far away

<sup>1</sup>A pseudometric space differs from a metric space in that different points may have distance zero in the former and not in the latter.

from the initial state are given less weight than the probabilities of transitions that are near. In the undiscounted setting where  $\delta$  equals one, all differences in behaviour, whether in the near or far future, are given equal weight. The intuition behind discounting is that the far away future is less important than near future. (See, for example, [AHM03] for more details.) Pseudometrics for the undiscounted and discounted future usually provide different quantitative information. For systems that (in principle) run forever, all the differences whether near or far should contribute equally to the distance and, hence, for such systems the undiscounted setting is more useful.

Desharnais, Gupta, Jagadeesan and Panangaden [DGJP99, DGJP04] introduced a family of behavioural pseudometrics for probabilistic transition systems. These pseudometrics assign a distance to each pair of states of the probabilistic transition system. The distance is zero if and only if the states are *probabilistic bisimilar*. Based on their pseudometric, the distance between states  $s_1$  and  $s_2$  for different values of the discount factor  $\delta$  for the probabilistic transition system given in Figure 1.2 is 0.01 if  $\delta = 0.25$ , 0.06 if  $\delta = 0.5$ , 0.15 if  $\delta = 0.75$  and 0.32 if  $\delta = 1$ . The more

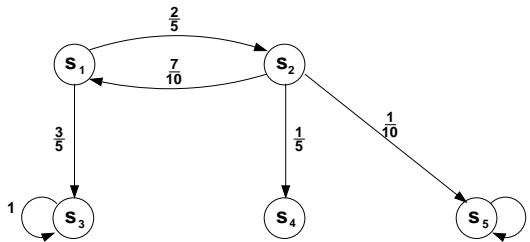


Figure 1.2: An example

we discount the future, the more  $s_1$  and  $s_2$  behave alike.

In [DGJP99], Desharnais et al. presented an *algorithm to approximate* the behavioural distances for  $\delta$  smaller than one. Van Breugel and Worrell presented a pseudometric in [BW01b, BW05] that coincides with the pseudometric of Desharnais et al. They also gave the first polynomial time approximation algorithm [BW01a, BW06] for  $\delta$  smaller than one.

## 1.1 Problem Statement and Contributions

The aim of this work is to give an algorithm for approximating distances between states of a probabilistic transition system when the discount factor  $\delta$  equals one. We use the notation  $d_1(s_1, s_2)$  to denote the distance between states  $s_1$  and  $s_2$  in the undiscounted setting. The approximation algorithms mentioned above cannot be modified in an obvious way to handle the case of  $\delta = 1$ . This is because there is a fundamental difference between pseudometrics that discount the future and one that does not. This is, for example, reflected by the fact that all pseudometrics that discount the future give rise to the same topology, whereas the pseudometric that does not discount the future gives rise to a different topology. (See, for example, [DGJP04, page 350].)

The main contribution of this thesis is proving that the problem of approximating  $d_1$  is decidable by reducing it to deciding a formula of the first order theory over reals. We present an implementation of the first algorithm to approximate distances in the undiscounted setting. Our approach makes use of several known results such as Tarski’s fixpoint theorems, Kantorovich-Rubinstein’s duality theorem and most importantly Tarski’s theorem which states that the first order theory over reals is decidable. In this thesis we state and prove our results for unlabelled finite probabilistic transition systems where a state has either no outgoing transitions or the

sum of probabilities of all outgoing transitions is one. An example of such a system is shown in Figure 1.2. Later we show that our approach can be generalized to handle several types of quantitative models including labelled probabilistic transition systems and also systems where the set of actions (labels) forms a metric space. An overview of our approach follows.

Based on the pseudometric of Desharnais et al. we first provide an alternate characterization of the pseudometric  $d_1$ . In particular, we first define a function  $\Delta$  from a complete lattice (of pseudometrics) to itself. Using Tarski’s fixpoint theorems, we prove that  $d_1$  is the greatest fixpoint of  $\Delta$ . We express this result in a formula of the first order theory over reals. Using the decision procedure for the first order theory over reals, we give an algorithm to approximate  $d_1$ . The existence of an algorithm to compute/approximate distances in the undiscounted setting has been unknown since Desharnais et al. introduced a family of pseudometrics for probabilistic transition systems. We believe ours is the first algorithm to approximate  $d_1$ .

We have implemented the algorithm as a Java program that takes as input the description of the probabilistic transition system and produces as output the simplified formula in a format that can be fed to *Mathematica*. Mathematica provides a decision procedure for checking the satisfiability of a first order formula over reals. However, due to the doubly-exponential time complexity of this decision procedure, we were able to approximate distances for very small systems.

## 1.2 Related Work

In this section, we present a brief overview of work done in the area of behavioural pseudometrics for several different types of systems.

Giacalone Jou and Smolka suggested using metrics for approximate analysis and defined the notion of  $\epsilon$ -bisimilarity [GJS90] for the class of deterministic labelled probabilistic transition systems. We will discuss  $\epsilon$ -bisimilarity in detail in Chapter 3.

As discussed earlier Desharnais et al. defined behavioural pseudometrics for probabilistic transition systems in [DGJP99, DGJP04]. In fact they presented the pseudometric for labelled Markov chains and labelled concurrent Markov chains. Probabilistic transition systems, as considered by us, are contained in the class of systems considered by them. Our work is based on their pseudometric and, hence, we will discuss it in more detail in Chapter 3.

Van Breugel and Worrell presented a behavioural pseudometric for reactive probabilistic transition systems in [BW01b, BW05]. They also showed that their pseudometric coincides with the pseudometric of Desharnais et al. Their approach is based on category theory and the theory of coalgebras. They gave an algorithm to approximate their pseudometric in [BW01a, BW06]. We will discuss more about their work in Chapter 3.

Deng, Chothia, Palamidessi and Pang [DCPP06] extended the approach of [DGJP02, DGJP04] to action-labelled quantitative transition systems. These systems subsume probabilistic automata, simple probabilistic automata [Seg95], fully probabilistic models [BH97], reactive models, generative models [GSS95] and weighted automata [Eil74, Moh03]. They defined a notion of metrics called *state-metrics* based on the Kantorovich distance. They showed that the greatest state-metric corresponds to bisimilarity and characterized it as the greatest fixpoint of an order-preserving function on state-metrics.

Gupta, Jagadeesan and Panangaden [GJP04] developed a pseudometric analogue of bisimulation for generalized semi-Markov processes (GSMP). GSMPs are real-time probabilistic systems that combine continuous time and probability. GSMPs include finite state continuous time Markov chains while also permitting general probability distributions. The formal definition of GSMP can

be found in [Whi80]. Informally, in each state of a GSMP, there are possibly several events that can be executed. Each event has its own clock, running down at its own rate, and when the first one reaches zero that event is selected for execution. Then a probabilistic transition determines the final state and any new clocks are set according to given probability distributions, defined by conditional density functions.

Mitra and Lynch [ML] proposed approximate simulations for *task-structured probabilistic I/O automata*. A probabilistic I/O automaton (PIAO) is informally defined as a countable-state automaton model that allows nondeterministic and probabilistic choices in state transitions. A task-PIAO adds a task structure on the locally controlled actions of a PIAO as a means for restricting the nondeterminism in the model. (See, for example, [ML, Definition 2.1] for a formal definition of a task-PIAO.) Mitra and Lynch define a notion of similarity of traces based on a metric on trace distributions. A trace distribution is a probability distribution over the set of traces of a task-PIAO. Informally, a task-PIAO  $A_1$  is approximately similar to a task-PIAO  $A_2$ , if every trace distribution of  $A_1$  is close to some trace distribution of  $A_2$  where *closeness* is defined by some metric on trace distributions. Formally, let  $\mu_1$  and  $\mu_2$  be probability distributions over executions of task-PIAOs  $A_1$  and  $A_2$ . An approximate simulation from  $A_1$  to  $A_2$  is a function  $d$  mapping each  $\mu_1, \mu_2$  pair to a nonnegative real number. This number is a measure of similarity of  $\mu_1$  and  $\mu_2$  in terms of producing similar trace distributions.

Van Breugel [Bre05] introduced a behavioural pseudometric for metric-labelled transition systems. Metric-labelled transition systems are labelled transition systems whose states and actions contain quantitative data such as time. These systems can model a large class of timed transition systems such as systems with uncountably many states and uncountable nondeterminism. A fix-point, logical and coinductive characterizations are presented for the pseudometric. In this thesis, we demonstrate how our approach can be used to approximate their pseudometric for finite state systems.

In [DGJP02] Desharnais, Jagadeesan, Gupta and Panangaden presented the metric analogue of weak bisimulation for systems called labelled concurrent Markov chains (LCMC). They gave a fixpoint characterization of their metric based on the definition of weak bisimulation for LCMCs given in [PLS00]. They also presented a logical characterization of their metric. An LCMC consists of a finite set of states partitioned into nondeterministic states and probabilistic states. The transitions from nondeterministic states are finitely branching and labelled with action symbols (chosen from a finite set). Transitions from probabilistic states are labelled with numbers denoting probabilities. (See, for example, [PLS00, DGJP02] for a formal definition of an LCMC.)

Girard and Pappas developed notions of approximate language inclusion, approximate simulation and approximate bisimulation for metric transition systems in [GP]. Metric transition systems as defined by them, are labeled transition systems with a set of observations associated with the states. Moreover, they are equipped with pseudometrics on the state space and observation space. These systems enable modelling of both discrete and continuous systems with either deterministic or nondeterministic dynamics. Based on the observation metric, they developed a hierarchy of pseudometrics for capturing the above mentioned equivalences. They also proposed algorithms for computing the proposed pseudometrics exactly for discrete systems and approximately for deterministic and nondeterministic continuous systems. Julius and Pappas presented the extension of approximate bisimulation by including a pseudometric in the set of labels of the metric transition systems and also introduced the notion of approximate synchronization in [JP06]. In [GP05b] Girard and Pappas extended this work by developing bisimulation functions for constrained linear systems and in [GP05a], the method is generalized to the class of metric transition systems generated by nonlinear but deterministic systems. Girard, Julius and Pappas [GJP06] presented  $\delta$ -approximate simulation relation for hybrid systems [ATHP00] and an

algorithm to approximate these simulation relations. Similar work for approximate reduction of dynamical systems and hierarchical control using approximate simulation relations is described in [TAJP06] and [GP06] respectively.

Ferns, Panangaden and Precup [FPP04, FPP05] presented metrics for measuring state similarity in Markov decision processes with finitely many states, infinitely many states and processes with continuous state spaces.

De Alfaro, Faella and Stoelinga [AFS04] defined distances to quantify trace inclusion and bisimulation for quantitative transition systems. A quantitative transition system is a transition system consisting of finite number of states, transitions and propositions and a function that assigns a real number (between zero and one) to every state for each proposition. They also present a logical characterization of the distances in terms of quantitative versions of LTL and the  $\mu$ -calculus.

In [AHM03] de Alfaro, Henzinger and Majumdar presented discounted quantitative  $\mu$ -calculus. They considered concurrent probabilistic game structures which generalize nondeterministic transition systems, Markov decision processes and deterministic two-player games. They showed that in the discounted setting, the bisimilarity distance between two states is equal to the supremum, over all  $\mu$ -calculus formulas, of the difference between the values of a formula at the two states.

In [Bro98] Broucke defined a quantitative notion (Skorohod metric) of equivalence on trajectories of hybrid systems.

### 1.3 Organization of Thesis

The remainder of this thesis is organized as follows.

- In Chapter 2 we describe various types of models for non-probabilistic and probabilistic systems. We also discuss various types of behavioural equivalences such as trace equivalence, simulation preorder and bisimulation for non-probabilistic systems and probabilistic bisimulation for probabilistic systems.
- In Chapter 3 we introduce the concept of metric spaces and pseudometrics. We then present a brief survey of the work done earlier to compute/approximate behavioural equivalences for probabilistic transition systems. In particular, we discuss Giacalone, Jou and Smolka's idea of quantitative analysis, Desharnais, Gupta, Jagadeesan and Panangaden's logical characterization of probabilistic bisimilarity and Van Breugel and Worrell's approach based on coalgebras and category theory.
- In Chapter 4 we present our main theoretical results, in particular, the characterization of  $d_1$  as the fixpoint of an order-preserving function. We also present the iterative approach for reaching the greatest fixpoint. The work described in this chapter has been done in collaboration with Van Breugel and Worrell.
- In Chapter 5 we present the reduction of the problem of approximating pseudometric  $d_1$  to deciding a formula in the first order theory over reals. We also present our algorithm and discuss several optimizations to reduce the size of the formula.
- In Chapter 6 we present an extension of our work to partial-probabilistic transition systems, labelled probabilistic transition systems and metric-labelled transition systems.
- In Chapter 7 we survey a few decision procedures for checking satisfiability of a formula of the first order theory over reals.

- Finally, we conclude in Chapter 8 by summarizing our ideas and contributions.

## 2 Probabilistic Systems and Equivalences

In this chapter we describe a few probabilistic and non-probabilistic models. One can choose the appropriate model depending on the level of abstraction required. We also present a few types of equivalences for probabilistic and non-probabilistic systems. Specifically we discuss probabilistic bisimilarity of Larsen and Skou which is related to our work.

### 2.1 Modelling Probabilistic Systems

We first define a few non-probabilistic models which can be extended to handle probabilities.

**Definition 2.1.1.** *A transition system is a tuple  $\langle S, \rightarrow \rangle$ , consisting of*

- *a finite set  $S$  of states and*
- *a transition relation  $\rightarrow \subseteq S \times S$ .*

A transition system can be represented by means of a transition diagram.

**Example 2.1.2.** *The system  $\langle S, \rightarrow \rangle$  where  $S = \{s_1, s_2, s_3, s_4\}$  and  $\rightarrow = \{(s_1, s_2), (s_1, s_3), (s_2, s_4)\}$ , is represented as shown below. The states  $s_3$  and  $s_4$  are terminal*

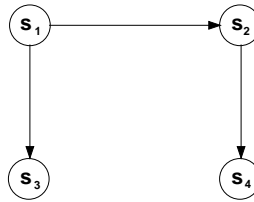


Figure 2.1: A transition system

*states with no outgoing transitions.*

Labelled transition systems evolve from transition systems and are used to model the change of state in a system as a result of executing an *action*. There are two ways to incorporate labels in a transition system namely by labelling the states (usually with some values of variables, or a set of propositions true in the state), or by labelling the transitions with actions. In this thesis we consider the latter type of labelling.

**Definition 2.1.3.** *A labelled transition system is a tuple  $\langle S, A, \rightarrow \rangle$ , consisting of*

- *a finite set  $S$  of states,*
- *a finite set  $A$  of actions and*



- a labelled transition relation  $\rightarrow \subseteq S \times A \times S$ .

We use  $s \xrightarrow{a} s'$  to denote that  $(s, a, s') \in \rightarrow$ .  $s \xrightarrow{a}$  expresses that state  $s$  has an outgoing  $a$ -transition.  $s \not\xrightarrow{a}$  expresses that this is not the case. For  $\sigma$ , a string of actions  $a_1 \dots a_n$ , we write

$$s \xrightarrow{\sigma} s' \text{ if } \exists s_1, \dots, s_{n-1}. s' \xrightarrow{a_1} s_1 \xrightarrow{a_2} \dots \xrightarrow{a_{n-1}} s_{n-1} \xrightarrow{a_n} s'.$$

**Example 2.1.4.** Consider the labelled transition system shown in Figure 2.2. In this example  $S = \{s_0, s_1, s_2, s_3, s_4, s_5\}$  and  $A = \{a, b, c\}$ .

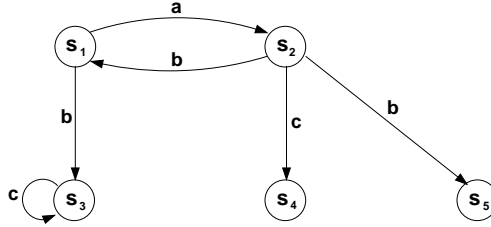


Figure 2.2: A labelled transition system

Often in the literature, an initial state is also included in the tuple. Sometimes we define the transition relation as  $\rightarrow \subseteq S \times (A \cup \{\tau\}) \times S$  where  $\tau$  denotes *unobservable* or *silent* actions.

Systems which exhibit probabilistic behaviour can be described using probabilistic models. Also probability is an abstraction mechanism that is sometimes used to hide inessential or unknown details. Below we define a few variants of probabilistic models.

**Definition 2.1.5.** A probabilistic transition system is a tuple  $\langle S, \pi \rangle$  consisting of

- a finite set  $S$  of states and
- a function  $\pi : S \times S \rightarrow [0, 1] \cap \mathbb{Q}$  satisfying  $\sum_{s' \in S} \pi(s, s') \in \{0, 1\}$ .

We write  $s \rightarrow$  if  $\sum_{s' \in S} \pi(s, s') = 1$  and  $s \not\rightarrow$  if  $\sum_{s' \in S} \pi(s, s') = 0$ .

$\pi(s, s')$  denotes the probability of making a transition from state  $s$  to state  $s'$ . Note that we restrict the probabilities to rational numbers.

A state  $s$  such that  $s \not\rightarrow$  is called a *terminal* state. The example shown in Figure 1.2 of Chapter 1 is a probabilistic transition system.

A *partial-probabilistic system* can model the situation where a state  $s$  may refuse to make a transition with some probability.

**Definition 2.1.6.** A partial-probabilistic transition system is a tuple  $\langle S, \pi \rangle$  consisting of

- a finite set  $S$  of states and
- a function  $\pi : S \times S \rightarrow [0, 1] \cap \mathbb{Q}$  satisfying  $\sum_{s' \in S} \pi(s, s') \leq 1$ .

**Example 2.1.7.** In the partial-probabilistic system shown in Figure 2.3, states  $s_1$  and  $s_3$  refuse to make a transition with probabilities  $\frac{1}{10}$  and  $\frac{1}{5}$ , respectively.

Discrete-time Markov chains are a variant of probabilistic transition systems.

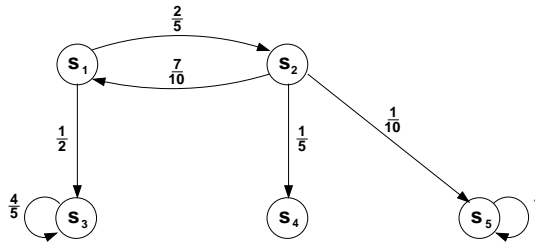


Figure 2.3: A partial-probabilistic system

**Definition 2.1.8.** A discrete-time Markov chain (DTMC) is a tuple  $\langle S, \pi \rangle$  consisting of

- a finite set  $S$  of states and
- a function  $\pi : S \times S \rightarrow [0, 1]$  satisfying  $\sum_{s' \in S} \pi(s, s') = 1$  for all states  $s$  in  $S$ .

A DTMC therefore is also a probabilistic transition system. Note that in DTMCs there are no terminal states. In the literature, a state having a self-loop with probability 1 is sometimes called a *terminating* state.

We can also incorporate labels or *actions* in these models.

**Definition 2.1.9.** A labelled probabilistic transition system is a tuple  $\langle S, A, \pi \rangle$  consisting of

- a finite set  $S$  of states,
- a finite set  $A$  of actions and
- a function  $\pi : S \times A \times S \rightarrow [0, 1] \cap \mathbb{Q}$  satisfying  $\sum_{u \in S} \pi(s, a, u) \in \{0, 1\}$  for each  $a \in A$ .

We use the short form  $\pi_a(s, u)$  instead of  $\pi(s, a, u)$ . For  $a \in A$ , we use  $s \xrightarrow{a, p} s'$  if  $\pi_a(s, s') = p$ ,  $s \xrightarrow{a}$  if  $\sum_{u \in S} \pi_a(s, u) = 1$  and  $s \not\xrightarrow{a}$  if  $\sum_{u \in S} \pi_a(s, u) = 0$ . We say that a state is terminal (written  $s \dashrightarrow$ ) if  $s \not\xrightarrow{a}$  for all  $a \in A$ .

**Example 2.1.10.** The system shown in Figure 2.4 is a labelled probabilistic transition system.

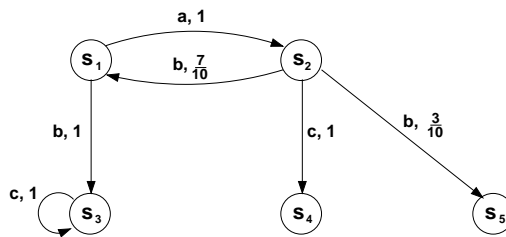


Figure 2.4: A labelled probabilistic transition system

A labelled probabilistic transition system is also called labelled Markov decision process (MDP). An MDP can model both nondeterministic and probabilistic behaviour and therefore it allows us to model the behaviour of a number of probabilistic systems running in parallel.

**Example 2.1.11 ([RKNP04]).** Figure 2.5 shows an example of an MDP and how it can be used to model the parallel composition of two DTMCs. The figure includes two three state DTMCs  $D_1$  and  $D_2$ , and the corresponding nine state MDP  $M_{12}$ .

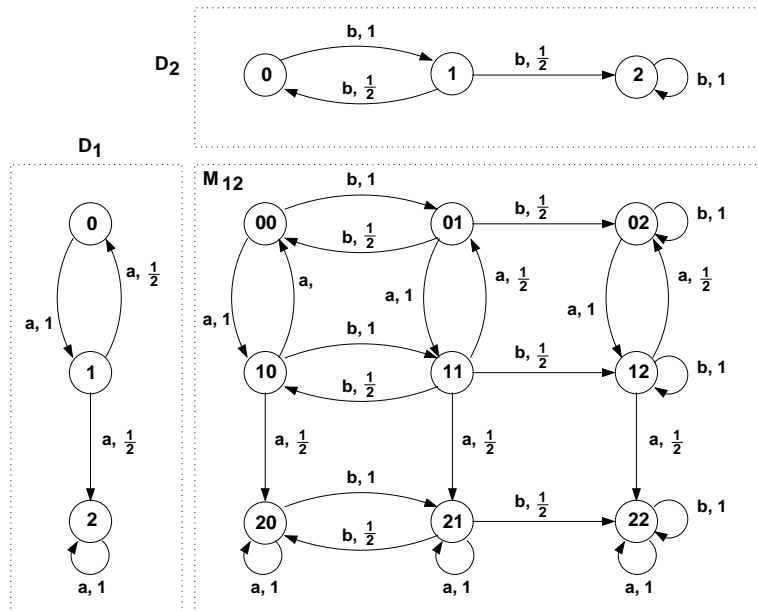


Figure 2.5: An MDP  $M_{12}$  representing the parallel composition of DTMCs  $D_1$  and  $D_2$

### 2.1.1 Reactive and Generative Models

For nonprobabilistic systems the reactive model of process behaviour is of interest. In the reactive model a system reacts to stimuli presented by its environment. Milner [Mil80] presented a mechanistic view of the reactive model for nonprobabilistic systems in terms of *button pushing experiments*. The idea is to consider a labelled transition system as a black-box with buttons (one for each action) as its interface to the outside world. The environment or user interacts with the system by attempting to depress one of several buttons. The interaction succeeds if the button is unlocked and therefore goes down; otherwise, the experiment fails. In response to a successful experiment, the system makes an internal state transition and is then ready for further experimentation.

**Example 2.1.12.** Consider the labelled transition system shown in Figure 2.2. If the system is in state  $s_1$  and the user depresses button  $a$  (or the system receives input event  $a$ ), it makes a transition to state  $s_2$ . In state  $s_1$  button  $c$  is locked (cannot be depressed). However, if the system is in state  $s_2$ , the user has buttons  $b$  and  $c$  unlocked and when the  $b$  button is depressed the system non-deterministically chooses to go to either state  $s_1$  or state  $s_5$ .

The reactive model has been adopted by Larsen and Skou [LS91] for probabilistic systems. In the reactive model, a button-pressing experiment succeeds with probability one or fails. If successful, the system makes an internal state transition according to a probability distribution associated with the depressed button and the current state of the system.

**Example 2.1.13.** Consider the probabilistic labelled transition system shown in Figure 2.4. If the system is in state  $s_2$  and the  $b$  button is depressed, it goes to state  $s_1$  with probability  $\frac{7}{10}$  or to state  $s_5$  with probability  $\frac{3}{10}$ .

Van Glabbeek et al. [GSS95] introduced the *generative* model for probabilistic systems. In the generative model, a system defines in each of its states a probability distribution over a set of enabled actions. An observer may attempt to depress more than one button at a time. Now the system decides, according to a prescribed probability distribution, which button if any will go down. In response to a successful outcome, this same probability distribution, conditioned by the system's choice of button, will govern the internal state transition made by the system.

**Example 2.1.14** ([GSS95]). *Consider the generative model of a system shown in Figure 2.6. If the observer at state  $s_1$  simultaneously attempts to depress the  $a$  and  $b$  buttons, the system will*

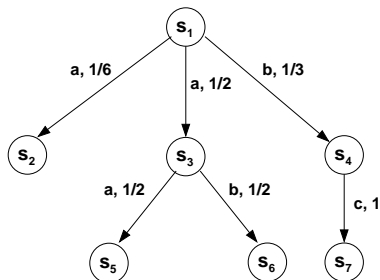


Figure 2.6: A generative model

*unlock its  $a$ -button with probability  $\frac{2}{3}$  and its  $b$ -button by probability  $\frac{1}{3}$ . In the former case, the system will branch left with probability  $\frac{1}{4}$  and right with probability  $\frac{3}{4}$ .*

The reactive model can be obtained from the generative model by abstraction (for details see [GSS95]) and the nonprobabilistic model is an abstraction of the reactive model. We consider the reactive model for probabilistic systems in the work presented in this thesis.

## 2.2 Equivalences

Given two systems (or two states of a system), we are interested in determining whether they are behaviourally equivalent or not. Several notions of behavioural equivalence exist. They can be broadly classified as linear time and branching time equivalences (see, for instance, [BBKM84, Pnu85]). In the former, equivalence is determined by possible executions of the systems, whereas in the latter the branching structure of systems is also taken into account. The standard example of a linear time equivalence is *trace equivalence* as employed in [Hoa80]; the standard example of a branching time equivalence is *observation equivalence* or *bisimulation equivalence* as defined by Milner [Mil80, Mil83, Mil89] and Park [Par81]. There are several decorated trace equivalences (see [Gla01] for an overview), that preserve part of the branching structure of systems.

In the following sections we discuss trace equivalence, bisimulation equivalence, simulation preorder and probabilistic bisimulation. These equivalence relations are defined on the set of states of a single system. For two different transition systems these notions of equivalences can be used by forming the disjoint composition of the two systems. For example, consider two labelled transition systems  $(S', A', \rightarrow')$  and  $(S'', A'', \rightarrow'')$  with initial states  $s'$  and  $s''$ . Their disjoint composition is a labelled transition system  $(S, A, \rightarrow)$  such that  $S = S' \cup S''$ ,  $A = A' \cup A''$  and  $\rightarrow = \rightarrow' \cup \rightarrow''$ , where  $\cup$  is the disjoint union operator between sets. Then, two systems are said to be equivalent if their initial states  $s'$  and  $s''$  are equivalent in the new system  $(S, A, \rightarrow)$ .

### 2.2.1 Trace Equivalence

Hoare [Hoa80] suggested the idea of distinguishing systems based on observing *traces*, which are *maximal* sequences of visible actions performed by a process. Two systems are considered *trace equivalent* if and only if both can execute the same action sequences.

**Definition 2.2.1.** Let  $\langle S, A, \rightarrow \rangle$  be a labelled transition system. For every  $s \in S$ , let

$$Tr(s) = \{\sigma \in A^* \mid s' \in S \text{ and } s \xrightarrow{\sigma} s'\}.$$

Two states  $s, s' \in S$  are said to be trace equivalent iff  $Tr(s) = Tr(s')$ . Two labelled transition systems are considered trace equivalent if their initial states are trace equivalent.

**Example 2.2.2.** Consider the two systems shown in Figure 2.7 with  $s_1$  and  $t_1$  as initial states.  $Tr(s_1) = Tr(t_1) = \{ab, ac\}$ . Hence they are trace equivalent.

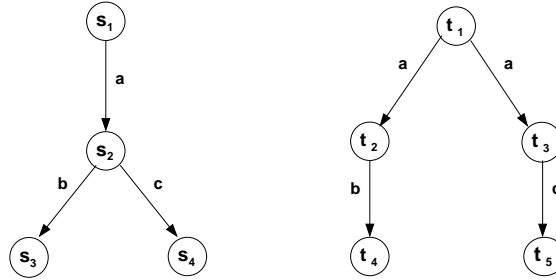


Figure 2.7: An example of trace equivalent systems

### 2.2.2 Bisimulation Equivalence

Bisimulation is a stronger form of equivalence. A bisimulation is a relation between transition systems, associating systems which behave in the same way in the sense that one system simulates the other and vice-versa. Intuitively, two systems are bisimilar if they match each other's moves. In this sense, each of the systems cannot be distinguished from the other by an observer.

**Definition 2.2.3.** Let  $\langle S, A, \rightarrow \rangle$  be a labelled transition system. Then a bisimulation relation  $\mathcal{R}$  is a binary relation  $\mathcal{R} \subseteq S \times S$  such that whenever  $s_1 \mathcal{R} s_2$  and  $a \in A$  then the following holds:

- if  $s_1 \xrightarrow{a} s'_1$ , then  $s_2 \xrightarrow{a} s'_2$  for some  $s'_2$  such that  $s'_1 \mathcal{R} s'_2$ , and
- if  $s_2 \xrightarrow{a} s'_2$ , then  $s_1 \xrightarrow{a} s'_1$  for some  $s'_1$  such that  $s'_1 \mathcal{R} s'_2$ .

Two states  $s_1$  and  $s_2$  are said to be bisimilar, written  $s_1 \sim s_2$ , if  $(s_1, s_2)$  is contained in some bisimulation  $\mathcal{R}$ . Two labelled transition systems are said to be bisimilar if their initial states are bisimilar.

The bisimilarity relation  $\sim$  is an equivalence relation. It is also the largest bisimulation relation over a given transition system.

**Example 2.2.4.** The systems shown in Figure 2.7 are not bisimilar.

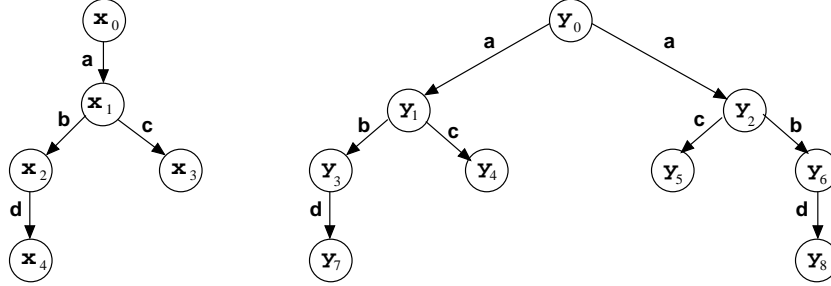


Figure 2.8: An example of bisimilar systems

**Example 2.2.5.** *The two systems shown in Figure 2.8 with initial states  $x_0$  and  $y_0$  are bisimilar.*

There are tools that calculate bisimulation equivalence classes (see, for example, [CS96]). Process algebra is another approach in which an equational theory is used to express the equivalence of processes. Equational and logical reasoning are used to obtain equivalence from the syntax of a system instead of finding the semantics [BPS01]. Another way of reasoning about bisimulation relations is by giving a logical characterization. A logical language, usually some form of modal logic, is given and states are shown to be bisimilar when they satisfy the same formulas. Showing that two states are not bisimilar can then be done by giving a distinguishing formula in the logic.

Hennesy and Milner [HM85] gave a logical characterization of bisimilarity. They defined a logic called Hennessy-Milner logic (HML) and proved that two states are bisimilar if and only if they satisfy exactly the same HML formulas. We briefly recall HML from [HM85].

**Definition 2.2.6.** *The formulas of HML are given by the following syntax:*

$$F ::= \text{true} \mid \langle a \rangle F \mid F_1 \wedge F_2 \mid \neg F$$

where  $a$  is an action.

The satisfaction relation  $s \models F$  where  $s$  is a state of a labelled transition system  $\langle S, A, \rightarrow \rangle$ , is defined as:

- $s \models \text{true}$  holds for every state,
- $s \models \langle a \rangle F$  if  $s \xrightarrow{a} s'$  and  $s' \models F$  for some  $s'$ ,
- $s \models F_1 \wedge F_2$  holds if  $s \models F_1$  and  $s \models F_2$  and
- $s \models \neg F$  holds if  $s \not\models F$ .

Bisimulation relations are useful in proving correctness of distributed algorithms and protocols. Furthermore, bisimulation can be used to reduce a system. The idea is to partition the state space of a transition system such that bisimilar states are placed in the same equivalence class. The resulting system is called the *bisimulation quotient* and the resulting state space is called the *quotient space*. The bisimulation quotient is used for abstraction and minimization purposes.

**Definition 2.2.7.** *Let  $S$  be a set and  $\mathcal{R}$  an equivalence relation on  $S$ . We write  $[s]_{\mathcal{R}}$  to denote the  $\mathcal{R}$ -equivalence class of state  $s$ , i.e.  $[s]_{\mathcal{R}} = \{s' \in S \mid (s, s') \in \mathcal{R}\}$ . The quotient space of  $S$  under  $\mathcal{R}$  is the set  $S_{\mathcal{R}} = \{[s]_{\mathcal{R}} \mid s \in S\}$  consisting of all  $\mathcal{R}$ -equivalence classes.*

Partitioning-splitter algorithms [PT87] can be used to generate the bisimulation quotient for a given finite transition system.

The bisimulation relation defined above is also sometimes referred to as *strong bisimulation*. *Weak bisimulation* is a variation where ‘internal’ and ‘external’ actions are distinguished and only the external behaviour has to be the same.

### 2.2.3 Simulation Preorder

Simulation preorder is a *uni-directed* variant of bisimulation. Intuitively, a transition system  $T_1$  is simulated by another transition system  $T_2$  if each step of  $T_1$  can be matched by a step of  $T_2$  but the converse might not hold. We now define the simulation relation on the states of a single labelled transition system.

**Definition 2.2.8.** Let  $\langle S, A, \rightarrow \rangle$  be a labelled transition system. Then a simulation relation  $\mathcal{R}$  is a binary relation  $\mathcal{R} \subseteq S \times S$  such that for every pair of elements  $s_1, s_2 \in S$ , if  $s_1 \mathcal{R} s_2$  then for all  $a \in A$ , and for all  $s'_1 \in S$ ,

- if  $s_1 \xrightarrow{a} s'_1$ , then  $s_2 \xrightarrow{a} s'_2$  such that  $s'_1 \mathcal{R} s'_2$ .

Given two states  $s_1$  and  $s_2$  in  $S$ ,  $s_2$  simulates  $s_1$ , written  $s_1 \preceq s_2$  if there is a simulation  $\mathcal{R}$  such that  $(s_1, s_2) \in \mathcal{R}$ . In such a case,  $s_1$  and  $s_2$  are said to be similar and  $\preceq$  is called the similarity relation. Two labelled transition systems are said to be similar if initial states are similar.

The similarity relation is the largest simulation relation over a given transition system. The simulation relation is transitive and reflexive but not symmetric. Hence, this equivalence is also called *simulation preorder*.

**Example 2.2.9.** Consider the systems shown in Figure 2.9 modelling two vending machines. The vending machine on the left simulates the one on the right.

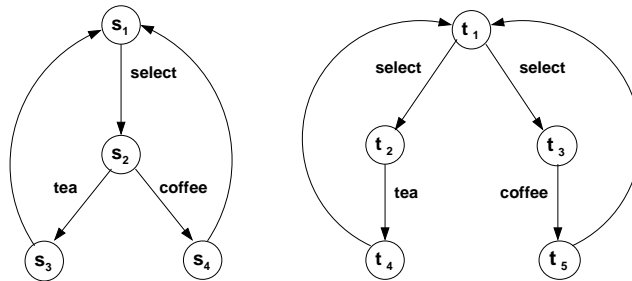


Figure 2.9: An example: vending machines

Simulation preorder induces an equivalence which is coarser than bisimulation equivalence and, hence, yields a better abstraction (i.e., a smaller quotient space).

## 2.2.4 Probabilistic Bisimulation

Larsen and Skou [LS91] extended the notion of bisimulation to probabilistic systems and defined *probabilistic bisimulation*. The idea is to match not only the transitions but also the probabilities with which they are taken.

**Definition 2.2.10.** Let  $\langle S, \pi \rangle$  be a probabilistic transition system. An equivalence relation  $\mathcal{R}$  on the set of states  $S$  is a probabilistic bisimulation if  $s_1 \mathcal{R} s_2$  implies  $\sum_{s \in E} \pi(s_1, s) = \sum_{s \in E} \pi(s_2, s)$  for all  $\mathcal{R}$ -equivalence classes  $E$ . States  $s_1$  and  $s_2$  are probabilistic bisimilar, denoted  $s_1 \sim s_2$ , if  $s_1 \mathcal{R} s_2$  for some probabilistic bisimulation  $\mathcal{R}$ .

Probabilistic bisimulation is a considerably stronger notion than Milner's bisimulation, since two states are required not only to derive the same equivalence classes, but must do so exactly with the same probability.

**Example 2.2.11.** Consider the probabilistic transition system of Figure 2.10. The equivalence

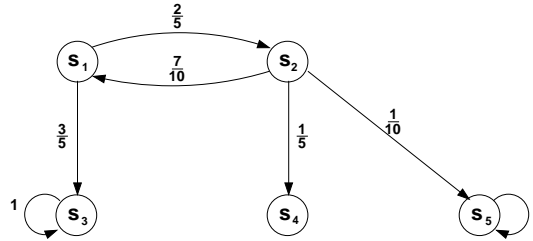


Figure 2.10: Probabilistic bisimilar states  $s_3$  and  $s_5$

relation containing  $(s_3, s_5)$  is a probabilistic bisimulation. Hence, the states  $s_3$  and  $s_5$  are probabilistic bisimilar.

**Example 2.2.12.** The systems shown in Figure 2.11 are probabilistic bisimilar.

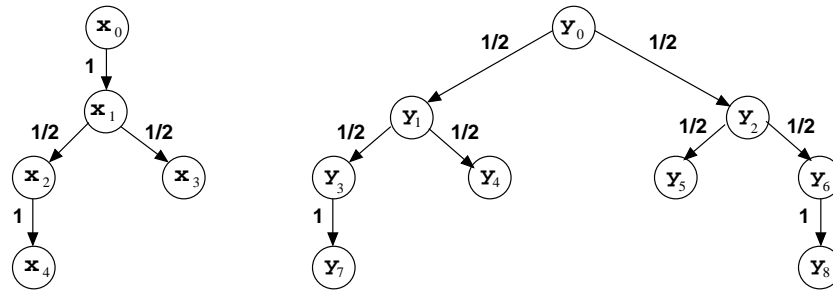


Figure 2.11: Probabilistic bisimilar systems

In [LS91], Larsen and Skou also presented an extension of HML for probabilistic systems called *Probabilistic Modal Logic* (PML).

**Definition 2.2.13.** The formulas of PML are given by the following syntax:

$$F ::= \text{true} \mid \langle a \rangle_q F \mid F_1 \wedge F_2 \mid \neg F$$



where  $a \in A$  and  $q \in [0, 1] \cap \mathbb{Q}$ .

The satisfaction relation  $s \models F$ , where  $s$  is a state of a probabilistic transition system  $\langle S, \pi \rangle$ , is defined as:

- $s \models \text{true}$  holds for all states,
- $s \models \langle a \rangle_q F$  holds if  $\sum_{s' \models F} \pi_a(s, s') \geq q$ ,
- $s \models F_1 \wedge F_2$  holds if  $s \models F_1$  and  $s \models F_2$  and
- $s \models \neg F$  holds if  $s \not\models F$ .

They proved that two states of a probabilistic transition system are probabilistic bisimilar just in case they satisfy the same PML formulas. In [DEP02], Desharnais et al. proved that negation is not needed for the characterization of probabilistic bisimulation.

Larsen and Skou defined probabilistic bisimulation for the reactive model of probabilistic systems. Van Glabbeek et al. [GSS95] extended this notion to the generative model of probabilistic systems.

We consider states of a probabilistic transition system behaviourally equivalent if they are probabilistic bisimilar [LS91]. Probabilistic bisimulation is regarded as the strongest form of behavioural equivalence. It is, however, a very rigid notion because two systems are either bisimilar or not.

**Example 2.2.14.** Consider the systems shown in Figure 2.12. Even if  $\epsilon$  is a very small non-zero number, according to the definition of probabilistic bisimilarity, the states  $x_1$  and  $y_2$  are not probabilistic bisimilar. Hence, the systems are non-bisimilar.

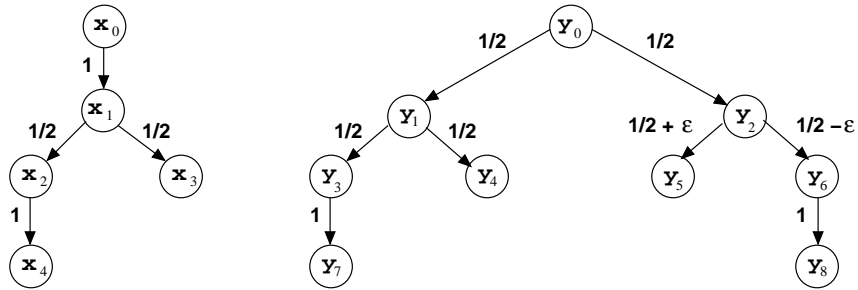


Figure 2.12: Probabilistic non-bisimilar systems

The notion of probabilistic bisimulation is therefore restrictive in the presence of quantitative data such as probabilities. Probabilities are mostly estimates or averages and do not represent exact information. Hence, a more *robust* notion is needed for defining similarity on probabilistic systems.

## 3 Behavioural Pseudometrics

We saw in Chapter 2 that probabilistic bisimulation is a very rigid notion for quantitative data. Therefore a notion of *approximate* equality instead of exact equivalence is more useful in practice for systems involving quantitative data such as for probabilistic transition systems. In this chapter we present an approach based on *pseudometrics* which allows us to quantify the behavioural similarity between two systems.

First, in Section 3.1, we briefly recall some definitions related to metric spaces which are used in the later parts of this chapter. In Section 3.2 we present the work of Giacalone, Jou and Smolka who first suggested the idea of using metrics to quantify *distance* between two systems. In Section 3.3 and Section 3.4 we present previous work in the area of behavioural pseudometrics which is closely related to our work.

### 3.1 Metric Spaces

A metric space is a set where a notion of distance between elements of the set is defined. For example, consider the set of points in the three-dimensional Euclidean space. The Euclidean metric of this space defines the distance between two points as the length of the straight line connecting them.

**Definition 3.1.1.** A metric space  $(X, d)$  is a set  $X$  with a function  $d : X \times X \rightarrow [0, \infty)$  such that for all  $x, y, z \in X$

1.  $d(x, y) = 0$  if and only if  $x = y$  (identity)
2.  $d(x, y) = d(y, x)$  (symmetry)
3.  $d(x, z) \leq d(x, y) + d(y, z)$  (triangle inequality).

The function  $d$  is called a metric or a distance on  $X$ . Instead of  $(X, d)$  we often write  $X$ .

If we relax the first condition to allow zero distance between two distinct points, then the space is known as *pseudometric* space.

**Example 3.1.2.** If  $G$  is an undirected connected graph, then the set  $V$  of vertices of  $G$  can be turned into a metric space by defining  $d(x, y)$  to be the length of a shortest path connecting the vertices  $x$  and  $y$ .

**Definition 3.1.3.** A function  $d : X \times X \rightarrow [0, \infty)$  is called 1-bounded if  $d(x, y) \leq 1$  for all  $x, y \in X$ . The metric space  $(X, d)$  is called 1-bounded if the metric  $d$  is 1-bounded.

**Example 3.1.4.** The real numbers  $\mathbb{R}$  with the usual metric  $d(x, y) = |x - y|$  is a metric space but not a 1-bounded metric space.

For any metric space  $(X, d)$  the space  $(X, d_1)$  with  $d_1(x, y) = \min\{d(x, y), 1\}$  is a 1-bounded metric space.

**Example 3.1.5.** Let  $X$  be a set. The discrete metric  $d : X \times X \rightarrow [0, 1]$  defined by

$$d(x_1, x_2) = \begin{cases} 0 & \text{if } x_1 = x_2 \\ 1 & \text{otherwise.} \end{cases}$$

is 1-bounded.

Combining Definition 3.1.1 and 3.1.3 we get the following

**Definition 3.1.6.** A 1-bounded pseudometric space is a pair  $(X, d)$  consisting of a set  $X$  and a distance function  $d : X \times X \rightarrow [0, 1]$  such that for all  $x, y, z \in X$ ,

1.  $d(x, x) = 0$ ,
2.  $d(x, y) = d(y, x)$ , and
3.  $d(x, z) \leq d(x, y) + d(y, z)$ .

A (1-bounded) pseudometric space differs from a (1-bounded) metric space in that different points may have distance zero in the former and not in the latter. Since different states of a system may behave the same, such states will have distance zero in our behavioural pseudometrics.

**Definition 3.1.7.** A metric space is compact if each sequence has a converging subsequence.

**Example 3.1.8.** Every finite space is compact. Since the space is finite, every sequence will have at least one state which is repeated infinitely many times. This subsequence is convergent.

In the characterization of a behavioural pseudometric nonexpansive functions play a key role.

**Definition 3.1.9.** Let  $(X, d)$  be a 1-bounded pseudometric space. A function  $f : X \rightarrow [0, 1]$  is nonexpansive if for all  $x_1, x_2 \in X$ ,

$$|f(x_1) - f(x_2)| \leq d(x_1, x_2).$$

The set of nonexpansive functions from  $X$  to  $[0, 1]$  is denoted by  $X \dashrightarrow [0, 1]$ .

**Example 3.1.10.** If the set  $X$  is endowed with the discrete metric, then every function from  $X$  to  $[0, 1]$  is nonexpansive.

## 3.2 $\epsilon$ -Bisimilarity

To overcome the rigidity of probabilistic bisimulation Giacalone et al. suggested using approximate analysis and defined the notion of  $\epsilon$ -bisimilarity [GJS90]. A slightly restricted definition of  $\epsilon$ -bisimilarity for the class of *deterministic labelled* probabilistic transition systems (if  $s \in S$ , then for all  $a \in A$ ,  $s$  has at most one transition on  $a$ ) is given below.

**Definition 3.2.1.** For  $\epsilon \in [0, 1)$ , a relation  $\mathcal{R}_\epsilon \subseteq S \times S$  is called an  $\epsilon$ -bisimulation if  $(s_1, s_2) \in \mathcal{R}_\epsilon$  implies for all  $a \in A$ ,

- (i) if  $s_1 \xrightarrow{a,p} s'_1$  then for some  $s'_2 \in S$ ,  $s_2 \xrightarrow{a,q} s'_2$ ,  $|p - q| \leq \epsilon$ , and  $(s'_1, s'_2) \in \mathcal{R}_\epsilon$
- (ii) if  $s_2 \xrightarrow{a,q} s'_2$  then for some  $s'_1 \in S$ ,  $s_1 \xrightarrow{a,p} s'_1$ ,  $|p - q| \leq \epsilon$ , and  $(s'_1, s'_2) \in \mathcal{R}_\epsilon$

Two states  $s_1$  and  $s_2$  are said to be  $\epsilon$ -bisimilar (written  $s_1 \stackrel{\epsilon}{\sim} s_2$ ) if there exists an  $\epsilon$ -bisimulation  $\mathcal{R}_\epsilon$  such that  $(s_1, s_2) \in \mathcal{R}_\epsilon$ . Formally,  $\stackrel{\epsilon}{\sim}$  is defined as

$$\stackrel{\epsilon}{\sim} = \bigcup \{ \mathcal{R}_\epsilon \mid \mathcal{R}_\epsilon \text{ is an } \epsilon\text{-bisimulation} \}.$$

Thus we obtain a family of binary relations  $\{ \stackrel{\epsilon}{\sim} \subseteq S \times S \mid \epsilon \in [0, 1] \}$ . These relations are not necessarily equivalences. Intuitively,  $s_1 \stackrel{\epsilon}{\sim} s_2$  if  $s_1$  and  $s_2$  can simulate each other with a bound  $\epsilon$  of deviation in probability.

**Example 3.2.2 ([GJS90]).** Consider the three labelled probabilistic transition systems shown in Figure 3.1. We have  $p \stackrel{0.1}{\sim} q$ , since  $(p, q)$  belongs to the 0.1-bisimulation

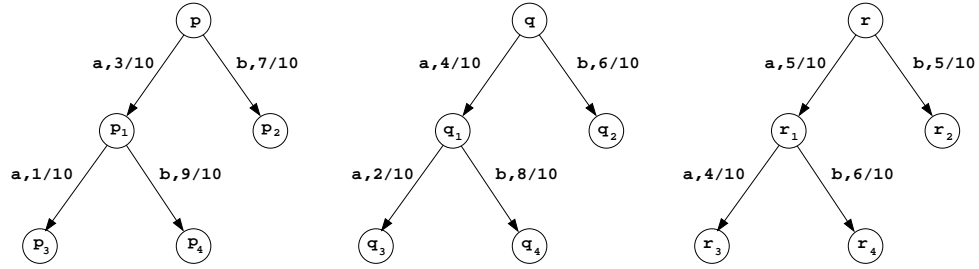


Figure 3.1: An example

$$R_{0.1} = \{ (p, q), (p_1, q_1), (p_2, q_2), (p_3, q_3), (p_4, q_4) \}.$$

Also, we can see that  $q \stackrel{0.2}{\sim} r$  and  $p \stackrel{0.3}{\sim} r$ .

The probabilistic bisimulation of Larsen and Skou coincides with  $\stackrel{0}{\sim}$ . The family of  $\stackrel{\epsilon}{\sim}$  relations is used to define a 1-bounded pseudometric on the set of states of a deterministic labelled probabilistic transition system. Intuitively, states  $s_1$  and  $s_2$  are distance  $\epsilon$  apart if they are related by  $\stackrel{\epsilon}{\sim}$  and by no other  $\stackrel{\epsilon'}{\sim}$  such that  $\epsilon' < \epsilon$ .

**Definition 3.2.3.** The distance function  $d : S \times S \rightarrow [0, 1]$  is given by

$$d(s_1, s_2) = \begin{cases} \inf \{ \epsilon \in [0, 1] \mid s_1 \stackrel{\epsilon}{\sim} s_2 \} & \text{if } s_1 \stackrel{\epsilon}{\sim} s_2 \text{ for some } \epsilon \in [0, 1] \\ 1 & \text{otherwise} \end{cases}$$

From the definition,  $d$  maps each pair of states to a non-negative real number. Also  $d(s_1, s_2) = 0$  iff  $s_1$  is bisimilar to  $s_2$ . Since  $\stackrel{\epsilon}{\sim}$  is symmetric, we have for all  $s_1, s_2 \in S$ ,  $d(s_1, s_2) = d(s_2, s_1)$ , i.e.,  $d$  is symmetric. Also  $d$  satisfies triangular inequality. Therefore,  $(S, d)$  forms a pseudometric space.

Note that  $\epsilon$ -bisimilarity has been defined only for deterministic labelled probabilistic transition systems. For the nondeterministic case, it can be shown that the distance function  $d$  does not satisfy the triangle inequality and therefore  $(S, d)$  is not a pseudometric space.

Also, note that the class of deterministic labelled probabilistic transition systems presents the generative model. In the reactive model, for all  $s \in S$  and for all  $a \in A$ ,  $\sum_{u \in S} \pi_a(s, u) \in \{0, 1\}$ . Since nondeterminism is not allowed, all transitions have to be taken with probability one. Hence  $\epsilon$ -bisimilarity is not applicable for the reactive model of probabilistic systems.

### 3.3 Logical Characterization of Equivalence

Desharnais, Gupta, Jagadeesan and Panangaden [DGJP04] introduced a family of behavioural pseudometrics for probabilistic transitions systems. Below, we will briefly review the key ingredients of their definition.

Larsen and Skou [LS91] introduced a logic that captures probabilistic bisimilarity. That is, states are probabilistic bisimilar if and only if they satisfy the same formulas. To define their behavioural pseudometrics, Desharnais et al. introduced (a variation on) the following logic.

**Definition 3.3.1.** *The logic  $\mathcal{L}$  is defined by*

$$\varphi ::= \text{true} \mid \diamond\varphi \mid \varphi \wedge \psi \mid \neg\varphi \mid \varphi \ominus q$$

where  $q$  is a rational in  $[0, 1]$ .

The key difference between the above logic and the one introduced by Desharnais et al. is that we use  $\diamond\varphi$  whereas they use  $\langle a \rangle\varphi$ . They consider labelled transitions whereas we restrict our attention to unlabelled transitions. The main difference between the above logic and the one of Larsen and Skou is that we have  $\diamond\varphi$  and  $\varphi \ominus q$  whereas they combine the operators  $\diamond$  and  $\ominus q$  into one. Since they consider labelled transitions, they use the notation  $\langle a \rangle_q$  for this combined operator.

Desharnais et al. provided a family of real-valued interpretations of the logic. That is, given a probabilistic transition system and a discount factor  $\delta$ , the interpretation gives a quantitative measure of the validity of a formula  $\varphi$  of the logic in a state  $s$  of the system. The interpretation  $\llbracket \varphi \rrbracket_\delta(s)$  is a real number in the interval  $[0, 1]$ . It measures the validity of the formula  $\varphi$  in the state  $s$ . This real number can roughly be thought of as the probability that  $\varphi$  is true in  $s$ .

**Definition 3.3.2.** *Given a probabilistic transition system  $\langle S, \pi \rangle$  and a discount factor  $\delta \in (0, 1]$ , for each  $\varphi \in \mathcal{L}$ , the function  $\llbracket \varphi \rrbracket_\delta : S \rightarrow [0, 1]$  is defined by*

$$\begin{aligned} \llbracket \text{true} \rrbracket_\delta(s) &= 1 \\ \llbracket \diamond\varphi \rrbracket_\delta(s) &= \delta \sum_{s' \in S} \pi(s, s') \llbracket \varphi \rrbracket_\delta(s') \\ \llbracket \varphi \wedge \psi \rrbracket_\delta(s) &= \min\{\llbracket \varphi \rrbracket_\delta(s), \llbracket \psi \rrbracket_\delta(s)\} \\ \llbracket \neg\varphi \rrbracket_\delta(s) &= 1 - \llbracket \varphi \rrbracket_\delta(s) \\ \llbracket \varphi \ominus q \rrbracket_\delta(s) &= \max\{\llbracket \varphi \rrbracket_\delta(s) - q, 0\} \end{aligned}$$

**Example 3.3.3.** *Consider the probabilistic transition system shown in Figure 3.2. For this sys-*

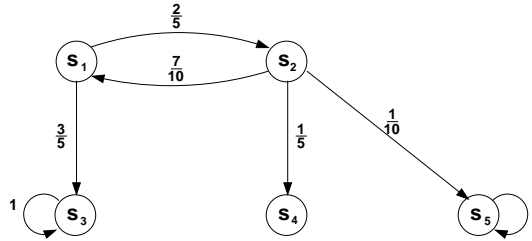


Figure 3.2: An example

tem,  $\llbracket \diamond\text{true} \rrbracket_\delta(s_3) = \delta$  and  $\llbracket \diamond\text{true} \rrbracket_\delta(s_4) = 0$ .

Given a discount factor  $\delta \in (0, 1]$ , the behavioural pseudometric  $d_\delta$  assigns a distance, a real number in the interval  $[0, 1]$ , to every pair of states of a probabilistic transition system. The distance is defined in terms of the logical formulas and their interpretation. Roughly speaking, the distance is captured by the logical formula that distinguishes the states the most.

**Definition 3.3.4.** *Given a probabilistic transition system  $\langle S, \pi \rangle$  and a discount factor  $\delta \in (0, 1]$ , the distance function  $d_\delta : S \times S \rightarrow [0, 1]$  is defined by*

$$d_\delta(s_1, s_2) = \sup_{\varphi \in \mathcal{L}} \{ \llbracket \varphi \rrbracket_\delta(s_1) - \llbracket \varphi \rrbracket_\delta(s_2) \}.$$

**Example 3.3.5.** *Consider the probabilistic transition system shown in Figure 3.2. The states  $s_3$  and  $s_4$  are  $\delta$  apart. This distance is witnessed by the formula  $\Diamond \text{true}$ .*

*The distances are collected in the following table. Since a distance function is symmetric and the distance from a state to itself is zero, we do not give all the entries.*

	$s_1$	$s_2$	$s_3$	$s_4$
$s_2$	$\frac{25\delta^2 - 2\delta^4}{125 - 25\delta - 35\delta^2 + 7\delta^3}$			
$s_3$	$\frac{2\delta^3}{25 - 7\delta^2}$	$\frac{5\delta^2}{25 - 7\delta^2}$		
$s_4$	$\delta$	$\delta$	$\delta$	
$s_5$	$\frac{2\delta^3}{25 - 7\delta^2}$	$\frac{5\delta^2}{25 - 7\delta^2}$	$0$	$\delta$

**Proposition 3.3.6.**  *$d_\delta$  is a 1-bounded pseudometric space.*

*Proof.* First, observe that, from Definition 3.3.2

$$\llbracket \varphi \rrbracket_\delta(s_1) - \llbracket \varphi \rrbracket_\delta(s_2) = \llbracket \neg\varphi \rrbracket_\delta(s_2) - \llbracket \neg\varphi \rrbracket_\delta(s_1).$$

As a consequence, we can replace  $\llbracket \varphi \rrbracket_\delta(s_1) - \llbracket \varphi \rrbracket_\delta(s_2)$  in the definition of  $d_\delta$  with  $|\llbracket \varphi \rrbracket_\delta(s_1) - \llbracket \varphi \rrbracket_\delta(s_2)|$ . Checking now that  $d_\delta$  satisfies the three conditions of Definition 3.1.6 is straightforward.  $\square$

A similar result is presented in [DGJP04, Theorem 5.2].

Each behavioural pseudometric  $d_\delta$  is a quantitative analogue of probabilistic bisimilarity. This behavioural equivalence is exactly captured by those states that have distance zero.

**Proposition 3.3.7.** *Given a probabilistic transition system  $\langle S, \pi \rangle$  and a discount factor  $\delta \in (0, 1]$ ,*

$$d_\delta(s_1, s_2) = 0 \text{ if and only if } s_1 \sim s_2$$

*for all  $s_1, s_2 \in S$ .*

*Proof.* We split the proof in two parts.

- Assume that  $s_1 \sim s_2$ . It suffices to show that  $\llbracket \varphi \rrbracket_\delta(s_1) = \llbracket \varphi \rrbracket_\delta(s_2)$  for all  $\varphi$ . We can prove this by structural induction on  $\varphi$ . We focus here on the only nontrivial case:  $\Diamond\varphi$ . Assume that  $\llbracket \varphi \rrbracket_\delta(s_1) = \llbracket \varphi \rrbracket_\delta(s_2)$ . We have to show that  $\llbracket \Diamond\varphi \rrbracket_\delta(s_1) = \llbracket \Diamond\varphi \rrbracket_\delta(s_2)$ . Let  $\{E_i \mid i \in I\}$  be the  $\sim$ -equivalence classes. Assume that  $e_i$  is an element of  $E_i$ . By induction, the function  $\llbracket \varphi \rrbracket_\delta$  restricted to  $E_i$  is constant. Hence,

$$\llbracket \Diamond\varphi \rrbracket_\delta(s_1) = \delta \sum_{s \in S} \pi(s_1, s) \llbracket \varphi \rrbracket_\delta(s)$$

$$\begin{aligned}
&= \delta \sum_{i \in I} \sum_{s \in E_i} \pi(s_1, s) \llbracket \varphi \rrbracket_\delta(s) \\
&= \delta \sum_{i \in I} \llbracket \varphi \rrbracket_\delta(e_i) \sum_{s \in E_i} \pi(s_1, s) \\
&= \delta \sum_{i \in I} \llbracket \varphi \rrbracket_\delta(e_i) \sum_{s \in E_i} \pi(s_2, s) \quad [s_1 \sim s_2] \\
&= \llbracket \Diamond \varphi \rrbracket_\delta(s_2).
\end{aligned}$$

- We show that the relation

$$\mathcal{R} = \{ (s_1, s_2) \mid d_\delta(s_1, s_2) = 0 \}$$

is a probabilistic bisimulation. Obviously,  $\mathcal{R}$  is an equivalence relation. Assume that  $s_1 \mathcal{R} s_2$ . That is,  $d_\delta(s_1, s_2) = 0$ . Let  $E$  be an  $\mathcal{R}$ -equivalence class. We denote the equivalence class that contains the state  $s$  by  $[s]_{d_\delta}$ . Therefore, we may assume that  $E$  is of the form  $[s]_{d_\delta}$ . All states in  $[s]_{d_\delta}$  assign the same value to each formula. For each state  $s' \notin [s]_{d_\delta}$  there exists a formula  $\varphi_{s'}$  such that  $\llbracket \varphi_{s'} \rrbracket_\delta(s) \neq \llbracket \varphi_{s'} \rrbracket_\delta(s')$ . Since  $\llbracket \neg \varphi \rrbracket_\delta(s) = 1 - \llbracket \varphi \rrbracket_\delta(s)$ , we may assume that  $\llbracket \varphi_{s'} \rrbracket_\delta(s) > \llbracket \varphi_{s'} \rrbracket_\delta(s')$ . Hence, there exists a rational  $q_{s'}$  in  $[0, 1]$  such that  $\llbracket \varphi_{s'} \ominus q_{s'} \rrbracket_\delta(s') = 0$  and  $\llbracket \varphi_{s'} \ominus q_{s'} \rrbracket_\delta(s) > 0$ . Now consider the formula

$$\varphi = \bigwedge_{s' \notin [s]_{d_\delta}} \varphi_{s'} \ominus q_{s'}.$$

Then  $\llbracket \varphi \rrbracket_\delta(s'') > 0$  iff  $s'' \in [s]_{d_\delta}$ . As a consequence,

$$\begin{aligned}
&\delta \llbracket \varphi \rrbracket_\delta(s) \sum_{s' \in [s]_{d_\delta}} \pi(s_1, s') \\
&= \delta \sum_{s' \in [s]_{d_\delta}} \pi(s_1, s') \llbracket \varphi \rrbracket_\delta(s') \\
&= \delta \sum_{s'' \in S} \pi(s_1, s'') \llbracket \varphi \rrbracket_\delta(s'') \quad [\llbracket \varphi \rrbracket_\delta(s'') = 0 \text{ for all } s'' \notin [s]_{d_\delta}] \\
&= \llbracket \Diamond \varphi \rrbracket_\delta(s_1) \\
&= \llbracket \Diamond \varphi \rrbracket_\delta(s_2) \quad [d_\delta(s_1, s_2) = 0] \\
&= \delta \llbracket \varphi \rrbracket_\delta(s) \sum_{s' \in [s]_{d_\delta}} \pi(s_2, s').
\end{aligned}$$

Therefore,  $\sum_{s' \in [s]_{d_\delta}} \pi(s_1, s') = \sum_{s' \in [s]_{d_\delta}} \pi(s_2, s')$  and, hence,  $\mathcal{R}$  is a probabilistic bisimulation. □

This result has also been proved in [DGJP04, Theorem 4.10].

In [DGJP99], Desharnais et al. presented a decision procedure for the behavioural pseudometric  $d_\delta$  when  $\delta$  is smaller than one. Let us briefly sketch their algorithm. They define the depth of a

logical formula as follows.

$$\begin{aligned}
\text{depth}(\text{true}) &= 0 \\
\text{depth}(\diamond\varphi) &= \text{depth}(\varphi) + 1 \\
\text{depth}(\varphi \wedge \psi) &= \max\{\text{depth}(\varphi), \text{depth}(\psi)\} \\
\text{depth}(\neg\varphi) &= \text{depth}(\varphi) \\
\text{depth}(\varphi \ominus q) &= \text{depth}(\varphi)
\end{aligned}$$

They show that

$$\llbracket\varphi\rrbracket_\delta(s_1) - \llbracket\varphi\rrbracket_\delta(s_2) \leq \delta^{\text{depth}(\varphi)}$$

for each  $\varphi \in \mathcal{L}$ . Let  $n$  be a natural number. Clearly, there exist infinitely many logical formulas  $\varphi$  with  $\text{depth}(\varphi) \leq n$ . Desharnais et al. show how to construct a finite subset  $\mathcal{F}_n$  of the logical formulas of at most depth  $n$  such that

$$d_\delta(s_1, s_2) - \sup_{\varphi \in \mathcal{F}_n} \{\llbracket\varphi\rrbracket_\delta(s_1) - \llbracket\varphi\rrbracket_\delta(s_2)\} \leq \delta^n.$$

In this way,  $d_\delta(s_1, s_2)$  can be approximated up to arbitrary accuracy provided that  $\delta$  is smaller than one.

### 3.4 Coalgebraic Approach

Van Breugel and Worrell presented a behavioural pseudometric for reactive probabilistic transition systems in [BW01b, BW05]. They also showed that their pseudometric coincides with the pseudometric of Desharnais et al. Their approach is based on category theory and the theory of coalgebras. For an overview of the theory of coalgebras see [Rut00]. We present a brief overview of their approach in this section.

They first defined an endofunctor  $P$  on the category of 1-bounded pseudometric spaces and nonexpansive maps. The definition of  $P$  is based on a metric on Borel probability measures. The metric is known as Kantorovich metric [Kan42]. The details of the functor  $P$  can be found in [BW05, Page 126]. They showed that all discrete probabilistic transition systems can be represented as  $P$ -coalgebras [BW05, Proposition 24]). A  $P$ -coalgebra consists of a pseudometric space  $S$ , called the *carrier*, together with a nonexpansive function  $t : S \rightarrow P(S)$ . The space  $S$  corresponds to the set of states of the probabilistic transition system and the nonexpansive function  $t$  characterizes the transitions of the system.

Using Rutten and Turi's (ultra)metric terminal coalgebra theorem [TR98], they showed that there exists a *terminal*  $P$ -coalgebra. A terminal coalgebra is a canonical representation of a system and there is a unique map  $\phi$  from the carrier of an arbitrary coalgebra to the carrier of the terminal coalgebra. This map preserves and reflects transitions. The terminal  $P$ -coalgebra carries a metric and states are mapped by  $\phi$  to the same element in the terminal  $P$ -coalgebra if they are probabilistically bisimilar, i.e., bisimilar states have distance zero in the terminal coalgebra. They proved that this map  $\phi$  can be approximated by a sequence of functions  $\phi_0, \phi_1, \dots, \phi_n$ . These approximations  $\phi_n$  therefore induce pseudometrics  $d_{\phi_n}$  on the carrier of the  $P$ -coalgebra, i.e., the states of a probabilistic transition system.

They showed that to calculate distances up to an accuracy  $\epsilon$ , one has to calculate  $\phi_1, \phi_2, \dots, \phi_{\lceil \log_\delta(\epsilon/2) \rceil}$  distances where  $\delta$  is the discount factor. They showed that  $\phi_{n+1}$  can be computed from  $\phi_n$  by solving a linear programming problem. Therefore the time-complexity of their algorithm to approximate distances is polynomial. However the above discussion holds only for the case when the discount factor  $\delta$  is less than one.



## 4 Our Work: A Fixpoint Characterization of $d_1$

As seen in Chapter 3, there exist algorithms for approximating  $d_\delta$  when  $\delta$  is less than one. We can view a probabilistic transition system as an unfolded (possibly infinite level) tree with loops removed. In the discounted setting, the transitions that are far away in this tree are given less weight than those that are near. Hence, for approximating distances up to some accuracy  $\alpha$ , the transitions after a certain level in the tree, become non-significant. Hence, even in the presence of loops in the original system, we get a finite level tree. This is not the case when  $\delta$  equals one, hence approximating distances in the discounted setting is easier compared to the undiscounted case. The aim of this work is to give an algorithm for approximating distances between states of a probabilistic transition system when the discount factor  $\delta$  equals one. We use the notation  $d_1(s_1, s_2)$  to denote the distance between states  $s_1$  and  $s_2$  in the undiscounted setting.

In this chapter, we present the first step of our approach to approximate  $d_1$ . First we recall some definitions and results in Section 4.1 which are useful in understanding the later part of this chapter. In Section 4.2 we provide an alternate characterization of  $d_1$  as a fixpoint of a function from a complete lattice to itself. This characterization is based on the definition of the pseudometric given by Desharnais et al.[DGJP04]. Finally in Section 4.3 we present an iterative approach to obtain the fixpoint.

### 4.1 Lattices and Fixpoint Theorems

In this section we first define partially ordered sets, order-preserving functions and lattices and then state two important theorems for complete lattices.

**Definition 4.1.1.** *Let  $P$  be a set. A partial order on  $P$  is a binary relation  $\leq$  on  $P$  such that, for all  $x, y, z \in P$ ,*

- $x \leq x$ ,
- $x \leq y$  and  $y \leq x$  imply  $x = y$ ,
- $x \leq y$  and  $y \leq z$  imply  $x \leq z$ .

*A set  $P$  equipped with a partial order relation  $\leq$  is said to be a partially ordered set or poset.*

**Example 4.1.2.** *The set  $N$  of natural numbers with its natural order is a poset.*

**Example 4.1.3.** *Let  $X$  be a set. The power set  $P(X)$  consisting of all subsets of  $X$  is a poset. It is ordered by set inclusion, that is, for  $A, B \in P(X)$ , we define  $A \leq B$  if and only if  $A \subseteq B$ .*

Let  $P$  be a finite poset. We can represent  $P$  by a Hasse diagram. (See, for example, [DP90, Section 1.9].)



Figure 4.1: Hasse diagrams for P

**Example 4.1.4.** Let  $P = \{a, b, c, d\}$  in which  $a \leq c, a \leq d, b \leq c$  and  $b \leq d$ . Figure 4.1 shows two alternative diagrams for P.

**Definition 4.1.5.** Let P be a poset. The greatest element of P, if it exists, is called the top element of P and written as  $\top$ . Similarly, the least element of P, if such exists, is called the bottom element and denoted  $\perp$ .

**Example 4.1.6.** In the power set  $P(X)$ , we have  $\top = X$  and  $\perp = \emptyset$ .

**Example 4.1.7.** The poset P of Example 4.1.4 does not have any top or bottom element.

**Example 4.1.8.** In  $\mathbb{N}$ , the set of natural numbers, 0 is the bottom element but there is no top element.

**Definition 4.1.9.** Let P be a poset and let  $S \subseteq P$ . An element  $x \in P$  is an upper bound of S if  $s \leq x$  for all  $s \in S$ . A lower bound is defined dually.

We denote the set of all upper bounds of S by  $S^u$  and the set of all lower bounds by  $S^l$ . An element  $x$  of P is called the least upper bound of S if  $x \in S^u$ , and  $x \leq y$  for all  $y \in S^u$ . Symmetrically, we can define the greatest lower bound of S.

Since least elements and greatest elements are unique, least upper bounds and greatest lower bounds are unique when they exist. In a poset P if  $x, y \in P$ , we denote the least upper bound of  $\{x, y\}$  as  $x \sqcup y$  (read as ‘x join y’) when it exists and the greatest lower bound as  $x \sqcap y$  (read as ‘x meet y’) when it exists. Similarly if  $S \subseteq P$ , we write  $\bigsqcup S$  to denote the least upper bound of S and  $\bigsqcap S$  to denote the greatest lower bound of S, when these exist.

**Example 4.1.10.** Consider poset P of Example 4.1.4. We observe that  $\{a, b\}^u = \{c, d\}$  and thus  $a \sqcup b$  does not exist as  $\{a, b\}^u$  has no least element.  $\{a, b\}^l = \emptyset$  and hence  $a \sqcap b$  also does not exist. Similarly  $\{c, d\}$  does not have a least upper bound or a greatest lower bound.

**Example 4.1.11.** Consider the poset denoted by the Hasse diagram shown in Figure 4.2. In this

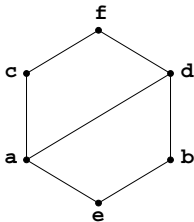


Figure 4.2: An example

case  $\{a, b\}^u = \{d, f\}$  and d is the least upper bound of  $\{a, b\}$  as it is the least element.  $\{a, b\}^l = \{e\}$  hence e is the greatest lower bound.

**Definition 4.1.12.** Let  $P$  be a non-empty poset.

- If  $x \sqcup y$  and  $x \sqcap y$  exist for all  $x, y \in P$ , then  $P$  is called a lattice.
- If  $\sqcup S$  and  $\sqcap S$  exist for all  $S \subseteq P$ , then  $P$  is called a complete lattice.

A complete lattice has a top and bottom element. The following lemma is often useful in proving that a poset is a complete lattice.

**Lemma 4.1.13.** Let  $P$  be a poset such that  $\sqcap S$  exists in  $P$  for every non-empty subset  $S$  of  $P$ . Then  $\sqcup S$  exists in  $P$  for every subset  $S$  of  $P$  which has an upper bound in  $P$ ; indeed  $\sqcup S = \sqcap S^u$ .

(See, for example, [DP90, Lemma 2.15] for a proof of this lemma.)

**Definition 4.1.14.** Let  $P$  and  $Q$  be posets. A function  $f : P \rightarrow Q$  is said to be order-preserving (or monotone) if  $x \leq y$  in  $P$  implies  $f(x) \leq f(y)$  in  $Q$ .

**Example 4.1.15.** The function  $f$  shown in Figure 4.3 is not order-preserving whereas the function  $g$  shown in Figure 4.4 is order-preserving.

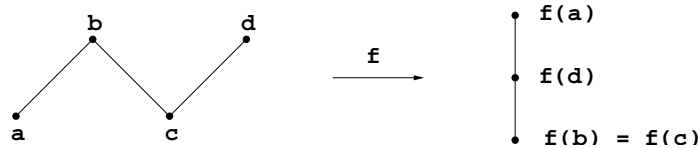


Figure 4.3: A non order-preserving function

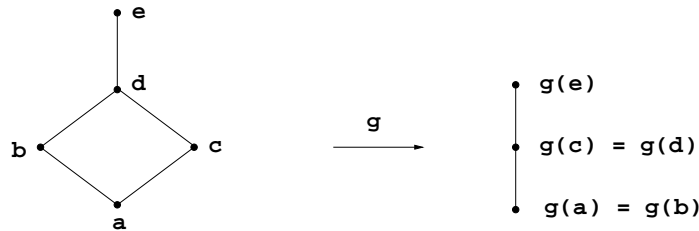


Figure 4.4: An order-preserving function

**Definition 4.1.16.** Let  $P$  be a poset and  $f : P \rightarrow P$  be a function, a point  $x \in P$  such that  $f(x) = x$  is called a fixpoint of  $f$ .

A point  $x \in P$  such that  $x \leq f(x)$  is called a pre-fixpoint of  $f$ . We use  $\text{fix}(f)$  to denote the set of all fixpoints of  $f$ .

Thus a fixpoint of  $f$  is also its pre-fixpoint. The following is an important theorem given by Tarski which holds for complete lattices.

**Theorem 4.1.17 (Lattice-Theoretical Fixpoint Theorem [Tar55]).** Let  $(L, \leq)$  be any complete lattice. Suppose  $f : L \rightarrow L$  is order-preserving. Let  $P$  be the set of all fixpoints of  $f$ . Then the set  $P$  is not empty and the system  $(P, \leq)$  is a complete lattice.

This theorem guarantees the existence of at least one fixpoint of  $f$  and also guarantees that  $f$  has a greatest fixpoint and a least fixpoint. The following theorem provides an approach to systematic search for fixpoints.

**Theorem 4.1.18 (The Knaster-Tarski Fixpoint Theorem [Tar55]).** *Let  $L$  be a complete lattice and  $f : L \rightarrow L$  an order-preserving function. Then*

$$\bigsqcup \{x \in L \mid x \leq f(x)\} \in \text{fix}(f).$$

The theorem states that the least upper bound of the set of *all* pre-fixpoints of  $f$  is a fixpoint of  $f$ . It is easily seen that the formula in Theorem 4.1.18 finds the greatest fixpoint of  $f$  in the set  $\{x \mid x \leq f(x)\}$  of *pre-fixpoints*.

A fixpoint of an order-preserving function on a complete lattice can be obtained by iteration (see, for example, [DP90, Exercise 4.13]).

**Definition 4.1.19.** *Let  $L$  be a complete lattice and  $F : L \rightarrow L$  an order-preserving function. For each ordinal  $\alpha$ , the element  $x^\alpha$  of  $L$  is defined by*

$$\begin{aligned} x^0 &= \top \\ x^{\alpha+1} &= F(x^\alpha) \\ x^\beta &= \prod_{\alpha \in \beta} x^\alpha \quad \text{if } \beta \text{ is a limit ordinal} \end{aligned}$$

**Proposition 4.1.20.** *If  $F(x) = x$  then for any ordinal  $\alpha$ ,  $x \leq x^\alpha$ .*

*Proof.* Let  $F(x) = x$ .

We prove this by induction on ordinal  $\alpha$ .

- $x \leq x^0$  is vacuously true.
- Assume  $x \leq x^\gamma$  for all ordinals  $\gamma$  in  $\alpha$ . We have to prove 2 cases.
  - Case 1: When  $\alpha$  is a successor ordinal, (i.e.,  $\alpha = \beta + 1$  for some ordinal  $\beta$ ). This means  $\beta \in \alpha$ . From the induction hypothesis, we have  $x \leq x^\beta$ . Since  $F$  is an order-preserving function,  $F(x) \leq F(x^\beta)$  which means  $x \leq x^{\beta+1}$ , i.e.,  $x \leq x^\alpha$ .

Case 2: When  $\alpha$  is a limit ordinal. From Definition 4.1.19,

$$x^\alpha = \prod_{\gamma \in \alpha} x^\gamma.$$

Using induction hypothesis, we get  $x \leq x^\gamma$  for all  $\gamma \in \alpha$ . Therefore,  $x \leq x^\alpha$ .

□

## 4.2 A Fixpoint Characterization of $d_1$

In this section we present an alternative characterization of the pseudometric  $d_1$ . In particular, we characterize  $d_1$  as the greatest (pre-)fixpoint of a function from a complete lattice to itself. This characterization can be viewed as a quantitative analogue of the greatest fixpoint characterization of bisimilarity [Par81].

The idea is as follows. Let  $P$  be the set of all 1-bounded pseudometrics. First, we define a partial order  $\sqsubseteq$  on the set  $P$ . We prove that  $(P, \sqsubseteq)$  forms a complete lattice. Then we define a function  $\Delta : P \rightarrow P$  and prove that  $\Delta$  is order-preserving. Using Tarski's fixpoint theorem, we conclude that the function  $\Delta$  has a greatest fixpoint. Finally, we prove that the greatest fixpoint of  $\Delta$  is in fact  $d_1$ , the pseudometric of our interest.

For the rest of this chapter, we fix a probabilistic transition system  $\langle S, \pi \rangle$ . We endow the set of 1-bounded pseudometrics on  $S$  with the following order.

**Definition 4.2.1.** *The relation  $\sqsubseteq$  on 1-bounded pseudometrics on  $S$  is defined by*

$$d^1 \sqsubseteq d^2 \text{ if } d^1(s_1, s_2) \geq d^2(s_1, s_2) \text{ for all } s_1, s_2 \in S.$$

Note the reverse direction of  $\sqsubseteq$  and  $\geq$  in the above definition. We decided to make this reversal so that  $d_1$  is a greatest fixpoint, in analogy with the characterization of bisimilarity, rather than a least fixpoint. This choice has no impact on any results in this thesis.

**Proposition 4.2.2.** *The set of 1-bounded pseudometrics on  $S$  endowed with the order  $\sqsubseteq$  forms a complete lattice.*

*Proof.* Obviously,  $\sqsubseteq$  is a partial order. The top element is the 1-bounded pseudometric  $\top$  defined by

$$\top(s_1, s_2) = 0.$$

The bottom element is the 1-bounded pseudometric  $\perp$  defined by

$$\perp(s_1, s_2) = \begin{cases} 0 & \text{if } s_1 = s_2 \\ 1 & \text{otherwise.} \end{cases}$$

Let  $D$  be a nonempty set of 1-bounded pseudometrics on  $S$ . The meet of  $D$  is the 1-bounded pseudometric  $\bigsqcap D$  defined by

$$(\bigsqcap D)(s_1, s_2) = \sup_{d \in D} d(s_1, s_2).$$

The join of  $D$  can be expressed in terms of the meet of  $D$  (see Lemma 4.1.13). □

Next, we introduce a function from this complete lattice to itself of which the behavioural pseudometric  $d_1$  is the greatest fixpoint.

**Definition 4.2.3.** *Let  $d$  be a 1-bounded pseudometric on  $S$ . The distance function  $\Delta(d) : S \times S \rightarrow [0, 1]$  is defined by*

$$\Delta(d)(s_1, s_2) = \max \left\{ \sum_{s \in S} f(s)(\pi(s_1, s) - \pi(s_2, s)) \mid f \in (S, d) \twoheadrightarrow [0, 1] \right\}$$

if  $s_1 \rightarrow$  and  $s_2 \rightarrow$ , and

$$\Delta(d)(s_1, s_2) = \begin{cases} 0 & \text{if } s_1 \not\rightarrow \text{ and } s_2 \not\rightarrow \\ 1 & \text{otherwise.} \end{cases}$$

We now prove that  $\Delta(d)$  is also a 1-bounded pseudometric on  $S$ .

**Proposition 4.2.4.**  *$\Delta(d)$  is a 1-bounded pseudometric on  $S$ .*

*Proof.* Note that  $f \in (S, d) \dashrightarrow [0, 1]$  implies  $1 - f \in (S, d) \dashrightarrow [0, 1]$ . Furthermore, if  $s_1 \rightarrow$  and  $s_2 \rightarrow$  then

$$\begin{aligned}
& \sum_{s \in S} (1 - f)(s)(\pi(s_1, s) - \pi(s_2, s)) \\
&= \sum_{s \in S} \pi(s_1, s) - \sum_{s \in S} \pi(s_2, s) + \sum_{s \in S} f(s)(\pi(s_2, s) - \pi(s_1, s)) \\
&= \sum_{s \in S} f(s)(\pi(s_2, s) - \pi(s_1, s)) \quad [\sum_{s \in S} \pi(s_1, s) = \sum_{s \in S} \pi(s_2, s) = 1] \\
&= \sum_{s \in S} f(s)\pi(s_2, s) - \sum_{s \in S} f(s)\pi(s_1, s).
\end{aligned}$$

As a consequence, if  $s_1 \rightarrow$  and  $s_2 \rightarrow$  then

$$\Delta(d)(s_1, s_2) = \max \left\{ \left| \sum_{s \in S} f(s)\pi(s_1, s) - \sum_{s \in S} f(s)\pi(s_2, s) \right| \mid f \in (S, d) \dashrightarrow [0, 1] \right\}.$$

Now that we have this alternative representation of  $\Delta(d)$ , checking that it is a 1-bounded pseudometric (using the three conditions of Definition 3.1.6) is straightforward.  $\square$

To conclude that  $\Delta$  has a greatest fixpoint, it suffices to show that  $\Delta$  is order-preserving.

**Proposition 4.2.5.**  *$\Delta$  is order-preserving.*

*Proof.* Let  $d^1$  and  $d^2$  be 1-bounded pseudometrics on  $S$  with  $d^1 \sqsubseteq d^2$ . To prove that  $\Delta$  is order-preserving, we have to prove that  $\Delta(d^1) \sqsubseteq \Delta(d^2)$ .

Assume that  $f \in (S, d^2) \dashrightarrow [0, 1]$ . Then

$$\begin{aligned}
& |f(s_1) - f(s_2)| \\
& \leq d^2(s_1, s_2) \quad [f \in (S, d^2) \dashrightarrow [0, 1]] \\
& \leq d^1(s_1, s_2) \quad [d^1 \sqsubseteq d^2, d^2(s_1, s_2) \leq d^1(s_1, s_2)]
\end{aligned}$$

As a consequence,

$$(S, d^1) \dashrightarrow [0, 1] \supseteq (S, d^2) \dashrightarrow [0, 1]. \tag{4.1}$$

We have to show that  $\Delta(d^1)(s_1, s_2) \geq \Delta(d^2)(s_1, s_2)$ . We focus on the only nontrivial case:  $s_1 \rightarrow$  and  $s_2 \rightarrow$ . In this case,

$$\begin{aligned}
& \Delta(d^1)(s_1, s_2) \\
&= \sup \left\{ \sum_{s \in S} f(s)(\pi(s_1, s) - \pi(s_2, s)) \mid f \in (S, d^1) \dashrightarrow [0, 1] \right\} \\
&\geq \sup \left\{ \sum_{s \in S} f(s)(\pi(s_1, s) - \pi(s_2, s)) \mid f \in (S, d^2) \dashrightarrow [0, 1] \right\} \quad [\text{Equation 4.1}] \\
&= \Delta(d^2)(s_1, s_2).
\end{aligned}$$

$\square$

A similar result is presented in [BHMW, Proposition 38].

According to Theorem 4.1.17, the fixpoints of an order-preserving function on a complete lattice form a complete lattice and, hence, the function has a greatest fixpoint. We denote the greatest fixpoint of  $\Delta$  by  $\text{gfp}(\Delta)$ . This greatest fixpoint of  $\Delta$  is also the greatest pre-fixpoint of  $\Delta$  (see, for example, Theorem 4.1.18).

**Theorem 4.2.6.**  $d_1 = \text{gfp}(\Delta)$ .

*Proof.* We first prove that  $d_1$  is a pre-fixpoint of  $\Delta$ . That is, we show that  $d_1 \sqsubseteq \Delta(d_1)$ , i.e.,  $\Delta(d_1)(s_1, s_2) \leq d_1(s_1, s_2)$  for all  $s_1, s_2 \in S$ . To prove this, we distinguish the following three cases.

- If  $s_1 \not\rightarrow$  and  $s_2 \not\rightarrow$  then the property is vacuously true.
- If  $s_1 \not\rightarrow$  and  $s_2 \rightarrow$ , or  $s_1 \rightarrow$  and  $s_2 \not\rightarrow$ , then the formula  $\diamond\text{true}$  witnesses that the states  $s_1$  and  $s_2$  have distance one.
- Assume that  $s_1 \rightarrow$  and  $s_2 \rightarrow$ . According to [BW05, Proposition 39], the set  $\{\llbracket \varphi \rrbracket_1 \mid \varphi \in \mathcal{L}\}$  is dense in  $(S, d_1) \dashv\vdash [0, 1]$ . As a consequence,

$$\begin{aligned}
& \Delta(d_1)(s_1, s_2) \\
&= \max \left\{ \sum_{s \in S} f(s)(\pi(s_1, s) - \pi(s_2, s)) \mid f \in (S, d_1) \dashv\vdash [0, 1] \right\} \\
&= \max \left\{ \sum_{s \in S} \llbracket \varphi \rrbracket_1(s)(\pi(s_1, s) - \pi(s_2, s)) \mid \varphi \in \mathcal{L} \right\} \\
&= \max \left\{ \sum_{s \in S} \pi(s_1, s) \llbracket \varphi \rrbracket_1(s) - \sum_{s \in S} \pi(s_2, s) \llbracket \varphi \rrbracket_1(s) \mid \varphi \in \mathcal{L} \right\} \\
&= \max \left\{ \llbracket \diamond\varphi \rrbracket_1(s_1) - \llbracket \diamond\varphi \rrbracket_1(s_2) \mid \varphi \in \mathcal{L} \right\} \\
&\leq d_1(s_1, s_2). \quad [d_1(s_1, s_2) = \max \{ \llbracket \phi \rrbracket_1(s_1) - \llbracket \phi \rrbracket_1(s_2) \mid \phi \in \mathcal{L} \}]
\end{aligned}$$

Next we prove that  $d_1$  is the greatest pre-fixpoint of  $\Delta$ . Assume that  $d$  is a pre-fixpoint of  $\Delta$ , i.e.,  $d \sqsubseteq \Delta(d)$ . We have to show that  $d \sqsubseteq d_1$ . That is,  $d_1(s_1, s_2) \leq d(s_1, s_2)$  for all  $s_1, s_2 \in S$ . We restrict our attention to the case that  $s_1 \rightarrow$  and  $s_2 \rightarrow$ . Since

$$d_1(s_1, s_2) = \max \left\{ \llbracket \phi \rrbracket_1(s_1) - \llbracket \phi \rrbracket_1(s_2) \mid \phi \in \mathcal{L} \right\},$$

it suffices to show that

$$\llbracket \varphi \rrbracket_1(s_1) - \llbracket \varphi \rrbracket_1(s_2) \leq d(s_1, s_2)$$

for all  $\varphi$ . This can be proved by structural induction on  $\varphi$ . We consider only the nontrivial case:  $\diamond\varphi$ .

$$\begin{aligned}
& \llbracket \diamond\varphi \rrbracket_1(s_1) - \llbracket \diamond\varphi \rrbracket_1(s_2) \\
&= \sum_{s \in S} \pi(s_1, s) \llbracket \varphi \rrbracket_1(s) - \sum_{s \in S} \pi(s_2, s) \llbracket \varphi \rrbracket_1(s)
\end{aligned}$$

$$\begin{aligned}
&= \sum_{s \in S} \llbracket \varphi \rrbracket_1(s) (\pi(s_1, s) - \pi(s_2, s)) \\
&\leq \max \left\{ \sum_{s \in S} f(s) (\pi(s_1, s) - \pi(s_2, s)) \mid f \in (S, d) \twoheadrightarrow [0, 1] \right\} \\
&\quad \text{[by induction, } \llbracket \varphi \rrbracket_1 \in (S, d) \twoheadrightarrow [0, 1]\text{]} \\
&= \Delta(d)(s_1, s_2) \\
&\leq d(s_1, s_2) \quad [d \text{ is a pre-fixpoint of } \Delta, \Delta(d)(s_1, s_2) \leq d(s_1, s_2)]
\end{aligned}$$

□

A similar result can be obtained by combining Theorem 40 and 44 of [BHMW].

### 4.3 Reaching a Fixpoint by Iteration

In this section we present an iterative approach to compute  $d_1$ . Using definition 4.1.19, we know that  $d^\omega = \prod_{n \in \omega} d^n$ . We prove that the closure ordinal of  $\Delta$  is  $\omega$ , that is, we show that  $\Delta(d^\omega) = d^\omega$ . As a consequence,  $d^\omega$  is the greatest fixpoint of  $\Delta$ . The proof makes use of a number of results which are proved below.

States having distance zero defines an equivalence relation.

**Definition 4.3.1.** *The relation  $\equiv_d$  on states is defined by*

$$s_1 \equiv_d s_2 \text{ if } d(s_1, s_2) = 0.$$

Obviously,  $\equiv_d$  is an equivalence relation. As earlier, we denote the equivalence class that contains the state  $s$  by  $[s]_d$ . From each equivalence class  $[s]_d$  we pick a designated state which we denote by  $\langle s \rangle_d$ . Hence,  $\langle s \rangle_d \in [s]_d$  and also  $d(s, \langle s \rangle_d) = 0$ .

**Proposition 4.3.2.** *For all  $s_1, s_2 \in S$ ,*

$$d(\langle s_1 \rangle_d, \langle s_2 \rangle_d) = d(s_1, s_2).$$

*Proof.*

$$\begin{aligned}
&d(\langle s_1 \rangle_d, \langle s_2 \rangle_d) \\
&\leq d(\langle s_1 \rangle_d, s_1) + d(s_1, \langle s_2 \rangle_d) \quad \text{[triangle inequality]} \\
&\leq d(\langle s_1 \rangle_d, s_1) + d(s_1, s_2) + d(s_2, \langle s_2 \rangle_d) \quad \text{[triangle inequality]} \\
&= d(s_1, s_2) \quad [d(\langle s_1 \rangle_d, s_1) = d(s_2, \langle s_2 \rangle_d) = 0] \\
&\leq d(s_1, \langle s_1 \rangle_d) + d(\langle s_1 \rangle_d, \langle s_2 \rangle_d) + d(\langle s_2 \rangle_d, s_2) \quad \text{[triangle inequality twice]} \\
&= d(\langle s_1 \rangle_d, \langle s_2 \rangle_d).
\end{aligned}$$

Therefore we have proved that  $d(\langle s_1 \rangle_d, \langle s_2 \rangle_d) = d(s_1, s_2)$ .

□

**Definition 4.3.3.** *Let  $d^1 \sqsubseteq d^2$ . The ratio  $\rho(d^1, d^2)$  of  $d^1$  and  $d^2$  is defined by*

$$\rho(d^1, d^2) = \min_{s_1, s_2 \in S} \left\{ \frac{d^2(s_1, s_2)}{d^1(s_1, s_2)} \mid d^2(s_1, s_2) > 0 \right\}$$



We use the convention that the minimum of the empty set is one and the maximum of the empty set is zero. Note that we never divide by zero since  $d^1 \sqsubseteq d^2$  and, hence,  $d^1(s_1, s_2) \geq d^2(s_1, s_2)$ , i.e.,  $d^1(s_1, s_2) \geq 0$ .

Given pseudometrics  $d^1$  and  $d^2$  such that  $d^1 \sqsubseteq d^2$  and given an  $f \in (S, d^1) \twoheadrightarrow [0, 1]$ , we next show that there exists a  $g_f \in (S, d^2) \rightarrow [0, 1]$  that is nonexpansive.

**Proposition 4.3.4.** *Let  $d^1 \sqsubseteq d^2$  and  $f \in (S, d^1) \twoheadrightarrow [0, 1]$ . Let  $g_f : S \rightarrow [0, 1]$  be defined by*

$$g_f(s) = \rho(d^1, d^2)f(\langle s \rangle_{d^2}).$$

*Then  $g_f \in (S, d^2) \twoheadrightarrow [0, 1]$ .*

*Proof.* Let  $s_1, s_2 \in S$ . We have to show that  $g_f$  is nonexpansive, i.e.,

$$|g_f(s_1) - g_f(s_2)| \leq d^2(s_1, s_2).$$

We distinguish two cases.

Case 1: If  $d^2(s_1, s_2) = 0$  then  $\langle s_1 \rangle_{d^2} = \langle s_2 \rangle_{d^2}$  and, hence,  $f(\langle s_1 \rangle_{d^2}) = f(\langle s_2 \rangle_{d^2})$ . Therefore

$$\begin{aligned} g_f(s_1) &= \rho(d^1, d^2)f(\langle s_1 \rangle_{d^2}) \\ &= \rho(d^1, d^2)f(\langle s_2 \rangle_{d^2}) \\ &= g_f(s_2). \end{aligned}$$

and, hence, the property is vacuously true.

Case 2: Let  $d^2(s_1, s_2) > 0$ . According to Proposition 4.3.2,  $d^2(\langle s_1 \rangle_{d^2}, \langle s_2 \rangle_{d^2}) = d^2(s_1, s_2)$ . Hence  $d^2(\langle s_1 \rangle_{d^2}, \langle s_2 \rangle_{d^2}) > 0$ . Also  $d^1(s_1, s_2) > 0$  since  $d^1 \sqsubseteq d^2$ , and

$$\begin{aligned} |g_f(s_1) - g_f(s_2)| &= |\rho(d^1, d^2)f(\langle s_1 \rangle_{d^2}) - \rho(d^1, d^2)f(\langle s_2 \rangle_{d^2})| \\ &= \rho(d^1, d^2)|f(\langle s_1 \rangle_{d^2}) - f(\langle s_2 \rangle_{d^2})| \\ &\leq \rho(d^1, d^2)d^1(\langle s_1 \rangle_{d^2}, \langle s_2 \rangle_{d^2}) \quad [f \in (S, d^1) \twoheadrightarrow [0, 1]] \\ &\leq \frac{d^2(\langle s_1 \rangle_{d^2}, \langle s_2 \rangle_{d^2})}{d^1(\langle s_1 \rangle_{d^2}, \langle s_2 \rangle_{d^2})}d^1(\langle s_1 \rangle_{d^2}, \langle s_2 \rangle_{d^2}) \quad [\text{Definition 4.3.3}] \\ &= d^2(\langle s_1 \rangle_{d^2}, \langle s_2 \rangle_{d^2}) \\ &= d^2(s_1, s_2) \quad [\text{Proposition 4.3.2}] \end{aligned}$$

Therefore  $|g_f(s_1) - g_f(s_2)| \leq d^2(s_1, s_2)$ , i.e.,  $g_f$  is nonexpansive.  $\square$

Next, we bound  $f - g_f$  from above.

**Proposition 4.3.5.** *Let  $d^1 \sqsubseteq d^2$ . Let  $\mu = \min_{s_1, s_2 \in S} \{d^1(s_1, s_2) \mid d^1(s_1, s_2) > 0\}$ . Then*

$$f(s) - g_f(s) \leq \frac{\mu + 1}{\mu} \max_{s'_1, s'_2 \in S} (d^1(s'_1, s'_2) - d^2(s'_1, s'_2))$$

*for all  $s \in S$ .*

*Proof.* Let  $s \in S$ . Then

$$\begin{aligned}
& f(s) - g_f(s) \\
&= f(s) - \rho(d^1, d^2) f(\langle s \rangle_{d^2}) \quad [\text{Proposition 4.3.4}] \\
&= (f(s) - f(\langle s \rangle_{d^2})) + (f(\langle s \rangle_{d^2}) - \rho(d^1, d^2) f(\langle s \rangle_{d^2})).
\end{aligned}$$

Furthermore,

$$\begin{aligned}
& f(s) - f(\langle s \rangle_{d^2}) \\
&\leq d^1(s, \langle s \rangle_{d^2}) \quad [f \in (S, d^1) \rightarrow [0, 1]] \\
&= d^1(s, \langle s \rangle_{d^2}) - d^2(s, \langle s \rangle_{d^2}) \quad [d^2(s, \langle s \rangle_{d^2}) = 0] \\
&\leq \max_{s'_1, s'_2 \in S} d^1(s'_1, s'_2) - d^2(s'_1, s'_2).
\end{aligned}$$

and

$$\begin{aligned}
& f(\langle s \rangle_{d^2}) - \rho(d^1, d^2) f(\langle s \rangle_{d^2}) \\
&= f(\langle s \rangle_{d^2}) (1 - \rho(d^1, d^2)) \\
&\leq 1 - \rho(d^1, d^2) \\
&= 1 - \min_{s_1, s_2 \in S} \left\{ \frac{d^2(s_1, s_2)}{d^1(s_1, s_2)} \mid d^2(s_1, s_2) > 0 \right\} \\
&= \max_{s_1, s_2 \in S} \left\{ 1 - \frac{d^2(s_1, s_2)}{d^1(s_1, s_2)} \mid d^2(s_1, s_2) > 0 \right\} \\
&= \max_{s_1, s_2 \in S} \left\{ \frac{d^1(s_1, s_2) - d^2(s_1, s_2)}{d^1(s_1, s_2)} \mid d^2(s_1, s_2) > 0 \right\} \\
&\leq \frac{\max_{s_1, s_2 \in S} \left\{ d^1(s_1, s_2) - d^2(s_1, s_2) \mid d^2(s_1, s_2) > 0 \right\}}{\min_{s_1, s_2 \in S} \left\{ d^1(s_1, s_2) \mid d^1(s_1, s_2) > 0 \right\}} \\
&\leq \frac{1}{\mu} \max_{s_1, s_2 \in S} \left\{ d^1(s_1, s_2) - d^2(s_1, s_2) \mid d^2(s_1, s_2) > 0 \right\} \\
&\leq \frac{1}{\mu} \max_{s'_1, s'_2 \in S} (d^1(s'_1, s'_2) - d^2(s'_1, s'_2)).
\end{aligned}$$

Therefore we get

$$f(s) - g_f(s) \leq \frac{\mu + 1}{\mu} \max_{s'_1, s'_2 \in S} (d^1(s'_1, s'_2) - d^2(s'_1, s'_2))$$

□

Now we can prove that  $\Delta$  is a Lipschitz function, that is, for  $d^1 \sqsubseteq d^2$ ,

$$\max_{s_1, s_2 \in S} (\Delta(d^1)(s_1, s_2) - \Delta(d^2)(s_1, s_2)) \leq \lambda \max_{s'_1, s'_2 \in S} (d^1(s'_1, s'_2) - d^2(s'_1, s'_2)).$$

for some constant  $\lambda$ .

**Proposition 4.3.6.** *Let  $d^1 \sqsubseteq d^2$ . For all  $s_1, s_2 \in S$ ,*

$$\Delta(d^1)(s_1, s_2) - \Delta(d^2)(s_1, s_2) \leq |S| \frac{\mu + 1}{\mu} \max_{s'_1, s'_2 \in S} (d^1(s'_1, s'_2) - d^2(s'_1, s'_2)).$$

*Proof.* Let  $s_1, s_2 \in S$ . Then

$$\begin{aligned} & \Delta(d^1)(s_1, s_2) - \Delta(d^2)(s_1, s_2) \\ &= \max \left\{ \sum_{s \in S} f(s)(\pi(s_1, s) - \pi(s_2, s)) \mid f \in (S, d^1) \not\rightarrow [0, 1] \right\} - \\ & \quad \max \left\{ \sum_{s \in S} g(s)(\pi(s_1, s) - \pi(s_2, s)) \mid g \in (S, d^2) \not\rightarrow [0, 1] \right\} \\ &= \max \left\{ \min \left\{ \sum_{s \in S} f(s)(\pi(s_1, s) - \pi(s_2, s)) - \sum_{s \in S} g(s)(\pi(s_1, s) - \pi(s_2, s)) \right. \right. \\ & \quad \left. \left. \mid g \in (S, d^2) \not\rightarrow [0, 1] \right\} \mid f \in (S, d^1) \not\rightarrow [0, 1] \right\} \\ &= \max \left\{ \min \left\{ \sum_{s \in S} (f(s) - g(s))(\pi(s_1, s) - \pi(s_2, s)) \right. \right. \\ & \quad \left. \left. \mid g \in (S, d^2) \not\rightarrow [0, 1] \right\} \mid f \in (S, d^1) \not\rightarrow [0, 1] \right\} \\ &\leq \max \left\{ \sum_{s \in S} (f(s) - g_f(s))(\pi(s_1, s) - \pi(s_2, s)) \mid f \in (S, d^1) \not\rightarrow [0, 1] \right\} \\ & \quad \text{[Proposition 4.3.4]} \\ &\leq \max \left\{ \sum_{s \in S} f(s) - g_f(s) \mid f \in (S, d^1) \not\rightarrow [0, 1] \right\} \quad [\pi(s_1, s) - \pi(s_2, s) \leq 1] \\ &\leq \sum_{s \in S} \left\{ \frac{\mu + 1}{\mu} \max_{s'_1, s'_2 \in S} d^1(s'_1, s'_2) - d^2(s'_1, s'_2) \right\} \quad \text{[Proposition 4.3.5]} \\ &\leq |S| \frac{\mu + 1}{\mu} \left( \max_{s'_1, s'_2 \in S} d^1(s'_1, s'_2) - d^2(s'_1, s'_2) \right) \end{aligned}$$

□

We now prove that, for  $\Delta$ , we need to iterate at most  $\omega$  times before reaching the greatest fixpoint. As we will see below, the fact that  $d^\omega$  is a fixpoint of  $\Delta$  follows from the facts that  $\Delta$  is order-preserving (Proposition 4.2.5) and Lipschitz (Proposition 4.3.6). A seemingly obvious way to try to prove that the closure ordinal of  $\Delta$  is  $\omega$  is trying to prove that  $\Delta$  is continuous. However, this seems just as difficult if not more difficult than our current proof.

**Proposition 4.3.7.**  $\Delta(d^\omega) = d^\omega$ .

*Proof.* First, we show that  $\Delta(d^\omega) \sqsubseteq d^\omega$ . By definition,  $d^\omega = \prod_{n \in \omega} d^n$ . Hence  $d^\omega \sqsubseteq d^n$  for all  $n \in \omega$ . Since  $\Delta$  is order-preserving,  $\Delta(d^\omega) \sqsubseteq \Delta(d^n) = d^{n+1}$  for all  $n \in \omega$ . Obviously,  $\Delta(d^\omega) \sqsubseteq d^0$ .

Therefore,  $\Delta(d^\omega)$  is a lowerbound of  $\{d^n \mid n \in \omega\}$ . Since  $d^\omega$  is the greatest lowerbound by definition, we get,  $\Delta(d^\omega) \sqsubseteq d^\omega$ .

We now have left to show that  $d^\omega \sqsubseteq \Delta(d^\omega)$ , i.e.,

$$\Delta(d^\omega)(s_1, s_2) \leq d^\omega(s_1, s_2) \text{ for all } s_1, s_2 \in S. \quad (4.2)$$

Let  $s_1, s_2 \in S$ . Let  $\epsilon > 0$ . We prove that it suffices to show that there exists an  $n$  such that

$$\Delta(d^\omega)(s_1, s_2) - d^{n+1}(s_1, s_2) \leq \epsilon. \quad (4.3)$$

Towards a contradiction, assume that Equation 4.3 holds and Equation 4.2 does not hold which means  $\Delta(d^\omega)(s_1, s_2) > d^\omega(s_1, s_2)$  for all  $s_1, s_2 \in S$ . Let  $\epsilon = \Delta(d^\omega)(s_1, s_2) - d^\omega(s_1, s_2)$ . Since Equation 4.2 does not hold,  $\epsilon > 0$ . Now consider Equation 4.3 for  $\frac{\epsilon}{2}$ . Then  $\Delta(d^\omega)(s_1, s_2) - d^{n+1}(s_1, s_2) \leq \frac{\epsilon}{2}$  for some  $n$ . But we know that  $d^{n+1}(s_1, s_2) \leq d^\omega(s_1, s_2)$ . So we have a contradiction.

Let  $\mu = \min\{d^\omega(s_1, s_2) \mid d^\omega(s_1, s_2) > 0\}$ . Since the set  $S$  is finite, for every  $\delta > 0$  there exists an  $n$  such that for all  $s'_1, s'_2 \in S$ ,

$$d^\omega(s'_1, s'_2) - d^n(s'_1, s'_2) \leq \delta. \quad (4.4)$$

We prove this as follows: Let  $\delta > 0$ . Let  $s'_1, s'_2 \in S$ . By definition,  $d^\omega(s'_1, s'_2) = \sup_{n \in \omega} d^n(s'_1, s'_2)$ . Hence, there exists an  $N_{s'_1, s'_2}$  such that for all  $n \geq N_{s'_1, s'_2}$ ,  $d^\omega(s'_1, s'_2) - d^n(s'_1, s'_2) \leq \delta$ . Now take  $N = \max\{N_{s'_1, s'_2} \mid s'_1, s'_2 \in S\}$ . Note that this maximum exists since the set  $S$  is finite. For all  $n \geq N$  and for all  $s'_1, s'_2 \in S$ , we have that

$$d^\omega(s'_1, s'_2) - d^n(s'_1, s'_2) \leq \delta.$$

Here we pick  $\delta$  to be  $\frac{\mu\epsilon}{(\mu+1)|S|}$ . From Proposition 4.3.6 we can conclude that

$$\begin{aligned} & \Delta(d^\omega)(s_1, s_2) - d^{n+1}(s_1, s_2) \\ &= \Delta(d^\omega)(s_1, s_2) - \Delta(d^n)(s_1, s_2) \\ &\leq |S| \frac{\mu+1}{\mu} \max_{s'_1, s'_2 \in S} d^\omega(s'_1, s'_2) - d^n(s'_1, s'_2) \\ &\leq |S| \frac{\mu+1}{\mu} \delta \quad [\text{Equation 4.4}] \\ &\leq |S| \frac{\mu+1}{\mu} \frac{\mu\epsilon}{(\mu+1)|S|} \\ &\leq \epsilon. \end{aligned}$$

□

We now prove that  $d^\omega$  is the greatest fixpoint of  $\Delta$ .

**Corollary 4.3.8.**  $\text{gfp}(\Delta) = d^\omega$ .

*Proof.* From Proposition 4.3.7,  $d^\omega = \Delta(d^\omega)$  which means  $d^\omega$  is a fixpoint of  $\Delta$ . From Proposition 4.1.20,  $d \sqsubseteq d^\omega$  for any other fixpoint  $d$  of  $\Delta$ . □

**Example 4.3.9.** Consider the probabilistic transition system shown in Figure 1.2. It can be shown that the exact distance between states  $s_2$  and  $s_3$  cannot be computed in less than  $\omega$  iterations.

## 5 The Algorithm: Reduction to First Order Theory over Reals

In this chapter we present the reduction of the problem of approximating  $d_1$  to deciding the satisfiability of a formula of the first order theory over reals. It has been proved by Tarski that the first order theory over reals is decidable and a number of algorithms exist to decide the satisfiability of a formula of this theory. In this chapter we present our reduction and our algorithm for approximating  $d_1$  based on this reduction. We believe that ours is the first algorithm to approximate  $d_1$ . The algorithm is based on expressing the fact  $d_1 = \text{gfp}(\Delta)$  in the theory.

First, in Section 5.1, we present the dualization of the definition of  $\Delta$ . The dualization helps us in obtaining a formula in the existential fragment of the first order theory over reals. For this fragment, there exist more efficient algorithms. In Section 5.2 we briefly recall the first order theory over reals. We give the algorithm in Section 5.3. We also present some optimizations and simplifications of the formula in Section 5.4.

### 5.1 Duality Theorem

In this section we dualize the definition of  $\Delta$  exploiting the Kantorovich-Rubinstein duality theorem [KR58]. As we will see in Section 5.3, this dual characterization will allow us to define  $\Delta$  as the solution to a minimization problem rather than a maximization problem. In turn this will allow us to capture the fact that a pseudometric is a pre-fixpoint of  $\Delta$  in the existential fragment of the first order theory over reals.

Let us recall (a minor variation of) the Kantorovich-Rubinstein duality theorem. Let  $X$  be a 1-bounded compact pseudometric space. Let  $\mu_1$  and  $\mu_2$  be Borel probability measures on  $X$ . We denote the set of Borel probability measures on the product space with marginals  $\mu_1$  and  $\mu_2$  by  $\mu_1 \otimes \mu_2$ , that is, the Borel probability measures  $\mu$  on  $X^2$  such that for all Borel subsets  $B$  of  $X$ ,

$$\mu(B \times X) = \mu_1(B) \text{ and } \mu(X \times B) = \mu_2(B).$$

(See, for example, [Bil95] for an overview of Borel probability measures.) The Kantorovich-Rubinstein duality theorem tells us that

$$\max \left\{ \int_X f d\mu_1 - \int_X f d\mu_2 \mid f \in X \rightarrow [0, 1] \right\} = \min \left\{ \int_{X^2} d_X d\mu \mid \mu \in \mu_1 \otimes \mu_2 \right\}.$$

The following proposition, which is a consequence of the Kantorovich-Rubinstein duality theorem, defines  $\Delta(d)$  as a minimum as opposed to the maximum in Definition 4.2.3.

**Proposition 5.1.1.** *Let  $d$  be a 1-bounded pseudometric on  $S$ . Let  $s_1, s_2 \in S$  such that  $s_1 \rightarrow$  and  $s_2 \rightarrow$ . Then*

$$\Delta(d)(s_1, s_2) = \min \left\{ \sum_{(s_i, s_j) \in S^2} d(s_i, s_j) \mu(s_i, s_j) \mid \mu \in \pi(s_1, \cdot) \otimes \pi(s_2, \cdot) \right\}$$

where  $\mu \in \pi(s_1, \cdot) \otimes \pi(s_2, \cdot)$  if

$$\forall s_j \in S \sum_{s_i \in S} \mu(s_i, s_j) = \pi(s_1, s_j) \wedge \forall s_i \in S \sum_{s_j \in S} \mu(s_i, s_j) = \pi(s_2, s_i).$$

*Proof.* Since the set  $S$  is finite, the space  $(S, d)$  is compact. The probability distributions  $\pi(s_1, \cdot)$  and  $\pi(s_2, \cdot)$  define Borel probability measures on  $(S, d)$ . Applying the Kantorovich-Rubinstein duality theorem gives us the desired result.  $\square$

A similar result is presented in [BW06, Corollary 19].

In the next section we give a brief overview of the first order theory over reals.

## 5.2 First Order Theory over Reals

The first order theory over reals is also called *elementary algebra of real numbers* in the literature. A sentence in the first order theory over reals can have

- variables representing real numbers,
- constants denoting individual integers like ‘0’, ‘1’ and ‘-1’,
- symbols denoting elementary operations and relations like  $+$ ,  $\cdot$ ,  $-$ ,  $<$ ,  $>$  and  $=$ , and
- expressions of elementary logic such as ‘and’ ( $\wedge$ ), ‘or’ ( $\vee$ ), ‘not’ ( $\neg$ ), ‘for some  $x$ ’ ( $\exists x$ ), and ‘for all  $x$ ’ ( $\forall x$ ).

We can use algebraic equations and inequalities combined by means of logical expressions to obtain sentences of this theory. For example, following are a few examples of formulas of this theory.

- $0 < (1 + 1) + (1 + 1)$ .
- $\neg(x > 1) \wedge (\exists y)(x = y \cdot y)$ .
- $\forall a \forall b \forall c \forall d. a \neq 0 \rightarrow \exists x. a \cdot x \cdot x + b \cdot x \cdot x + c \cdot x + d = 0$ .

A variable  $x$  is called a *free variable* if it is not inside the scope of a quantifier ( $\exists x$  or  $\forall x$ ). A formula is called a sentence if it contains no free variables. A sentence is **true** or **false** whereas a formula with free variables will be satisfied by some values of the free variables and not satisfied by the others.

Tarski [Tar51] gave the first decision method for deciding the truth of sentences of this theory. Given any sentence  $\theta$ , a decision method can always decide in a finite number of steps whether  $\theta$  is **true** or **false**.

**Theorem 5.2.1** ([Tar51]). *There is a decision method for the class of all true sentences of elementary algebra of real numbers.*

### 5.3 The Reduction

Now we present a reduction of the problem of approximating  $d_1$  for the states of a probabilistic transition system to the problem of deciding a sentence in the first order theory over reals. We show that the fact that a pseudometric is a pre-fixpoint of  $\Delta$  can be expressed in (the existential fragment of) the first order theory over real numbers. This will allow us to exploit Tarski's decision procedure to approximate the behavioural pseudometric.

For the rest of this chapter, we assume that the probabilistic transition system  $\langle S, \pi \rangle$  has  $N$  states  $s_1, s_2, \dots, s_N$ . Instead of  $\pi(s_i, s_j)$  we will write  $\pi_{ij}$ . We represent a 1-bounded pseudometric on the set  $S$  of states of the probabilistic transition system, as (the values of) a collection of real valued variables  $d_{ij}$ .

The fact that  $d$  is a 1-bounded pseudometric can now be captured in the first order theory over reals as follows. This expression is based on Definition 3.1.6.

**Definition 5.3.1.** *The predicate  $\text{pseudo}(d)$  is defined by*

$$\begin{aligned} \text{pseudo}(d) \equiv & \bigwedge_{1 \leq k, l \leq N} d_{kl} \geq 0 \wedge d_{kl} \leq 1 \wedge \\ & \bigwedge_{1 \leq k \leq N} d_{ik} = 0 \wedge \bigwedge_{1 \leq k, l \leq N} d_{kl} = d_{lk} \wedge \bigwedge_{1 \leq h, k, l \leq N} d_{hl} \leq d_{hk} + d_{kl} \end{aligned}$$

Furthermore, the fact that  $d$  is a pre-fixpoint of  $\Delta$  which means  $d \sqsubseteq \Delta(d)$  or  $\Delta(d)(s_1, s_2) \leq d(s_1, s_2)$  for all  $s_1, s_2 \in S$  can be captured as follows. In the following expression we make use of the expression of  $\Delta$  given in Definition 5.1.1.

**Definition 5.3.2.** *The predicate  $\text{pre-fix}(d)$  is defined by*

$$\text{pre-fix}(d) \equiv \bigwedge_{1 \leq i, j \leq N} \text{pre-fix}_1(d, i, j) \vee \text{pre-fix}_2(d, i, j) \vee \text{pre-fix}_3(d, i, j)$$

where

$$\begin{aligned} \text{pre-fix}_1(d, i, j) \equiv & \sum_{1 \leq k \leq N} \pi_{ik} > 0 \wedge \sum_{1 \leq l \leq N} \pi_{jl} > 0 \wedge \\ & \exists (\mu_{kl})_{1 \leq k, l \leq N} \bigwedge_{1 \leq k, l \leq N} \mu_{kl} \geq 0 \wedge \mu_{kl} \leq 1 \wedge \\ & \bigwedge_{1 \leq l \leq N} \sum_{1 \leq k \leq N} \mu_{kl} = \pi_{il} \wedge \\ & \bigwedge_{1 \leq k \leq N} \sum_{1 \leq l \leq N} \mu_{kl} = \pi_{jk} \wedge \\ & \sum_{1 \leq k, l \leq N} d_{kl} \mu_{kl} \leq d_{ij} \\ \text{pre-fix}_2(d, i, j) \equiv & \sum_{1 \leq k \leq N} \pi_{ik} = 0 \wedge \sum_{1 \leq l \leq N} \pi_{jl} = 0 \wedge 0 \leq d_{ij} \\ \text{pre-fix}_3(d, i, j) \equiv & \left( \left( \sum_{1 \leq k \leq N} \pi_{ik} > 0 \wedge \sum_{1 \leq l \leq N} \pi_{jl} = 0 \right) \vee \right. \end{aligned}$$

$$\left( \sum_{1 \leq k \leq N} \pi_{ik} = 0 \wedge \sum_{1 \leq l \leq N} \pi_{jl} > 0 \right) \wedge 1 \leq d_{ij}$$

The expression  $\text{pre-fix}_1(d, i, j)$  corresponds to the case where  $s_1 \rightarrow$  and  $s_2 \rightarrow$ . (See Definition 4.2.3.) The expressions  $\text{pre-fix}_2(d, i, j)$  and  $\text{pre-fix}_3(d, i, j)$  correspond to the other two cases of Definition 4.2.3. Therefore, the formula  $\text{pseudo}(d) \wedge \text{pre-fix}(d)$  captures the fact that  $d$  is a 1-bounded pseudometric and it is a pre-fixpoint of  $\Delta$ .

**Example 5.3.3.** *The formula  $\text{pseudo}(d) \wedge \text{pre-fix}(d)$  for the probabilistic transition system of Example 5.4.2 is given in Appendix B.*

Now we are ready to present our algorithm. In our algorithm, we use the decision method `tarski` that takes as input a sentence of the first order theory over reals and decides the truth or falsity of the given sentence. Our algorithm called `approximate` is a recursive algorithm similar to binary search. It is invoked for every pair  $s_i$  and  $s_j$  of states where  $s_i \rightarrow$  and  $s_j \rightarrow$ . In the other cases the computation of the distance is trivial as shown in the next section.

The aim of the algorithm is to compute  $d_1(s_i, s_j)$  with accuracy  $\epsilon$  where  $\epsilon$  is a small real number greater than 0. The input to the algorithm is an interval  $[\ell, u]$  and it starts with the initial interval  $[0, 1]$  and returns an interval  $[\ell_0, u_0] \subseteq [0, 1]$  such that  $u_0 - \ell_0 \leq \epsilon$  and  $d_1(s_i, s_j) \in [\ell_0, u_0]$ . We are interested in finding the greatest pre-fixpoint of  $\Delta$ , i.e., if there is another fixpoint  $d'$  then  $d' \sqsubseteq d_1$ , that is  $d_1(s_1, s_2) \leq d'(s_1, s_2)$  for all  $s_1, s_2 \in S$ . This essentially means that we want the smallest value for  $d(s_i, s_j)$  which also satisfies the formula  $\text{pseudo}(d) \wedge \text{pre-fix}(d)$ . Therefore in the algorithm given below we first check the existence of a pre-fixpoint in the left half of the interval  $[\ell, u]$ .

```

approximate( $\ell, u$ ):
pre-condition:  $d_{ij} \in [\ell, u]$ 
  if  $u - \ell \leq \epsilon$ 
    return  $[\ell, u]$ 
  else
     $m = \frac{\ell + u}{2}$ 
    if tarski( $\exists d \text{pseudo}(d) \wedge \text{pre-fix}(d) \wedge d_{ij} \leq m$ )
      return approximate( $\ell, m$ )
    else
      return approximate( $m, u$ )

```

Note that the argument of `tarski` is a sentence that is in the existential fragment of the first order theory over reals. For this fragment there are more efficient decision procedures than for the general theory. (See, for example, [BPR96].)

Let us sketch a correctness proof of our algorithm. Assume that  $d_1(s_i, s_j) \in [\ell, u]$ . We distinguish the following three cases.

- If  $u - \ell \leq \epsilon$ , then the algorithm obviously returns the desired result.
- Assume that  $u - \ell > \epsilon$  and suppose that `tarski` returns true. Then there exists a 1-bounded pseudometric  $d$  that is a pre-fixpoint of  $\Delta$  and  $d(s_i, s_j) \leq m$ . Since  $d_1$  is the greatest pre-fixpoint of  $\Delta$ , we have that  $d \sqsubseteq d_1$ . Hence,  $d_1(s_i, s_j) \leq d(s_i, s_j) \leq m$ . Therefore,  $d_1(s_i, s_j) \in [\ell, m]$ .



- Assume that  $u - \ell > \epsilon$  and suppose that `tarski` returns false. Then  $d(s_i, s_j) > m$  for every 1-bounded pseudometric  $d$  that is a pre-fixpoint of  $\Delta$ . Since  $d_1$  is also a pre-fixpoint of  $\Delta$ , we have that  $d_1(s_i, s_j) > m$ . Therefore,  $d_1(s_i, s_j) \in [m, u]$ .

Obviously, the algorithm terminates.

Current implementations of the decision procedure for the first order theory over reals have time-complexity doubly exponential in the number of variables in the formula. Therefore it is desirable to reduce the number of variables.

## 5.4 Optimizations

In this section we present techniques to optimize the formula by minimizing the number of states of a probabilistic transition system and thereby reduce the size of the formula. First, we classify the states of a probabilistic transition system into three types and show that the distance between certain types of states can be computed directly, i.e., without making use of the algorithm given in the previous section. We also prove that the state space can be minimized by collapsing states that are at zero distance from each other. Then we present some simplifications that can further reduce the size of a formula. The optimized formula for Example 5.4.2 is given at the end of this section.

### 5.4.1 Classification of States

We classify the states of a probabilistic transition system based on their *probabilities of termination* (denoted as  $\tau_\omega$ ), i.e., the probabilities of reaching a terminal state. A terminal state is one which has no outgoing transitions.

Given a state  $s$  and  $n \in \omega + 1$ ,  $\tau_n(s)$  is the probability of terminating in less than  $n$  transitions when started in  $s$ .

**Definition 5.4.1.** For each  $n \in \omega + 1$ , the function  $\tau_n : S \rightarrow [0, 1]$  is defined by

$$\begin{aligned} \tau_0(s) &= 0 \\ \tau_{n+1}(s) &= \begin{cases} 1 & \text{if } s \text{ is a terminal state} \\ \sum_{s' \in S} \pi(s, s') \tau_n(s') & \text{otherwise} \end{cases} \\ \tau_\omega(s) &= \sup_{n \in \omega} \tau_n(s) \end{aligned}$$

**Example 5.4.2.** Consider the probabilistic transition system shown in Figure 5.1. Then we have

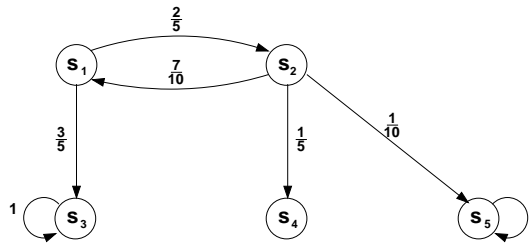


Figure 5.1: An example

that  $\tau_\omega(s_1) = \frac{1}{9}$ ,  $\tau_\omega(s_2) = \frac{5}{18}$ ,  $\tau_\omega(s_3) = 0$ ,  $\tau_\omega(s_4) = 1$  and  $\tau_\omega(s_5) = 0$ .

The classification of states is as follows:

- $S^1 = \{s \in S \mid s \not\rightarrow\}$ : These are states that have no outgoing transitions. We call such states *terminal* states. For such states we know that  $\tau_\omega(s) = 1$ .
- $S^0 = \{s \in S \mid \tau_\omega(s) = 0\}$ : These are states that are not connected to any terminal state. We call such states *non-terminating* states.
- $S^* = \{s \in S \mid s \rightarrow \wedge \tau_\omega(s) > 0\}$ : Such states have at least one path to a terminal state. We call such states *terminating* states.

Obviously,  $S^0$ ,  $S^1$  and  $S^*$  form a partition of  $S$ .

**Example 5.4.3.** Consider the probabilistic transition system shown in Figure 5.1. In this case  $S^1 = \{s_4\}$ ,  $S^0 = \{s_3, s_5\}$  and  $S^* = \{s_1, s_2\}$ .

Now, we prove that the distance between two terminal states is zero and distance between a terminal state and any non-terminal state is one.

**Proposition 5.4.4.**

- If  $s_1 \in S^1$  and  $s_2 \in S^1$  then  $d_1(s_1, s_2) = 0$ .
- If  $s_1 \in S^1$  and  $s_2 \in S^0 \cup S^*$ , or  $s_1 \in S^0 \cup S^*$  and  $s_2 \in S^1$  then  $d_1(s_1, s_2) = 1$ .

*Proof.* We only consider the first case. The second one can be proved similarly. If  $s_1 \not\rightarrow$  and  $s_2 \not\rightarrow$  then from Definition 4.2.3,  $\Delta(d_1)(s_1, s_2) = 0$ . Hence  $d_1(s_1, s_2) = 0$ .  $\square$

**Example 5.4.5.** Consider the probabilistic transition system of Example 5.4.2. State  $s_4$  has distance one to all other states.

Next, we prove that the distance between a non-terminating state and any other state is the probability of termination of the other state.

**Proposition 5.4.6.** If  $s_2 \in S^0$  then  $d_1(s_1, s_2) = \tau_\omega(s_1)$ .

*Proof.* Assume that  $\tau_\omega(s_2) = 0$ . We prove that for all  $n \in \omega + 1$ ,

$$d^n(s_1, s_2) = \tau_n(s_1)$$

by induction on  $n$ .

- Obviously,  $d^0(s_1, s_2) = 0 = \tau_0(s_1)$ .
- Assume  $d^n(s_1, s_2) = 0 = \tau_n(s_1)$ . We have to prove that  $d^{n+1}(s_1, s_2) = \tau_{n+1}(s_1)$ . We distinguish the following two cases.
  - If  $s_1 \not\rightarrow$  then  $d^{n+1}(s_1, s_2) = 1 = \tau_{n+1}(s_1)$ .
  - Now let us assume that  $s_1 \rightarrow$ . First we show that  $\tau_n$  as a function from  $(S, d^n)$  to  $[0, 1]$  is nonexpansive. For all  $s, s' \in S$ ,

$$\begin{aligned} |\tau_n(s) - \tau_n(s')| &= |d^n(s, s_2) - d^n(s', s_2)| \quad [\text{induction}] \\ &\leq d^n(s, s') \quad [\text{triangle inequality}] \end{aligned}$$

Since

$$\begin{aligned}
d^{n+1}(s_1, s_2) &= \Delta(d^n)(s_1, s_2) \\
&\geq \sum_{s \in S} \tau_n(s)(\pi(s_1, s) - \pi(s_2, s)) \\
&\quad [\tau_n \text{ is nonexpansive and using Definition 4.2.3}] \\
&= \sum_{s \in S} \tau_n(s)\pi(s_1, s) - \sum_{s \in S} \tau_n(s)\pi(s_2, s) \\
&= \tau_{n+1}(s_1) - \tau_{n+1}(s_2) \\
&= \tau_{n+1}(s_1) \quad [\tau_\omega(s_2) = 0 \text{ and, hence, } \tau_{n+1}(s_2) = 0]
\end{aligned}$$

We now prove that  $d^{n+1}(s_1, s_2) \leq \tau_{n+1}(s_1)$ . Let  $f \in (S, d^n) \rightarrow [0, 1]$ . For all  $s \in S$ ,

$$f(s) - f(s_2) \leq |f(s) - f(s_2)| \leq d^n(s, s_2) = \tau_n(s). \quad [\text{induction}]$$

As a consequence,

$$\begin{aligned}
&\sum_{s \in S} f(s)(\pi(s_1, s) - \pi(s_2, s)) \\
&= \sum_{s \in S} f(s)\pi(s_1, s) - \sum_{s \in S} f(s)\pi(s_2, s) \\
&= \sum_{s \in S} f(s)\pi(s_1, s) - \sum_{s \in S} f(s)\pi(s_2, s) \\
&\quad - \sum_{s \in S} f(s_2)\pi(s_1, s) + \sum_{s \in S} f(s_2)\pi(s_2, s) \\
&\quad [s_1 \rightarrow \text{ and } s_2 \rightarrow \text{ hence } \sum_{s \in S} \pi(s_1, s) = \sum_{s \in S} \pi(s_2, s) = 1] \\
&= \sum_{s \in S} (f(s) - f(s_2))\pi(s_1, s) - \sum_{s \in S} (f(s) - f(s_2))\pi(s_2, s) \\
&= \sum_{s \in S} (f(s) - f(s_2))(\pi(s_1, s) - \pi(s_2, s)) \\
&\leq \sum_{s \in S} \tau_n(s)(\pi(s_1, s) - \pi(s_2, s)) \\
&= \tau_{n+1}(s_1).
\end{aligned}$$

Since  $f$  was chosen arbitrarily, we can conclude that

$$d^{n+1}(s_1, s_2) \leq \tau_{n+1}(s_1).$$

– Finally,

$$\begin{aligned}
d^\omega(s_1, s_2) &= \sup_n d^n(s_1, s_2) \\
&= \sup_n \tau_n(s_1) \quad [\text{induction}] \\
&= \tau_\omega(s_1).
\end{aligned}$$

From Proposition 4.2.6 and 4.3.8 we can conclude that  $d_1(s_1, s_2) = d^\omega(s_1, s_2) = \tau_\omega(s_1)$ .  $\square$

**Example 5.4.7.** Consider the probabilistic transition system of Example 5.4.2. From Proposition 5.4.6 we can conclude that  $d_1(s_1, s_3) = \frac{1}{9}$ ,  $d_1(s_2, s_3) = \frac{5}{18}$ ,  $d_1(s_4, s_3) = 1$  and  $d_1(s_5, s_3) = 0$ .

We can compute the probability of termination using standard techniques as described in, for example, [GS97, Section 11.2]. In particular, this involves inverting the matrix  $I - P$ , where  $I$  is the identity matrix and  $P$  is a matrix based on  $\pi$ . The details are given in Appendix A.

### 5.4.2 State Space Minimization

Now we show that the state space can be minimized by collapsing states that are at zero distance from each other. Given a probabilistic bisimulation  $\mathcal{R}$ , we can quotient the probabilistic transition system  $\langle S, \pi \rangle$  as follows.

**Definition 5.4.8.** Let  $\mathcal{R}$  be a probabilistic bisimulation. The probabilistic transition system  $\langle S_{\mathcal{R}}, \pi_{\mathcal{R}} \rangle$  consists of

- the set  $S_{\mathcal{R}} = \{ [s]_{\mathcal{R}} \mid s \in S \}$  of  $\mathcal{R}$ -equivalence classes and
- the function  $\pi_{\mathcal{R}} : S_{\mathcal{R}} \times S_{\mathcal{R}} \rightarrow [0, 1]$  defined by

$$\pi_{\mathcal{R}}([s]_{\mathcal{R}}, [s']_{\mathcal{R}}) = \sum_{s'' \mathcal{R} s'} \pi(s, s'').$$

To avoid cluttering of notation, we drop the subscript from  $[s]_{\mathcal{R}}$  in the following discussion. Note that the function  $\pi_{\mathcal{R}}$  is well-defined since  $\mathcal{R}$  is a probabilistic bisimulation. We will apply the above quotient construction for the following bisimulation.

**Proposition 5.4.9.** The smallest equivalence relation containing  $\{ \langle s_1, s_2 \rangle \mid s_1 \in S^1 \wedge s_2 \in S^1 \}$  and  $\{ \langle s_1, s_2 \rangle \mid s_1 \in S^0 \wedge s_2 \in S^0 \}$  is a probabilistic bisimulation.

*Proof.* Let us denote the relation by  $\mathcal{R}$ . Assume that  $s_1 \mathcal{R} s_2$ , i.e.,  $s_1 \not\rightarrow \wedge s_2 \not\rightarrow$  or  $\tau_\omega(s_1) = 0 \wedge \tau_\omega(s_2) = 0$ . From Definition 2.2.10, it suffices to show that  $\sum_{s' \mathcal{R} s} \pi(s_1, s') = \sum_{s' \mathcal{R} s} \pi(s_2, s')$  for each  $s \in S$ . We distinguish the following three cases.

- If  $s_1 = s_2$  then it is vacuously true.
- If  $s_1 \not\rightarrow$  and  $s_2 \not\rightarrow$  then  $\sum_{s' \mathcal{R} s} \pi(s_1, s') = 0$  and  $\sum_{s' \mathcal{R} s} \pi(s_2, s') = 0$ .
- Assume  $\tau_\omega(s_1) = 0$  and  $\tau_\omega(s_2) = 0$ . This means  $s_1$  and  $s_2$  have transitions only to non-terminating states i.e. states  $s$  for which  $\tau_\omega(s) = 0$ . Therefore, if  $\tau_\omega(s) = 0$  then  $\sum_{s' \mathcal{R} s} \pi(s_1, s') = 1$  and  $\sum_{s' \mathcal{R} s} \pi(s_2, s') = 1$ . Otherwise,  $\sum_{s' \mathcal{R} s} \pi(s_1, s') = 0$  and  $\sum_{s' \mathcal{R} s} \pi(s_2, s') = 0$ .

$\square$

The above relation essentially contains those pairs of states that are at distance zero (see Proposition 5.4.4 and Proposition 5.4.6).

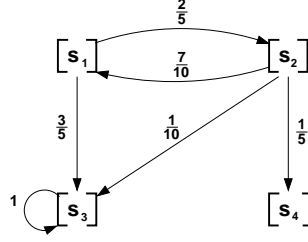


Figure 5.2: An example

**Example 5.4.10.** Consider the probabilistic transition system of Example 5.4.2. According to Proposition 5.4.9, the smallest equivalence relation containing  $\{\langle s_3, s_5 \rangle\}$  is a bisimulation. The resulting quotient can be depicted as

By quotienting, the number of states that need to be considered and, hence, the number of variables in the formula may be reduced. However, we still have to check that the quotiented system gives rise to the same distances. Next we relate the behavioural pseudometric  $d_1$  of the original system  $\langle S, \pi \rangle$  with the behavioural pseudometric  $d_{\mathcal{R}}$  of the quotiented system  $\langle S_{\mathcal{R}}, \pi_{\mathcal{R}} \rangle$ .

**Proposition 5.4.11.** For all  $s_1, s_2 \in S$ ,  $d_{\mathcal{R}}([s_1], [s_2]) = d_1(s_1, s_2)$ .

*Proof.* First of all, note that

$$\sum_{s' \in S} \pi(s, s') = \sum_{[s'] \in S_{\mathcal{R}}} \sum_{s'' \mathcal{R} s'} \pi(s, s'') = \sum_{[s'] \in S_{\mathcal{R}}} \pi_{\mathcal{R}}([s], [s']).$$

As a consequence, we have left to consider the case  $s_1 \rightarrow$  and  $s_2 \rightarrow$ . We prove that for all  $n \in \omega + 1$ ,  $d_{\mathcal{R}}^n([s_1], [s_2]) = d_1^n(s_1, s_2)$  by induction on  $n$ . We distinguish the following three cases.

- If  $n = 0$  then the property is vacuously true.
- Assume that  $d_{\mathcal{R}}^n([s'_1], [s'_2]) = d_1^n(s'_1, s'_2)$  for all  $s'_1, s'_2 \in S$ . Let  $s_1, s_2 \in S$ . We have to prove that  $d_{\mathcal{R}}^{n+1}([s_1], [s_2]) = d_1^{n+1}(s_1, s_2)$ . In the proof of this case, we make use of the following two observations. For each  $f \in (S_{\mathcal{R}}, d_{\mathcal{R}}^n) \rightarrow [0, 1]$ , there exists a  $g \in (S, d_1^n) \rightarrow [0, 1]$  such that  $g(s) = f([s])$  for all  $s \in S$ , since

$$\begin{aligned} |g(s) - g(s')| &= |f([s]) - f([s'])| \\ &\leq d_{\mathcal{R}}^n(s, s') \quad [f \text{ is nonexpansive}] \\ &= d_1^n(s, s') \quad [\text{induction}]. \end{aligned}$$

Similarly, we can show that for each  $g \in (S, d_1^n) \rightarrow [0, 1]$ , there exists  $f \in (S_{\mathcal{R}}, d_{\mathcal{R}}^n) \rightarrow [0, 1]$  such that  $f([s]) = g(s)$  for all  $s \in S$ . Note that if states  $s$  and  $s'$  are probabilistic bisimilar then  $d_1(s, s') = 0$  and, hence,  $d_1^n(s, s') = 0$  and, therefore,  $g(s) = g(s')$ , since  $g$  is nonexpansive.

$$\begin{aligned} &d_{\mathcal{R}}^{n+1}([s_1], [s_2]) \\ &= \Delta(d_{\mathcal{R}}^n)([s_1], [s_2]) \\ &= \max \left\{ \sum_{[s] \in S_{\mathcal{R}}} f([s]) (\pi_{\mathcal{R}}([s_1], [s]) - \pi_{\mathcal{R}}([s_2], [s])) \mid f \in (S_{\mathcal{R}}, d_{\mathcal{R}}^n) \rightarrow [0, 1] \right\} \end{aligned}$$

$$\begin{aligned}
&= \max \left\{ \sum_{[s] \in S_{\mathcal{R}}} f([s]) \sum_{s' \mathcal{R} s} (\pi(s_1, s') - \pi(s_2, s')) \mid f \in (S_{\mathcal{R}}, d_{\mathcal{R}}^n) \twoheadrightarrow [0, 1] \right\} \\
&= \max \left\{ \sum_{[s] \in S_{\mathcal{R}}} \sum_{s' \mathcal{R} s} f([s']) (\pi(s_1, s') - \pi(s_2, s')) \mid f \in (S_{\mathcal{R}}, d_{\mathcal{R}}^n) \twoheadrightarrow [0, 1] \right\} \\
&= \max \left\{ \sum_{s \in S} g(s) (\pi(s_1, s) - \pi(s_2, s)) \mid g \in (S, d_1^n) \twoheadrightarrow [0, 1] \right\} \\
&= \Delta(d_1^n)(s_1, s_2) \\
&= d_1^{n+1}(s_1, s_2).
\end{aligned}$$

- Furthermore,

$$\begin{aligned}
d_{\mathcal{R}}^{\omega}([s_1], [s_2]) &= \sup_n d_{\mathcal{R}}^n([s_1], [s_2]) \\
&= \sup_n d_1^n(s_1, s_2) \quad [\text{induction}] \\
&= d_1^{\omega}(s_1, s_2).
\end{aligned}$$

□

### 5.4.3 Simplifications

To simplify the formula even further, we exploit the following three observations.

- Since  $d$  is a pseudometric,  $d_1(s_i, s_i) = 0$  and  $d_1(s_i, s_j) = d_1(s_j, s_i)$ . Therefore, in  $\text{pseudo}(d) \wedge \text{pre-fix}(d)$  we can replace all  $d_{ii}$ 's with zero and all  $d_{ij}$ 's where  $i > j$  with  $d_{ji}$ 's. As a consequence, we only need to consider  $d_{ij}$ 's with  $i < j$ . This reduces the number of variables in the formula considerably.
- Let  $C$  be the set of pairs of states for which the distances have already been computed using Propositions 5.4.4 and 5.4.6. Then

$$\exists d \text{pseudo}(d) \wedge \text{pre-fix}(d) \wedge d_{i_0 j_0} \leq m$$

is equivalent to

$$\exists d \text{pseudo}(d) \wedge \text{pre-fix}(d) \wedge d_{i_0 j_0} \leq m \wedge \bigwedge_{(i,j) \in C} d_{ij} = d_1(s_i, s_j)$$

since  $d_1$  is the greatest pre-fixpoint. As a consequence, we can replace all  $d_{ij}$ 's where  $(i, j) \in C$  with their already computed distances  $d_1(s_i, s_j)$ . Again, the number of variables may be reduced.

- If  $\pi_{i_0 j} = 0$ , we can infer that  $\mu_{ij} = 0$  for all  $1 \leq i \leq N$ . As a consequence, we can replace the occurrences of all those  $\mu_{ij}$ 's with 0. Symmetrically, if  $\pi_{j_0 i} = 0$  we can simplify the formula similarly. This simplification also may reduce the number of variables.

#### 5.4.4 An Example Formula

Consider the system given in Example 5.4.2. The optimized formula to approximate distance between states  $s_1$  and  $s_2$  is as follows:

$$\begin{aligned}
& \exists d_{12}. \\
& (d_{12} \geq 0 \wedge d_{12} \leq 1) \wedge \\
& \left(\frac{1}{9} \leq d_{12} \leq \frac{5}{18}\right) \wedge \\
& \left(d_{12} \leq \frac{1}{9} + \frac{5}{18}\right) \wedge \\
& \exists(\mu_{12}, \mu_{13}, \mu_{32}, \mu_{42}, \mu_{43}, \mu_{33})\{ \\
& \quad (\mu_{12} \geq 0 \wedge \mu_{12} \leq 1) \wedge \\
& \quad (\mu_{13} \geq 0 \wedge \mu_{13} \leq 1) \wedge \\
& \quad (\mu_{32} \geq 0 \wedge \mu_{32} \leq 1) \wedge \\
& \quad (\mu_{42} \geq 0 \wedge \mu_{42} \leq 1) \wedge \\
& \quad (\mu_{43} \geq 0 \wedge \mu_{43} \leq 1) \wedge \\
& \quad (\mu_{12} + \mu_{32} + \mu_{42} = \frac{2}{5}) \wedge \\
& \quad (\mu_{13} + \mu_{43} + \mu_{33} = \frac{3}{5}) \wedge \\
& \quad (\mu_{12} + \mu_{13} = \frac{7}{10}) \wedge \\
& \quad (\mu_{32} + \mu_{33} = \frac{1}{10}) \wedge \\
& \quad (\mu_{42} + \mu_{43} = \frac{1}{5}) \wedge \\
& \quad \left(d_{12} * \mu_{12} + \frac{1}{9} * \mu_{13} + \frac{5}{18} * \mu_{32} + \mu_{42} + \mu_{43} \leq d_{12}\right)\} \wedge \\
& \exists(\mu_{21}, \mu_{23}, \mu_{24}, \mu_{31}, \mu_{33}, \mu_{34})\{ \\
& \quad (\mu_{21} \geq 0 \wedge \mu_{21} \leq 1) \wedge \\
& \quad (\mu_{23} \geq 0 \wedge \mu_{23} \leq 1) \wedge \\
& \quad (\mu_{24} \geq 0 \wedge \mu_{24} \leq 1) \wedge \\
& \quad (\mu_{31} \geq 0 \wedge \mu_{31} \leq 1) \wedge \\
& \quad (\mu_{34} \geq 0 \wedge \mu_{34} \leq 1) \wedge \\
& \quad (\mu_{21} + \mu_{31} = \frac{7}{10}) \wedge \\
& \quad (\mu_{23} + \mu_{33} = \frac{1}{10}) \wedge \\
& \quad (\mu_{24} + \mu_{34} = \frac{1}{5}) \wedge \\
& \quad (\mu_{21} + \mu_{23} + \mu_{24} = \frac{2}{5}) \wedge \\
& \quad (\mu_{31} + \mu_{33} + \mu_{34} = \frac{3}{5}) \wedge \\
& \quad \left(d_{12} * \mu_{21} + \frac{5}{18} * \mu_{23} + \mu_{24} + \frac{1}{9} * \mu_{31} + \mu_{34} \leq d_{12}\right)\} \wedge \\
& (0 \leq d_{12} \leq \frac{1}{2})
\end{aligned}$$

Note that we used the fact  $d_1(s_1, s_3) = \frac{1}{9}$  and  $d_1(s_2, s_3) = \frac{5}{18}$  in the above formula.

As seen from the unoptimized formula given in Appendix B, using optimizations, we obtained a reduction by a factor of more than thirty in size. This shows that the optimizations and simplifications discussed in this chapter are extremely useful. In fact, we were unable to solve the formula given in Appendix B using the solver *Mathematica*.

## 6 Extension of our Approach to other Models

In this chapter we show that our approach can be applied to approximate distances between other models of systems. Specifically, we show the applicability of our ideas and algorithm to partially defined probabilistic transition systems, labelled probabilistic transition systems and metric-labelled transition systems.

In each of the sections that follow, we first define the model, then a *distance* function for the model and finally, we show how this distance can be approximated using our algorithm.

### 6.1 Partially-defined Probabilistic Systems

In this section we consider partially-defined probabilistic transition systems (Definition 2.1.6), and show how distances between states of such a system can be computed. The logical characterization given by Desharnais et al. in [DGJP04] is also applicable to partial-probabilistic transition systems. Based on their logical characterization we define a distance function for such systems as follows.

**Definition 6.1.1.** *Let  $d$  be a 1-bounded pseudometric on  $S$ . The distance function  $\Delta_p(d) : S \times S \rightarrow [0, 1]$  is defined as*

$$\Delta_p(d)(s_1, s_2) = \max \left\{ \sum_{u \in S} f(u)(\pi(s_1, u) - \pi(s_2, u)) \mid f \in (S, d) \twoheadrightarrow [0, 1] \right\} \\ + \max \left\{ \sum_{u \in S} \pi(s_2, u) - \sum_{u \in S} \pi(s_1, u), 0 \right\}.$$

It can be proved that  $\Delta_p$  is order-preserving.

**Definition 6.1.2.** *The distance function  $d_p : S \times S \rightarrow [0, 1]$  is defined by*

$$d_p = \text{gfp}(\Delta_p).$$

Now we show how  $d_p$  can be computed using our algorithm for probabilistic transition systems. First we build a probabilistic transition system from the partial-probabilistic system using the following construction.

**Definition 6.1.3.** *Construct  $\langle S', \pi' \rangle$  from  $\langle S, \pi \rangle$  as follows:*

- $S' = S \cup \{s_N\}$ , where  $s_N$  is a new state,
- $\pi'(s, s_N) = 1 - \sum_{u \in S} \pi(s, u)$  where  $s \in S$  and  
 $\pi'(s_i, s_j) = \pi(s_i, s_j)$  where  $s_i, s_j \in S$ .

Observe that  $s_N \twoheadrightarrow$  and  $s \rightarrow$  for all  $s \in S$ , hence  $d_1(s_N, s) = 1$ .



**Example 6.1.4.** Consider the partial-probabilistic transition system shown in Figure 2.3. Using Definition 6.1.3, we add a new state  $s_N$  and construct the new system as shown in Figure 6.1. In this system, the sum of probabilities of outgoing transitions for all states except  $s_N$  is 1.

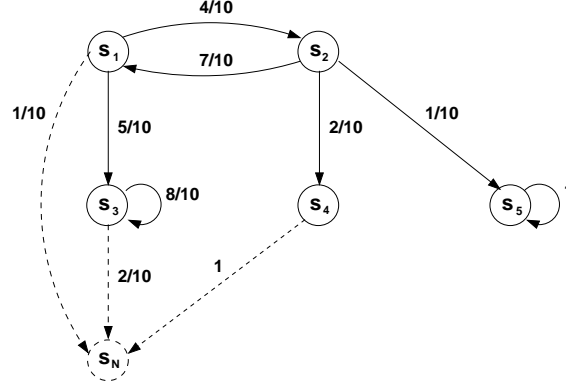


Figure 6.1:  $\langle S', \pi' \rangle$  for Figure 2.3

Next, we show that distances between states of a partial-probabilistic system are the same as distances between the corresponding states of the constructed system.

**Proposition 6.1.5.** For all  $s_1, s_2 \in S$ ,

$$d_p(s_1, s_2) = d_1(s_1, s_2).$$

*Proof.* Let  $d$  be a 1-bounded pseudometric on  $S$ . We define the pseudometric  $d^e$  on  $S'$  by

$$d^e(s, s') = \begin{cases} 0 & \text{if } s = s_N \text{ and } s' = s_N \\ d(s, s') & \text{if } s \in S \text{ and } s' \in S \\ 1 & \text{otherwise} \end{cases}$$

Next, we prove that for all  $s_1, s_2 \in S$ ,

$$\Delta(d^e)(s_1, s_2) = \Delta_p(d)(s_1, s_2).$$

For the probabilistic transition system  $\langle S', \pi' \rangle$ , we know that,

$$\begin{aligned} \Delta(d^e)(s_1, s_2) &= \max \left\{ \sum_{u \in S'} f(u)(\pi'(s_1, u) - \pi'(s_2, u)) \mid f \in (S', d^e) \rightarrow [0, 1] \right\} \\ &\quad \text{[Definition 4.2.3]} \\ &= \max \left\{ \sum_{u \in S} f(u)(\pi'(s_1, u) - \pi'(s_2, u)) \right. \\ &\quad \left. + f(s_N)(\pi'(s_1, s_N) - \pi'(s_2, s_N)) \mid f \in (S', d^e) \rightarrow [0, 1] \right\} \\ &= \max \left\{ \sum_{u \in S} f(u)(\pi(s_1, u) - \pi(s_2, u)) \right\} \end{aligned}$$

$$\begin{aligned}
& + f(s_N) \left( \left( 1 - \sum_{u \in S} \pi(s_1, u) \right) - \left( 1 - \sum_{u \in S} \pi(s_2, u) \right) \right) \\
& \left. \left| f \in (S', d^e) \not\rightarrow [0, 1] \right\} \quad [\text{Definition 6.1.3}] \right. \\
= & \max \left\{ \sum_{u \in S} f(u) (\pi(s_1, u) - \pi(s_2, u)) \right. \\
& \left. + f(s_N) \left( \sum_{u \in S} \pi(s_2, u) - \sum_{u \in S} \pi(s_1, u) \right) \right. \\
& \left. \left| f \in (S', d^e) \not\rightarrow [0, 1] \right\} \tag{6.1}
\end{aligned}$$

We know that  $d^e(s_N, s) = 1$  for all  $s \in S$  and since  $f$  is nonexpansive in  $(S', d^e)$ , from Definition 3.1.9 we have

$$|f(s_N) - f(s)| \leq 1 \quad \text{for all } s \in S.$$

The above condition will be satisfied irrespective of the values of  $f(s_N)$  and  $f(s)$ . Therefore, we can choose any value for  $f(s_N)$ . If

$$\sum_{u \in S} \pi(s_2, u) - \sum_{u \in S} \pi(s_1, u) > 0,$$

we can maximize the value of the expression on the right hand side of Equation 6.1 by choosing  $f(s_N) = 1$ , otherwise we choose  $f(s_N) = 0$ .

Also since  $S \subset S'$  we can define a function  $f^S$  as  $f$  restricted to  $S$  such that

$$f^S \in (S, d) \not\rightarrow [0, 1].$$

Therefore, we get

$$\begin{aligned}
\Delta(d^e)(s_1, s_2) & = \max \left\{ \sum_{u \in S} f^S(u) (\pi(s_1, u) - \pi(s_2, u)) \left| f^S \in (S, d) \not\rightarrow [0, 1] \right. \right\} \\
& \quad + \max \left\{ \sum_{u \in S} \pi(s_2, u) - \sum_{u \in S} \pi(s_1, u), 0 \right\} \\
& = \Delta_p(d)(s_1, s_2) \quad [\text{Definition 6.1.1}]
\end{aligned}$$

From the above, we can prove that

$$\Delta(d^e) = d^e \text{ iff } \Delta_p(d) = d.$$

Since  $d_p$  is a fixpoint of  $\Delta_p$  we can conclude that  $d_p^e$  is a fixpoint of  $\Delta$ . Since  $d_1$  is the greatest fixpoint of  $\Delta$ , we can deduce that  $d_p^e \sqsubseteq d_1$ .

Since  $d_1(s, s_N) = 1$ , we have  $(d_1^S)^e = d_1$  (where we use  $d_1^S$  as the restriction of  $d_1$  to  $S \times S$ ). Since  $d_1$  is a fixpoint of  $\Delta$ , we can conclude that  $d_1^S$  is a fixpoint of  $\Delta_p$ . Since  $d_p$  is the greatest fixpoint of  $\Delta_p$ , we can deduce that  $d_1^S \sqsubseteq d_p$ .  $\square$

This proves that distances between states of a partial-probabilistic system are the same as distances between the corresponding states of the constructed system. Hence,  $d_p$  can be approximated using our algorithm.

**Example 6.1.6.** Consider the example system of Figure 2.3. Using the formula given above, we get  $d_p(s_4, s_5) = 1$  and  $d_p(s_3, s_5) = 1$ .

## 6.2 Labelled Probabilistic Transition Systems

Labelled probabilistic transition systems were defined in Definition 2.1.9. Desharnais et al. considered labelled probabilistic transition system for giving their logical characterization. Based on their logic we arrive at the following distance function.

**Definition 6.2.1.** Let  $d$  be a 1-bounded pseudometric on  $S$ . The distance function  $\Delta_\ell(d) : S \times S \rightarrow [0, 1]$  is defined by

$$\Delta_\ell(d)(s_1, s_2) = \begin{cases} 0 & \text{if } s_1 \not\rightarrow \text{ and } s_2 \not\rightarrow \\ 1 & \text{if } s_1 \xrightarrow{a} \text{ and } s_2 \xrightarrow{a} \text{ or vice versa for some } a \in A, \text{ and} \\ & \left\{ \sum_{s \in S} f(s)(\pi_a(s_1, s) - \pi_a(s_2, s)) \mid f \in (S, d) \rightarrow [0, 1] \right\} \end{cases}$$

otherwise.

It can be proved that  $\Delta_\ell$  is order-preserving.

**Definition 6.2.2.** The distance function  $d_\ell : S \times S \rightarrow [0, 1]$  is defined by

$$d_\ell = \text{gfp}(\Delta_\ell).$$

Similar to Section 5.3, we present the formula in the first order theory of reals below.

The expression for  $\text{pseudo}(d)$  is the same as given in Definition 5.3.1.

**Definition 6.2.3.** Using Definition 6.2.1, the predicate  $\text{pre-fix}(d)$  is defined by

$$\text{pre-fix}(d) \equiv \bigwedge_{1 \leq i, j \leq N} \text{pre-fix}_1(d, i, j) \vee \text{pre-fix}_2(d, i, j) \vee \text{pre-fix}_3(d, i, j)$$

where

$$\text{pre-fix}_1(d, i, j) \equiv \bigwedge_{a \in A} \left( \exists (\mu_{kl})_{1 \leq k, l \leq N} \bigwedge_{1 \leq k, l \leq N} \mu_{kl} \geq 0 \wedge \mu_{kl} \leq 1 \wedge \right. \\ \bigwedge_{1 \leq l \leq N} \sum_{1 \leq k \leq N} \mu_{kl} = \pi_{a_{il}} \wedge \\ \bigwedge_{1 \leq k \leq N} \sum_{1 \leq l \leq N} \mu_{kl} = \pi_{a_{jk}} \wedge \\ \left. \sum_{1 \leq k, l \leq N} d_{kl} \mu_{kl} \leq d_{ij} \right)$$

$$\begin{aligned}
\text{pre-fix}_2(d, i, j) &\equiv \bigwedge_{a \in A} \left( \sum_{1 \leq k \leq N} \pi_{a_{ik}} = 0 \wedge \sum_{1 \leq l \leq N} \pi_{a_{jl}} = 0 \right) \wedge 0 \leq d_{ij} \\
\text{pre-fix}_3(d, i, j) &\equiv \bigvee_{a \in A} \left( \left( \sum_{1 \leq k \leq N} \pi_{a_{ik}} > 0 \wedge \sum_{1 \leq l \leq N} \pi_{a_{jl}} = 0 \right) \vee \right. \\
&\quad \left. \left( \sum_{1 \leq k \leq N} \pi_{a_{ik}} = 0 \wedge \sum_{1 \leq l \leq N} \pi_{a_{jl}} > 0 \right) \right) \wedge \\
&\quad 1 \leq d_{ij}
\end{aligned}$$

In the above expression,  $\text{pre-fix}_2(d, i, j)$  corresponds to the case when  $s_1 \not\rightarrow$  and  $s_2 \not\rightarrow$  in Definition 6.2.1. The expression  $\text{pre-fix}_3(d, i, j)$  corresponds to the case when  $s_1 \xrightarrow{a}$  and  $s_2 \xrightarrow{a}$  or vice versa for some  $a \in A$  and  $\text{pre-fix}_1(d, i, j)$  corresponds to the remaining case.

Since the formula  $\text{pseudo}(d) \wedge \text{prefix}(d)$  belongs to the first order theory over reals, we can now use our algorithm to approximate distances between states of a labelled probabilistic transition system.

### 6.3 Metric-labelled Transition Systems

Metric-labelled transition systems are labelled transition systems whose actions (labels) form a metric space. For example, systems where transitions are labelled with time and the distance between two labels is equal to the absolute difference between them can be modelled as metric-labelled transition systems. See [Bre05] for details.

**Definition 6.3.1.** *A metric-labelled transition system is a tuple  $\langle S, A, \rightarrow \rangle$ , consisting of*

- a finite set  $S$  of states,
- a finite metric space  $A$  of actions, and
- a labelled transition relation  $\rightarrow \subseteq S \times A \times S$ .

Instead of  $(s, a, s') \in \rightarrow$  we often write  $s \xrightarrow{a} s'$ .

A slightly modified (for finite states) definition of the distance function given by Van Breugel in [Bre05, Definition 7] follows.

**Definition 6.3.2.** *Let  $d$  be a pseudometric on the set  $S$  of states,  $d : S \times S \rightarrow [0, \infty]$ , and  $d_A$  be the metric on the set  $A$  of actions. The distance function  $\Delta_m(d) : S \times S \rightarrow [0, \infty]$  is defined by*

$$\Delta_m(d)(s_1, s_2) = \max \left\{ \begin{aligned} &\max_{s_1 \xrightarrow{a_1} s'_1} \min_{s_2 \xrightarrow{a_2} s'_2} ( d_A(a_1, a_2) + d(s'_1, s'_2) ), \\ &\max_{s_2 \xrightarrow{a_2} s'_2} \min_{s_1 \xrightarrow{a_1} s'_1} ( d_A(a_1, a_2) + d(s'_1, s'_2) ) \end{aligned} \right\}$$

Note that the pseudometric gives a real number between zero and infinity (including infinity). This is different from the pseudometrics defined earlier in this thesis where the pseudometrics were 1-bounded.

Van Breugel proved that  $\Delta_m$  is order-preserving and defined the distance function as follows.

**Definition 6.3.3.** *The distance function  $d_m : S \times S \rightarrow [0, \infty]$  is defined by*

$$d_m = \text{gfp}(\Delta_m).$$

**Example 6.3.4.** *Consider the metric-labelled transition system shown in Figure 6.2 where distance between two labels is equal to the absolute difference between them. In this case the distance*



Figure 6.2:  $d_m(s_1, s_2) = \infty$

between states  $s_1$  and  $s_2$  equals infinity.

Now, we present our reduction to the first order theory over reals. The fact that  $d$  is a pseudometric can be captured in the first order theory over reals as follows.

**Definition 6.3.5.** *The predicate  $\text{pseudo}(d)$  is defined by*

$$\begin{aligned} \text{pseudo}(d) \equiv & \bigwedge_{1 \leq i, j \leq N} d_{ij} \geq 0 \wedge \bigwedge_{1 \leq i \leq N} d_{ii} = 0 \wedge \bigwedge_{1 \leq i, j \leq N} d_{ij} = d_{ji} \wedge \\ & \bigwedge_{1 \leq h, i, j \leq N} d_{hj} \leq d_{hi} + d_{ij} \end{aligned}$$

Furthermore, the fact that  $d$  is a pre-fixpoint of  $\Delta_m$  which means  $d \sqsubseteq \Delta_m(d)$  or  $\Delta_m(d)(s_1, s_2) \leq d(s_1, s_2)$  for all  $s_1, s_2 \in S$  can be captured as follows. We make use of the expression of  $\Delta_m$  given in Definition 6.3.2.

**Definition 6.3.6.** *The predicate  $\text{pre-fix}(d)$  is defined by*

$$\begin{aligned} \text{pre-fix}(d) \equiv & \bigwedge_{1 \leq i, j \leq N} \left( \bigwedge_{i \xrightarrow{a_1} k} \bigvee_{j \xrightarrow{a_2} l} (d_A(a_1, a_2) + d_{kl} \leq d_{ij}) \wedge \right. \\ & \left. \bigwedge_{j \xrightarrow{a_2} l} \bigvee_{i \xrightarrow{a_1} k} (d_A(a_1, a_2) + d_{kl} \leq d_{ij}) \right) \end{aligned}$$

where  $d_{ij}$  is a real variable representing distance between states  $s_i$  and  $s_j$  and  $i \xrightarrow{a} j$  if  $(s_i, a, s_j) \in \rightarrow$ .

We can now use our algorithm to approximate  $d_m$  distances.

## 7 Implementation

We saw that a decision procedure for the first order theory over reals is the key component of our algorithm. In this chapter we discuss the complexity of various algorithms available for deciding a formula of the first order theory over reals. We present the implementation of our algorithm based on the solver *Mathematica* with the help of an example.

### 7.1 Decision Procedures for the First Order Theory over Reals

A decision procedure for the first order theory over reals based on quantifier elimination was first given by Tarski [Tar51] (which he had discovered in 1930). Tarski's decision method is often called the 'method of eliminating quantifiers'. It consists of two components. The first component is a procedure which, given any sentence in the first order theory over reals, finds in a mechanical way an equivalent sentence without quantifiers. The second component takes as input a sentence without quantifiers and decides in a mechanical way whether it is **true** or not. As noted by Tarski, any quantifier elimination method for this theory also provides a decision method. Tarski's algorithm for quantifier elimination has non-elementary complexity, meaning that no tower  $2^{2^{\dots^n}}$  can bound the execution time of the algorithm if  $n$  is the size of the problem. A number of algorithms have been developed thereafter for the theory (see, for example, [BPR96, Col75, Hör05]). Among them Collins' algorithm [Col75, CH91] based on *cylindrical algebraic decomposition* is regarded as most significant. Collins proved that the time complexity for his quantifier elimination algorithm for a formula  $\phi$  in prenex normal form is

$$O((md)^{c^n} \lambda^c)$$

where

- $m$  is the number of polynomials occurring in  $\phi$ ,
- $d$  is the maximum degree of any such polynomial in any variable,
- $n$  is the number of free and bound variables in  $\phi$ ,
- $\lambda$  is the maximum length of any integer coefficient of any polynomial, and
- $c$  is some constant.

A first order formula in prenex normal form can be expressed as

$$(Q_w X^w) \dots (Q_1 X^1) F(P_1, \dots, P_m)$$

where

- $Q_i \in \{\forall, \exists\}$ ,  $Q_i \neq Q_{i+1}$ ,
- $P_1, \dots, P_m$  are polynomials in  $n = k + l$  variables  $x_1, \dots, x_k, y_1, \dots, y_l$ , the degrees of the polynomials are bound by  $d$  and the coefficients are integers,
- $X^i$  is a block of  $k_i$  variables such that  $\sum_{1 \leq i \leq w} k_i = k$ ,
- $F(P_1, \dots, P_m)$  is a quantifier-free Boolean formula with atomic predicates of the form,

$$P_i(Y, X^w, \dots, X^1) \bowtie 0, \quad 1 \leq i \leq m,$$

where  $\bowtie \in \{>, <, =\}$  and  $Y = (y_1, \dots, y_l)$  is a block of free variables.

Basu, Pollack and Roy [BPR96] gave a more efficient algorithm in 1996. The time complexity of their algorithm is

$$O(m^{(l+1)\Pi_i(k_i+1)} d^{(l+1)\Pi_i c k_i}),$$

where  $c$  is some constant.

For the existential fragment of the first order theory over reals with no free variables ( $l = 0$ ), the formulas are of the form

$$\exists(x_1, \dots, x_r) F(P_1, \dots, P_m).$$

For such cases, the time complexity is

$$O((md)^{cn}).$$

It is obvious from the expressions of time complexity that the algorithm of Basu et al. is more efficient than Collins' algorithm.

## 7.2 Implementation based on Mathematica

A variant of Collins' algorithm [CH91] is implemented in the tool Mathematica which we use to decide the truth of our formulas. We have implemented the reduction in the form of a Java program that takes as input the probability matrix  $\pi$  and produces as output the simplified formula in a format that can be fed to Mathematica. The Java code is available at <http://www.cse.yorku.ca/~franck/research/pm2m/> and the description of APIs can be found at <http://www.cse.yorku.ca/~franck/research/pm2m/doc>.

**Example 7.2.1.** Consider the probabilistic transition system of Example 2.2.11. The optimized formula for this system is given below in the notation of Mathematica. This formula corresponds to the formula given in Section 5.4.4.

```

1 Reduce[
2   Exists[d12,
3     (0 <= d12 <= 1) &&&
4     (0.11112 <= d12 + 0.27778) &&&
5     (d12 <= 0.38889) &&&
6     Exists[{u12, u13, u32, u42, u43, u33},
7       (0 <= u12 <= 1) &&&
8       (0 <= u13 <= 1) &&&
9       (0 <= u32 <= 1) &&&
10      (0 <= u42 <= 1) &&&
11      (0 <= u43 <= 1) &&&

```

```

12      (u12 + u32 + u42 == 0.4) EE
13      (u13 + u43 + u33 == 0.6) EE
14      (u12 + u13 == 0.7) EE
15      (u32 + u33 == 0.1) EE
16      (u42 + u43 == 0.2) EE
17      (d12 * u12 + 0.11112 * u13 + 0.27778 * u32 + u42 + u43 <= d12)] EE
18      Exists[{u21, u23, u24, u31, u33, u34},
19      (0 <= u21 <= 1) EE
20      (0 <= u23 <= 1) EE
21      (0 <= u24 <= 1) EE
22      (0 <= u31 <= 1) EE
23      (0 <= u34 <= 1) EE
24      (u21 + u31 == 0.7) EE
25      (u23 + u33 == 0.1) EE
26      (u24 + u34 == 0.2) EE
27      (u21 + u23 + u24 == 0.4) EE
28      (u31 + u33 + u34 == 0.6) EE
29      (d12 * u21 + 0.27778 * u23 + u24 + 0.11112 * u31 + u34 <= d12)] EE
30      (0 <= d12 <= 0.5)]]

```

The variables starting with the letter  $u$  correspond to the variables starting with  $\mu$  in Example 5.4.4. The formula was reduced to true by Mathematica in 8.2 seconds on a 3GHz machine with 1GB RAM. Using our algorithm outlined in Section 5.3, we get the value for  $d_{12}$  as 0.31945 with an accuracy of 0.00001. For the formula given in Appendix B, Mathematica runs out of memory.

We also attempted to solve our formulas with a solver called QEPCAD B [Bro03] but the performance of Mathematica was better.

Using Mathematica we could not solve large formulas but, since our formulas belong to the existential fragment of the theory, an implementation of the algorithm given by Basu et al. may produce results for larger formulas.



## 8 Conclusion

Probabilistic transition systems and discrete-time Markov chains are routinely used in the modelling and analysis of probabilistic systems. Probabilistic bisimulation is often used to determine behavioural equivalence between two systems. It has been observed that quantitative analysis is more robust and hence more useful. A number of researchers work in the area of quantitative analysis using pseudometrics.

Our work is based on the behavioural pseudometrics defined by Desharnais et al. There are polynomial time algorithms to approximate distances in the discounted setting but the existence of an algorithm to compute/approximate distances in the undiscounted setting was unknown since 1999. We believe ours is the first algorithm that solves this problem and the techniques used by us can benefit a wide community of researchers in the area of modelling and verification of probabilistic systems.

Our main contribution is proving that the problem of approximating  $d_1$  is decidable. Our work combines a number of ingredients known already for a long time, including the Kantorovich-Rubinstein duality theorem of the fifties, Tarski's fixpoint theorem of the forties and Tarski's decision procedure for the first order theory over reals of the thirties.

Tarski's fixpoint theorem helped us to provide a nice characterization of the distance function as the greatest fixpoint of an order-preserving function. We regard the theorem that the set of all fixpoints of an order-preserving function over a complete-lattice is a complete lattice as one of the two key results used in this thesis. It helps us to prove the existence of a greatest fixpoint.

The other fundamental result used by us is that of the decidability of the first order theory over reals. The first order theory over reals has great expressive power and we could easily express our constraints in a formula of this theory. Based on the decidability of this theory we also proved the decidability of related problems of computing distances.

Kantorovich-Rubenstein's duality theorem allowed us to express our formula in the existential fragment of the first order theory over reals. There are efficient decision procedures for the existential fragment.

We now briefly summarize our main results and contributions.

- We defined an order-preserving distance function  $\Delta$  based on the definition of  $d_1$  given by Desharnais et al. and proved that  $d_1$  is the greatest fixpoint of  $\Delta$ .
- We proved the decidability of the problem of approximating distances in the undiscounted setting by reducing it to the problem of checking satisfiability of a formula in the first order theory over reals, which is known to be decidable.
- We gave the first algorithm to approximate the behavioural pseudometric  $d_1$  up to an arbitrary accuracy. We wrote a Java program to implement the algorithm using Mathematica.
- We proved that the closure ordinal for  $\Delta$  is  $\omega$  which means one will reach the greatest fixpoint within  $\omega$  iterations.

- We proved that the distance between a non-terminating state and any other state is the same as the probability of termination of the other state. We believe it to be a significant result as it can be used to compute exact distances between a number of states in polynomial time. These computations help in reducing the number of variables in the formula thereby simplifying it considerably.
- We demonstrated the general applicability of our techniques by reducing several other distance calculation problems to the problem of deciding a formula of the first order theory over reals.

The techniques exploited in this thesis can also be used to approximate other behavioural pseudometrics that do not discount the future like, for example, the one presented in [DCPP06]. Since the closure ordinal of  $\Delta$  is  $\omega$ , as proved in Proposition 4.3.8, a more efficient iterative algorithm might be feasible. Future work could be in the direction of computing the number of iterations required to approximate the greatest fixpoint. Also, we saw in Section 5.4.1 that the function  $\tau_\omega$  can be used to directly compute the distances between states when one of the states is a non-terminating state. Another direction of future work is to come up with a function for the remaining case, i.e., when both the states are terminating states.

Since the current solvers are based on the doubly-exponential algorithm of Collins and Hong, it is not surprising that our algorithm can only handle small examples as we have shown in Section 7.2. The decision procedure for the existential fragment, however, has exponential time-complexity and with future implementations based on the new algorithm of Basu et al. we hope to deal with larger systems. Also, since the maximum degree of the polynomials in our formulas is two, it might be possible that more efficient decision procedures could be developed for this case.

## Bibliography

- [AFS04] L. de Alfaro, M. Faella, and M. Stoelinga. Linear and branching metrics for quantitative transition systems. In J. Diaz, J. Karhumäki, A. Lepistö, and D. Sannella, editors, *Proceedings of the 31st International Colloquium on Automata, Languages and Programming*, volume 3142 of *Lecture Notes in Computer Science*, pages 97–109, Turku, Finland, July 2004. Springer-Verlag.
- [AHM03] L. de Alfaro, T.A. Henzinger, and R. Majumdar. Discounting the future in systems theory. In J.C.M. Baeten, J.K. Lenstra, J. Parrow, and G.J. Woeginger, editors, *Proceedings of 30th International Colloquium on Automata, Languages and Programming*, volume 2719 of *Lecture Notes in Computer Science*, pages 1022–1037, Eindhoven, June/July 2003. Springer-Verlag.
- [ATHP00] R. Alur, G. Lafferriere T.A. Henzinger, and G.J. Pappas. Discrete abstractions of hybrid systems. In *Proceedings of the IEEE*, volume 88, pages 971–984. IEEE, July 2000.
- [BBKM84] J.W. de Bakker, J.A. Bergstra, J.W. Klop, and J.C. Meyer. Linear time and branching time semantics for recursion with merge. *Theoretical Computer Science*, 34:135–156, 1984.
- [BDM01] M. Bicego, A. Dovier, and V. Murino. Designing the minimal structure of hidden Markov model by bisimulation. In M. Figueiredo, J. Zerubia, and A.K. Jain, editors, *Proceedings of the 3rd International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, volume 2134 of *Lecture Notes in Computer Science*, pages 75–90, Sophia Antipolis, September 2001. Springer-Verlag.
- [BH97] C. Baier and H. Hermanns. Weak bisimulation for fully probabilistic processes. In *Proceedings of the 9th International Conference on Computer Aided Verification*, volume 1254 of *Lecture Notes In Computer Science*, pages 119–130, London, UK, June 1997. Springer-Verlag.
- [BHMW] F. van Breugel, C. Hermida, M. Makkai, and J. Worrell. Recursively defined metric spaces without contraction. *Theoretical Computer Science*. To appear.
- [Bil95] P. Billingsley. *Probability and Measure*. John Wiley and Sons, Inc., New York, third edition, 1995.
- [BPR96] S. Basu, R. Pollack, and M.-F. Roy. On the combinatorial and algebraic complexity of quantifier elimination. *Journal of the ACM*, 43(6):1002–1045, November 1996.
- [BPS01] J.A. Bergstra, A. Ponse, and S.A. Smolka. *Handbook of Process Algebra*. Elsevier Science Inc., New York, 2001.

- [Bre05] F. van Breugel. A behavioural pseudometric for metric labelled transition systems. In M. Abadi and L. de Alfaro, editors, *Proceedings of the 16th International Conference on Concurrency Theory*, volume 3653 of *Lecture Notes in Computer Science*, pages 141–155, San Francisco, August 2005. Springer-Verlag.
- [Bro98] M. Broucke. Regularity of solutions and homotopic equivalence for hybrid systems. In *Proceedings of the 37th IEEE Conference on Decision and Control*, volume 4, pages 4283–4288, Tampa, December 1998. IEEE.
- [Bro03] C.W. Brown. An overview of QEPCAD B: a tool for real quantifier elimination and formula simplification. *Journal of Japan Society for Symbolic and Algebraic Computation*, 10(1):13–22, 2003.
- [BW01a] F. van Breugel and J. Worrell. An algorithm for quantitative verification of probabilistic transition systems. In K.G. Larsen and M. Nielsen, editors, *Proceedings of 12th International Conference on Concurrency Theory*, volume 2154 of *Lecture Notes in Computer Science*, pages 336–350, Aalborg, August 2001. Springer-Verlag.
- [BW01b] F. van Breugel and J. Worrell. Towards quantitative verification of probabilistic transition systems. In F. Orejas, P.G. Spirakis, and J. van Leeuwen, editors, *Proceedings of 28th International Colloquium on Automata, Languages and Programming*, volume 2076 of *Lecture Notes in Computer Science*, pages 421–432, Crete, July 2001. Springer-Verlag.
- [BW05] F. van Breugel and J. Worrell. A behavioural pseudometric for probabilistic transition systems. *Theoretical Computer Science*, 331(1):115–142, February 2005.
- [BW06] F. van Breugel and J. Worrell. Approximating and computing behavioural distances in probabilistic transition systems. *Theoretical Computer Science*, 360(1/3):373–385, August 2006.
- [CB02] P. Caspi and A. Benveniste. Toward an approximation theory for computerised control. In A.L. Sangiovanni-Vincentelli and J. Sifakis, editors, *Proceedings of the 2nd International Conference on Embedded Software*, volume 2491 of *Lecture Notes in Computer Science*, pages 294–304, Grenoble, October 2002. Springer-Verlag.
- [CH91] G.E. Collins and H. Hong. Partial cylindrical algebraic decomposition for quantifier elimination. *Journal of Symbolic Computation*, 12(3):299–328, September 1991.
- [Col75] G.E. Collins. Quantifier elimination for real closed fields by cylindrical algebraic decomposition. In H. Barkhage, editor, *Proceedings of the 2nd GI Conference on Automata Theory and Formal Languages*, volume 33 of *Lecture Notes in Computer Science*, pages 134–183, Kaiserslautern, May 1975. Springer-Verlag.
- [CS96] R. Cleaveland and S. Sims. The NCSU concurrency workbench. In R. Alur and T. Henzinger, editors, *Proceedings of the 8th International Conference on Computer Aided Verification*, volume 1102 of *Lecture Notes in Computer Science*, pages 394–397, New Brunswick, NJ, July 1996. Springer-Verlag.
- [DCPP06] Y. Deng, T. Chothia, C. Palamidessi, and J. Pang. Metrics for action-labelled quantitative transition systems. In *Proceedings of 3rd Workshop on Quantitative Aspects of Programming Languages*, volume 153 of *Electronic Notes in Theoretical Computer Science*, pages 79–96, Edinburgh, April 2006. Elsevier.

- [DEP02] J. Desharnais, A. Edalat, and P. Panangaden. Bisimulation for labelled Markov processes. *Information and Computation*, 179(2):163–193, December 2002.
- [DGJP99] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Metrics for labeled Markov systems. In J.C.M. Baeten and S. Mauw, editors, *Proceedings of 10th International Conference on Concurrency Theory*, volume 1664 of *Lecture Notes in Computer Science*, pages 258–273, Eindhoven, August 1999. Springer-Verlag.
- [DGJP02] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. The metric analogue of weak bisimulation for probabilistic processes. In *Proceedings of 17th Annual IEEE Symposium on Logic in Computer Science*, pages 413–422, Copenhagen, July 2002. IEEE.
- [DGJP04] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Metrics for labelled Markov processes. *Theoretical Computer Science*, 318(3):323–354, June 2004.
- [DP90] B.A. Davey and H.A. Priestley. *Introduction to Lattices and Order*. Cambridge University Press, Cambridge, 1990.
- [DPHW03] A. Di Pierro, C. Hankin, and H. Wiklicky. Quantitative relations and approximate process equivalences. In R. Amadio and D. Lugiez, editors, *Proceedings of the 14th International Conference on Concurrency Theory*, volume 2761 of *Lecture Notes in Computer Science*, pages 508–522, Marseille, September 2003. Springer-Verlag.
- [Eil74] S. Eilenberg. *Automata, Languages, and Machines*. Academic Press, Orlando, USA, 1974.
- [FPP04] N. Ferns, P. Panangaden, and D. Precup. Metrics for finite Markov decision processes. In *Proceedings of the 20th Annual Conference on Uncertainty in Artificial Intelligence*, pages 162–169, Banff, Canada, July 2004. AUAI Press.
- [FPP05] N. Ferns, P. Panangaden, and D. Precup. Metrics for Markov decision processes with infinite state spaces. In *Proceedings of the 21st Annual Conference on Uncertainty in Artificial Intelligence*, page 201, Edinburgh, UK, July 2005. AUAI Press.
- [GDG03] R. Givan, T. Dean, and M. Greig. Equivalence notions and model minimization in Markov decision processes. *Artificial Intelligence*, 147(1-2):163–223, July 2003.
- [GH02] H. Garavel and H. Hermanns. On combining functional verification and performance evaluation using CADP. In *Proceedings of the International Symposium of Formal Methods Europe on Formal Methods - Getting IT Right*, volume 2391 of *Lecture Notes in Computer Science*, pages 410–429, London, UK, July 2002. Springer-Verlag.
- [GJP04] Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Approximate reasoning for real-time probabilistic processes. In *Proceedings of the First International Conference on The Quantitative Evaluation of Systems*, pages 304–313, Washington, USA, September 2004. IEEE Computer Society.
- [GJP06] A. Girard, A.A. Julius, and G.J. Pappas. Approximate simulation relations for hybrid systems. In *Proceedings of 2nd IFAC Conference on Analysis and Design of Hybrid Systems*, Alghero, June 2006. To appear.

- [GJS90] A. Giacalone, C.-C. Jou, and S.A. Smolka. Algebraic reasoning for probabilistic concurrent systems. In *Proceedings of the IFIP WG 2.2/2.3 Working Conference on Programming Concepts and Methods*, pages 443–458, Sea of Gallilee, April 1990. North-Holland.
- [Gla01] R.J. van Glabbeek. The linear time – branching time spectrum I: The semantics of concrete, sequential processes. In J.A. Bergstra, A. Ponse, and S.A. Smolka, editors, *Handbook of Process Algebra*, chapter 1, pages 3–99. Elsevier, 2001.
- [GP] A. Girard and G. J. Pappas. Approximation metrics for discrete and continuous systems. *IEEE Transactions on Automatic Control*. To appear.
- [GP05a] A. Girard and G. J. Pappas. Approximate bisimulations for nonlinear dynamical systems. In *Proceedings of the 44th IEEE Conference on Decision and Control and the European Control Conference*, pages 684–689, Seville, December 2005. IEEE.
- [GP05b] A. Girard and G.J. Pappas. Approximate bisimulations for constrained linear systems. In *Proceedings of the 44th IEEE Conference on Decision and Control and the European Control Conference*, pages 4700–4705, Seville, December 2005. IEEE.
- [GP06] A. Girard and G.J. Pappas. Hierarchical control using approximate simulation relations. In *Proceedings of the IEEE Conference on Decision and Control*, San Diego, December 2006. To appear.
- [GS97] C.M. Grinstead and J.L. Snell. *Introduction to Probability*. AMS, Providence, second revised edition, 1997.
- [GSS95] R.J. van Glabbeek, S.A. Smolka, and B. Steffen. Reactive, generative, and stratified models of probabilistic processes. *Information and Computation*, 121(1):59–80, August 1995.
- [HM85] M. Hennessy and R. Milner. Algebraic laws for nondeterminism and concurrency. *Journal of the ACM*, 32(1):137–161, January 1985.
- [Hoa80] C.A.R. Hoare. A model for communicating sequential processes. In R.M. McKeag and A.M. MacNaughten, editors, *On the Construction of Programs*, pages 229–254. Cambridge University Press, 1980.
- [Hör05] L. Hörmander. *The Analysis of Linear Partial Differential Operators II: Differential Operators with Constant Coefficients*. Classics in Mathematics. Springer-Verlag, Berlin, 2005.
- [JP06] A.A. Julius and G.J. Pappas. Approximate equivalence and approximate synchronization of metric transition systems. In *Proceedings of the IEEE Conference on Decision and Control*, San Diego, December 2006. To appear.
- [Kan42] L. Kantorovich. On the transfer of masses (in Russian). *Doklady Akademii Nauk*, 37(2):227–229, 1942. Translated in *Management Science*, 5:1–4, 1959.
- [KR58] L.V. Kantorovich and G.S. Rubinstein. On the space of completely additive functions. *Vestnik Leningradskogo Universiteta*, 3(2):52–59, 1958. In Russian.

- [LS91] K.G. Larsen and A. Skou. Bisimulation through probabilistic testing. *Information and Computation*, 94(1):1–28, September 1991.
- [Mil80] R. Milner. *A Calculus of Communicating Systems*. Lecture Notes in Computer Science Vol. 92. Springer-Verlag, 1980.
- [Mil83] R. Milner. Calculi for synchrony and asynchrony. *Theoretical Computer Science*, 25:267–310, 1983.
- [Mil89] R. Milner. *Communication and concurrency*. Prentice-Hall Inc., NJ, USA, 1989.
- [ML] S. Mitra and N. Lynch. Approximate simulations for task-structured probabilistic i/o automata. Unpublished.
- [Moh03] M. Mohri. Edit-distance of weighted automata: General definitions and algorithms. *International Journal of Foundations of Computer Science*, 14(6):957–982, 2003.
- [MRST06] J.C Mitchell, A. Ramanathan, A. Scedrov, and V. Teague. A probabilistic polynomial-time process calculus for the analysis of cryptographic protocols. *Theoretical Computer Science*, 353(1/3):118–164, March 2006.
- [NSC<sup>+</sup>] S. Nejati, M. Sabetzadeh, M. Chechik, S. Easterbrook, and P. Zave. Matching and merging of statecharts specification. Unpublished.
- [Par81] D. Park. Concurrency and automata on infinite sequences. In P. Deussen, editor, *Proceedings of 5th GI-Conference on Theoretical Computer Science*, volume 104 of *Lecture Notes in Computer Science*, pages 167–183, Karlsruhe, March 1981. Springer-Verlag.
- [PLS00] A. Philippou, I. Lee, and O. Sokolsky. Weak bisimulation for probabilistic systems. In *Proceedings of the 11th International Conference on Concurrency Theory*, volume 1877 of *Lecture Notes In Computer Science*, pages 334–349, University Park, USA, August 2000. Springer-Verlag.
- [Pnu85] A. Pnueli. Linear and branching structures in the semantics and logics of reactive systems. In W. Brauer, editor, *Proceedings of the 12th Colloquium on Automata, Languages and Programming*, volume 194 of *Lecture Notes In Computer Science*, pages 15–32, Nafplion, Greece, July 1985. Springer-Verlag.
- [PT87] R. Paige and R.E. Tarjan. Three partition refinement algorithms. *SIAM Journal on Computing*, 16(6):973–989, December 1987.
- [RKNP04] J.J.M.M. Rutten, M. Kwiatkowska, G. Norman, and D. Parker. *Mathematical Techniques for Analyzing Concurrent and Probabilistic Systems*, volume 23 of *CRM Monograph Series*. American Mathematical Society, 2004.
- [Rut00] J.J.J.M. Rutten. Universal coalgebra: a theory of systems. *Theoretical Computer Science*, 249(1):3–80, October 2000.
- [Seg95] R. Segala. *Modeling and verification of randomized distributed real-time systems*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 1995. Also appears as technical report MIT/LCS/TR-676.

- [SKL06] O. Sokolsky, S. Kannan, and I. Lee. Simulation-based graph similarity. In H. Hermanns and J. Palsberg, editors, *Proceedings of the 12th. International Conference on Tools and Algorithms for the Construction and Analysis of Systems*, volume 3920 of *Lecture Notes in Computer Science*, pages 426–440, Vienna, Austria, March 2006. Springer-Verlag.
- [Str86] G. Strang. *Introduction to Applied Mathematics*. Wellesley-Cambridge Press, Wellesley, 1986.
- [TAJP06] P. Tabuada, A. Ames, A.A. Julius, and G.J. Pappas. Approximate reduction of dynamical systems. In *Proceedings of the 45th IEEE Conference on Decision and Control*, San Diego, December 2006. IEEE. To appear.
- [Tar51] A. Tarski. *A decision method for elementary algebra and geometry*. University of California Press, Berkeley, 1951.
- [Tar55] A. Tarski. A lattice-theoretic fixed point theorem and its applications. *Pacific Journal of Mathematics*, 5(2):285–309, June 1955.
- [TR98] D. Turi and J.J.M.M. Rutten. On the foundations of final coalgebra semantics: non-well-founded sets, partial orders, metric spaces. *Mathematical Structures in Computer Science*, 8(5):481–540, October 1998.
- [Whi80] W. Whitt. Continuity of generalized semi Markov processes. *Mathematics of Operation Research*, 5(4):494–501, 1980.
- [Yin02] M. Ying. Bisimulation indexes and their applications. *Theoretical Computer Science*, 275(1/2):1–68, March 2002.



## A Calculation of Termination Probabilities

In this appendix we present the key ingredients that are used to compute the termination probabilities. We already defined the termination probability  $\tau_\omega$  in Definition 5.4.1. It is the probability of reaching a state with no outgoing transitions.

We recall the classification of states given in Section 5.4.1.

1.  $S^1 = \{s \in S \mid s \not\rightarrow\}$ : These are states that have no outgoing transitions. We call such states *terminal* states. For such states we know that  $\tau_\omega(s) = 1$ .
2.  $S^0 = \{s \in S \mid \tau_\omega(s) = 0\}$ : These are states that are not connected to any terminal state. We call such states *non-terminating* states.
3.  $S^* = \{s \in S \mid s \rightarrow \wedge \tau_\omega(s) > 0\}$ : Such states have at least one path to a terminal state. We call such states *terminating* states.

We know that  $S^0$ ,  $S^1$  and  $S^*$  form a partition of  $S$ .

For a state  $s \in S^*$ , the probability of termination  $\tau_\omega(s)$  can be expressed as follows:

$$\tau_\omega(s) = \sum_{u \in S^*} \pi(s, s') \tau_\omega(s') + \sum_{s' \in S^1} \pi(s, s').$$

Let us rename the states such that  $S^* = \{s_1, \dots, s_M\}$  for some  $M \geq 0$ . Then the above equation can be expressed in matrix form as

$$T = P.T + R,$$

where  $T[i] = \tau_\omega(s_i)$ ,  $P[i, j] = \pi(s_i, s_j)$  and  $R[i] = \sum_{s' \in S^1} \pi(s_i, s')$  for  $1 \leq i, j \leq M$ . We have that

$$\begin{aligned} T &= P.T + R \\ \Leftrightarrow (I - P).T &= R \\ \Leftrightarrow T &= (I - P)^{-1}.R \end{aligned}$$

Next, we prove that  $(I - P)^{-1}$  exists and, therefore, the terminating probabilities can easily be computed using the above characterization.

**Proposition A.0.2.**  $\lim_{n \rightarrow \infty} P^n = 0$ .

*Proof.* This proof is a modification of the proof of [GS97, Theorem 11.3].

For each state  $s_i \in S^*$ , there exists a path from  $s_i$  to a state in  $S^0 \cup S^1$ . Let  $m_i$  be the minimum length of such a path and let  $p_i$  be the probability of staying in  $S^*$  in the first  $m_i$  transitions when started in  $s_i$ . Clearly,  $p_i < 1$ . Let  $m = \max_{1 \leq i \leq M} m_i$  and  $p = \max_{1 \leq i \leq M} p_i$ . Then  $p < 1$ . Obviously, the probability of staying in  $S^*$  in the first  $m$  transitions when started in  $s_i$  is at most

$p$ . Hence, the probability of staying in  $S^*$  in the first  $km$  transitions when started in  $s_i$  is at most  $p^k$ .

Note that  $P^m[i, j]$  is the probability of reaching state  $s_j$ , starting from state  $s_i$ , in  $m$  transitions. To reach  $s_j \in S^*$ , one has to stay in  $S^*$ . From the above we can conclude that  $P^{km}[i, j] \leq p^k$ , which implies the desired result.  $\square$

**Proposition A.0.3.**  $(I - P)^{-1}$  exists.

*Proof.* This proof is a modification of the proof of [GS97, Theorem 11.4].

The matrix  $I - P$  has an inverse iff 0 is not an eigenvalue of  $I - P$  (see, for example, [Str86, page 47]). We prove the latter by contradiction. Assume that 0 is an eigenvalue of  $I - P$ . Then  $(I - P).X = 0$  for some  $X \neq 0$ . Hence,

$$\begin{aligned}(I - P).X &= 0 \\ \Rightarrow I.X - P.X &= 0 \\ \Rightarrow I.X &= P.X \\ \Rightarrow X &= P.X \\ \Rightarrow X &= P^n.X.\end{aligned}$$

From Proposition A.0.2 we can conclude that  $X$  has to be 0, which contradicts the assumption that  $X \neq 0$ .  $\square$



$$\begin{aligned}
& \mu_{11} + \mu_{12} + \mu_{13} + \mu_{14} + \mu_{15} = 0.0 \wedge \\
& \mu_{21} + \mu_{22} + \mu_{23} + \mu_{24} + \mu_{25} = 0.4 \wedge \\
& \mu_{31} + \mu_{32} + \mu_{33} + \mu_{34} + \mu_{35} = 0.6 \wedge \\
& \mu_{41} + \mu_{42} + \mu_{43} + \mu_{44} + \mu_{45} = 0.0 \wedge \\
& \mu_{51} + \mu_{52} + \mu_{53} + \mu_{54} + \mu_{55} = 0.0 \wedge \\
& (d_{11}\mu_{11} + d_{12}\mu_{12} + d_{13}\mu_{13} + d_{14}\mu_{14} + d_{15}\mu_{15} + \\
& d_{21}\mu_{21} + d_{22}\mu_{22} + d_{23}\mu_{23} + d_{24}\mu_{24} + d_{25}\mu_{25} + \\
& d_{31}\mu_{31} + d_{32}\mu_{32} + d_{33}\mu_{33} + d_{34}\mu_{34} + d_{35}\mu_{35} + \\
& d_{41}\mu_{41} + d_{42}\mu_{42} + d_{43}\mu_{43} + d_{44}\mu_{44} + d_{45}\mu_{45} + \\
& d_{51}\mu_{51} + d_{52}\mu_{52} + d_{53}\mu_{53} + d_{54}\mu_{54} + d_{55}\mu_{55}) \leq d_{11} \\
& ) \\
& \vee \\
& ( \\
& 0.0 + 0.4 + 0.6 + 0.0 + 0.0 = 0 \wedge \\
& 0.0 + 0.4 + 0.6 + 0.0 + 0.0 = 0 \wedge \\
& 0 \geq d_{11} \\
& ) \\
& \vee \\
& ((( \\
& 0.0 + 0.4 + 0.6 + 0.0 + 0.0 > 0 \wedge \\
& 0.0 + 0.4 + 0.6 + 0.0 + 0.0 = 0 \\
& ) \vee ( \\
& 0.0 + 0.4 + 0.6 + 0.0 + 0.0 = 0 \wedge \\
& 0.0 + 0.4 + 0.6 + 0.0 + 0.0 > 0 \wedge \\
& )) \wedge \\
& 1 \leq d_{11} ) \\
& ) \\
& ( \\
& 0.0 + 0.4 + 0.6 + 0.0 + 0.0 > 0 \wedge \\
& 0.7 + 0.0 + 0.0 + 0.2 + 0.1 > 0 \wedge \\
& \exists (\mu_{11}\mu_{12}\mu_{13}\mu_{14}\mu_{15}\mu_{21}\mu_{22}\mu_{23}\mu_{24}\mu_{25}\mu_{31}\mu_{32}\mu_{33}\mu_{34}\mu_{35}\mu_{41}\mu_{42}\mu_{43}\mu_{44}\mu_{45}\mu_{51}\mu_{52}\mu_{53}\mu_{54}\mu_{55}) \{ \\
& \quad \mu_{11} \geq 0 \wedge \mu_{11} \leq 1 \wedge \mu_{12} \geq 0 \wedge \mu_{12} \leq 1 \wedge \mu_{13} \geq 0 \wedge \mu_{13} \leq 1 \wedge \mu_{14} \geq 0 \wedge \mu_{14} \leq 1 \wedge \mu_{15} \geq 0 \wedge \mu_{15} \leq 1 \wedge \\
& \quad \mu_{21} \geq 0 \wedge \mu_{21} \leq 1 \wedge \mu_{22} \geq 0 \wedge \mu_{22} \leq 1 \wedge \mu_{23} \geq 0 \wedge \mu_{23} \leq 1 \wedge \mu_{24} \geq 0 \wedge \mu_{24} \leq 1 \wedge \mu_{25} \geq 0 \wedge \mu_{25} \leq 1 \wedge \\
& \quad \mu_{31} \geq 0 \wedge \mu_{31} \leq 1 \wedge \mu_{32} \geq 0 \wedge \mu_{32} \leq 1 \wedge \mu_{33} \geq 0 \wedge \mu_{33} \leq 1 \wedge \mu_{34} \geq 0 \wedge \mu_{34} \leq 1 \wedge \mu_{35} \geq 0 \wedge \mu_{35} \leq 1 \wedge \\
& \quad \mu_{41} \geq 0 \wedge \mu_{41} \leq 1 \wedge \mu_{42} \geq 0 \wedge \mu_{42} \leq 1 \wedge \mu_{43} \geq 0 \wedge \mu_{43} \leq 1 \wedge \mu_{44} \geq 0 \wedge \mu_{44} \leq 1 \wedge \mu_{45} \geq 0 \wedge \mu_{45} \leq 1 \wedge \\
& \quad \mu_{51} \geq 0 \wedge \mu_{51} \leq 1 \wedge \mu_{52} \geq 0 \wedge \mu_{52} \leq 1 \wedge \mu_{53} \geq 0 \wedge \mu_{53} \leq 1 \wedge \mu_{54} \geq 0 \wedge \mu_{54} \leq 1 \wedge \mu_{55} \geq 0 \wedge \mu_{55} \leq 1 \wedge \\
& \quad \mu_{11} + \mu_{21} + \mu_{31} + \mu_{41} + \mu_{51} = 0.0 \wedge \\
& \quad \mu_{12} + \mu_{22} + \mu_{32} + \mu_{42} + \mu_{52} = 0.4 \wedge \\
& \quad \mu_{13} + \mu_{23} + \mu_{33} + \mu_{43} + \mu_{53} = 0.6 \wedge \\
& \quad \mu_{14} + \mu_{24} + \mu_{34} + \mu_{44} + \mu_{54} = 0.0 \wedge \\
& \quad \mu_{15} + \mu_{25} + \mu_{35} + \mu_{45} + \mu_{55} = 0.0 \wedge \\
& \quad \mu_{11} + \mu_{12} + \mu_{13} + \mu_{14} + \mu_{15} = 0.7 \wedge \\
& \quad \mu_{21} + \mu_{22} + \mu_{23} + \mu_{24} + \mu_{25} = 0.0 \wedge \\
& \quad \mu_{31} + \mu_{32} + \mu_{33} + \mu_{34} + \mu_{35} = 0.0 \wedge \\
& \quad \mu_{41} + \mu_{42} + \mu_{43} + \mu_{44} + \mu_{45} = 0.2 \wedge \\
& \quad \mu_{51} + \mu_{52} + \mu_{53} + \mu_{54} + \mu_{55} = 0.1 \wedge \\
& \quad (d_{11}\mu_{11} + d_{12}\mu_{12} + d_{13}\mu_{13} + d_{14}\mu_{14} + d_{15}\mu_{15} + \\
& \quad d_{21}\mu_{21} + d_{22}\mu_{22} + d_{23}\mu_{23} + d_{24}\mu_{24} + d_{25}\mu_{25} + \\
& \quad d_{31}\mu_{31} + d_{32}\mu_{32} + d_{33}\mu_{33} + d_{34}\mu_{34} + d_{35}\mu_{35} + \\
& \quad d_{41}\mu_{41} + d_{42}\mu_{42} + d_{43}\mu_{43} + d_{44}\mu_{44} + d_{45}\mu_{45} + \\
& \quad d_{51}\mu_{51} + d_{52}\mu_{52} + d_{53}\mu_{53} + d_{54}\mu_{54} + d_{55}\mu_{55}) \leq d_{12} \\
& \} \\
& ) \\
& \vee \\
& ( \\
& 0.0 + 0.4 + 0.6 + 0.0 + 0.0 = 0 \wedge \\
& 0.7 + 0.0 + 0.0 + 0.2 + 0.1 = 0 \wedge \\
& 0 \geq d_{12} \\
& ) \\
& \vee \\
& ((( \\
& 0.0 + 0.4 + 0.6 + 0.0 + 0.0 > 0 \wedge \\
& 0.7 + 0.0 + 0.0 + 0.2 + 0.1 = 0 \\
& ) \vee ( \\
& 0.0 + 0.4 + 0.6 + 0.0 + 0.0 = 0 \wedge \\
& 0.7 + 0.0 + 0.0 + 0.2 + 0.1 > 0 \wedge \\
& )) \wedge \\
& 1 \leq d_{12} ) \\
& )
\end{aligned}$$



$($   
 $0.0 + 0.4 + 0.6 + 0.0 + 0.0 = 0 \wedge$   
 $0.0 + 0.0 + 0.0 + 0.0 + 0.0 = 0 \wedge$   
 $0 \geq d_{14}$   
 $)$   
 $\vee$   
 $(($   
 $0.0 + 0.4 + 0.6 + 0.0 + 0.0 > 0 \wedge$   
 $0.0 + 0.0 + 0.0 + 0.0 + 0.0 = 0$   
 $) \vee ($   
 $0.0 + 0.4 + 0.6 + 0.0 + 0.0 = 0 \wedge$   
 $0.0 + 0.0 + 0.0 + 0.0 + 0.0 > 0 \wedge$   
 $)) \wedge$   
 $1 \leq d_{14}$   
 $)$   
 $($   
 $0.0 + 0.4 + 0.6 + 0.0 + 0.0 > 0 \wedge$   
 $0.0 + 0.0 + 0.0 + 0.0 + 1.0 > 0 \wedge$   
 $\exists (\mu_{11}\mu_{12}\mu_{13}\mu_{14}\mu_{15}\mu_{21}\mu_{22}\mu_{23}\mu_{24}\mu_{25}\mu_{31}\mu_{32}\mu_{33}\mu_{34}\mu_{35}\mu_{41}\mu_{42}\mu_{43}\mu_{44}\mu_{45}\mu_{51}\mu_{52}\mu_{53}\mu_{54}\mu_{55}) \{$   
 $\mu_{11} \geq 0 \wedge \mu_{11} \leq 1 \wedge \mu_{12} \geq 0 \wedge \mu_{12} \leq 1 \wedge \mu_{13} \geq 0 \wedge \mu_{13} \leq 1 \wedge \mu_{14} \geq 0 \wedge \mu_{14} \leq 1 \wedge \mu_{15} \geq 0 \wedge \mu_{15} \leq 1 \wedge$   
 $\mu_{21} \geq 0 \wedge \mu_{21} \leq 1 \wedge \mu_{22} \geq 0 \wedge \mu_{22} \leq 1 \wedge \mu_{23} \geq 0 \wedge \mu_{23} \leq 1 \wedge \mu_{24} \geq 0 \wedge \mu_{24} \leq 1 \wedge \mu_{25} \geq 0 \wedge \mu_{25} \leq 1 \wedge$   
 $\mu_{31} \geq 0 \wedge \mu_{31} \leq 1 \wedge \mu_{32} \geq 0 \wedge \mu_{32} \leq 1 \wedge \mu_{33} \geq 0 \wedge \mu_{33} \leq 1 \wedge \mu_{34} \geq 0 \wedge \mu_{34} \leq 1 \wedge \mu_{35} \geq 0 \wedge \mu_{35} \leq 1 \wedge$   
 $\mu_{41} \geq 0 \wedge \mu_{41} \leq 1 \wedge \mu_{42} \geq 0 \wedge \mu_{42} \leq 1 \wedge \mu_{43} \geq 0 \wedge \mu_{43} \leq 1 \wedge \mu_{44} \geq 0 \wedge \mu_{44} \leq 1 \wedge \mu_{45} \geq 0 \wedge \mu_{45} \leq 1 \wedge$   
 $\mu_{51} \geq 0 \wedge \mu_{51} \leq 1 \wedge \mu_{52} \geq 0 \wedge \mu_{52} \leq 1 \wedge \mu_{53} \geq 0 \wedge \mu_{53} \leq 1 \wedge \mu_{54} \geq 0 \wedge \mu_{54} \leq 1 \wedge \mu_{55} \geq 0 \wedge \mu_{55} \leq 1 \wedge$   
 $\mu_{11} + \mu_{21} + \mu_{31} + \mu_{41} + \mu_{51} = 0.0 \wedge$   
 $\mu_{12} + \mu_{22} + \mu_{32} + \mu_{42} + \mu_{52} = 0.4 \wedge$   
 $\mu_{13} + \mu_{23} + \mu_{33} + \mu_{43} + \mu_{53} = 0.6 \wedge$   
 $\mu_{14} + \mu_{24} + \mu_{34} + \mu_{44} + \mu_{54} = 0.0 \wedge$   
 $\mu_{15} + \mu_{25} + \mu_{35} + \mu_{45} + \mu_{55} = 0.0 \wedge$   
 $\mu_{11} + \mu_{12} + \mu_{13} + \mu_{14} + \mu_{15} = 0.0 \wedge$   
 $\mu_{21} + \mu_{22} + \mu_{23} + \mu_{24} + \mu_{25} = 0.0 \wedge$   
 $\mu_{31} + \mu_{32} + \mu_{33} + \mu_{34} + \mu_{35} = 0.0 \wedge$   
 $\mu_{41} + \mu_{42} + \mu_{43} + \mu_{44} + \mu_{45} = 0.0 \wedge$   
 $\mu_{51} + \mu_{52} + \mu_{53} + \mu_{54} + \mu_{55} = 1.0 \wedge$   
 $(d_{11}\mu_{11} + d_{12}\mu_{12} + d_{13}\mu_{13} + d_{14}\mu_{14} + d_{15}\mu_{15} +$   
 $d_{21}\mu_{21} + d_{22}\mu_{22} + d_{23}\mu_{23} + d_{24}\mu_{24} + d_{25}\mu_{25} +$   
 $d_{31}\mu_{31} + d_{32}\mu_{32} + d_{33}\mu_{33} + d_{34}\mu_{34} + d_{35}\mu_{35} +$   
 $d_{41}\mu_{41} + d_{42}\mu_{42} + d_{43}\mu_{43} + d_{44}\mu_{44} + d_{45}\mu_{45} +$   
 $d_{51}\mu_{51} + d_{52}\mu_{52} + d_{53}\mu_{53} + d_{54}\mu_{54} + d_{55}\mu_{55}) \leq d_{15}$   
 $\}$   
 $)$   
 $\vee$   
 $($   
 $0.0 + 0.4 + 0.6 + 0.0 + 0.0 = 0 \wedge$   
 $0.0 + 0.0 + 0.0 + 0.0 + 1.0 = 0 \wedge$   
 $0 \geq d_{15}$   
 $)$   
 $\vee$   
 $(($   
 $0.0 + 0.4 + 0.6 + 0.0 + 0.0 > 0 \wedge$   
 $0.0 + 0.0 + 0.0 + 0.0 + 1.0 = 0$   
 $) \vee ($   
 $0.0 + 0.4 + 0.6 + 0.0 + 0.0 = 0 \wedge$   
 $0.0 + 0.0 + 0.0 + 0.0 + 1.0 > 0 \wedge$   
 $)) \wedge$   
 $1 \leq d_{15}$   
 $)$   
 $($   
 $0.7 + 0.0 + 0.0 + 0.2 + 0.1 > 0 \wedge$   
 $0.0 + 0.4 + 0.6 + 0.0 + 0.0 > 0 \wedge$   
 $\exists (\mu_{11}\mu_{12}\mu_{13}\mu_{14}\mu_{15}\mu_{21}\mu_{22}\mu_{23}\mu_{24}\mu_{25}\mu_{31}\mu_{32}\mu_{33}\mu_{34}\mu_{35}\mu_{41}\mu_{42}\mu_{43}\mu_{44}\mu_{45}\mu_{51}\mu_{52}\mu_{53}\mu_{54}\mu_{55}) \{$   
 $\mu_{11} \geq 0 \wedge \mu_{11} \leq 1 \wedge \mu_{12} \geq 0 \wedge \mu_{12} \leq 1 \wedge \mu_{13} \geq 0 \wedge \mu_{13} \leq 1 \wedge \mu_{14} \geq 0 \wedge \mu_{14} \leq 1 \wedge \mu_{15} \geq 0 \wedge \mu_{15} \leq 1 \wedge$   
 $\mu_{21} \geq 0 \wedge \mu_{21} \leq 1 \wedge \mu_{22} \geq 0 \wedge \mu_{22} \leq 1 \wedge \mu_{23} \geq 0 \wedge \mu_{23} \leq 1 \wedge \mu_{24} \geq 0 \wedge \mu_{24} \leq 1 \wedge \mu_{25} \geq 0 \wedge \mu_{25} \leq 1 \wedge$   
 $\mu_{31} \geq 0 \wedge \mu_{31} \leq 1 \wedge \mu_{32} \geq 0 \wedge \mu_{32} \leq 1 \wedge \mu_{33} \geq 0 \wedge \mu_{33} \leq 1 \wedge \mu_{34} \geq 0 \wedge \mu_{34} \leq 1 \wedge \mu_{35} \geq 0 \wedge \mu_{35} \leq 1 \wedge$   
 $\mu_{41} \geq 0 \wedge \mu_{41} \leq 1 \wedge \mu_{42} \geq 0 \wedge \mu_{42} \leq 1 \wedge \mu_{43} \geq 0 \wedge \mu_{43} \leq 1 \wedge \mu_{44} \geq 0 \wedge \mu_{44} \leq 1 \wedge \mu_{45} \geq 0 \wedge \mu_{45} \leq 1 \wedge$   
 $\mu_{51} \geq 0 \wedge \mu_{51} \leq 1 \wedge \mu_{52} \geq 0 \wedge \mu_{52} \leq 1 \wedge \mu_{53} \geq 0 \wedge \mu_{53} \leq 1 \wedge \mu_{54} \geq 0 \wedge \mu_{54} \leq 1 \wedge \mu_{55} \geq 0 \wedge \mu_{55} \leq 1 \wedge$   
 $\mu_{11} + \mu_{21} + \mu_{31} + \mu_{41} + \mu_{51} = 0.7 \wedge$   
 $\mu_{12} + \mu_{22} + \mu_{32} + \mu_{42} + \mu_{52} = 0.0 \wedge$   
 $\mu_{13} + \mu_{23} + \mu_{33} + \mu_{43} + \mu_{53} = 0.0 \wedge$   
 $\}$   
 $)$

$$\begin{aligned}
& \mu_{14} + \mu_{24} + \mu_{34} + \mu_{44} + \mu_{54} = 0.2 \wedge \\
& \mu_{15} + \mu_{25} + \mu_{35} + \mu_{45} + \mu_{55} = 0.1 \wedge \\
& \mu_{11} + \mu_{12} + \mu_{13} + \mu_{14} + \mu_{15} = 0.0 \wedge \\
& \mu_{21} + \mu_{22} + \mu_{23} + \mu_{24} + \mu_{25} = 0.4 \wedge \\
& \mu_{31} + \mu_{32} + \mu_{33} + \mu_{34} + \mu_{35} = 0.6 \wedge \\
& \mu_{41} + \mu_{42} + \mu_{43} + \mu_{44} + \mu_{45} = 0.0 \wedge \\
& \mu_{51} + \mu_{52} + \mu_{53} + \mu_{54} + \mu_{55} = 0.0 \wedge \\
& (d_{11}\mu_{11} + d_{12}\mu_{12} + d_{13}\mu_{13} + d_{14}\mu_{14} + d_{15}\mu_{15} + \\
& d_{21}\mu_{21} + d_{22}\mu_{22} + d_{23}\mu_{23} + d_{24}\mu_{24} + d_{25}\mu_{25} + \\
& d_{31}\mu_{31} + d_{32}\mu_{32} + d_{33}\mu_{33} + d_{34}\mu_{34} + d_{35}\mu_{35} + \\
& d_{41}\mu_{41} + d_{42}\mu_{42} + d_{43}\mu_{43} + d_{44}\mu_{44} + d_{45}\mu_{45} + \\
& d_{51}\mu_{51} + d_{52}\mu_{52} + d_{53}\mu_{53} + d_{54}\mu_{54} + d_{55}\mu_{55}) \leq d_{21} \\
& ) \\
& ) \\
& \vee \\
& ( \\
& 0.7 + 0.0 + 0.0 + 0.2 + 0.1 = 0 \wedge \\
& 0.0 + 0.4 + 0.6 + 0.0 + 0.0 = 0 \wedge \\
& 0 \geq d_{21} \\
& ) \\
& ) \\
& \vee \\
& ((( \\
& 0.7 + 0.0 + 0.0 + 0.2 + 0.1 > 0 \wedge \\
& 0.0 + 0.4 + 0.6 + 0.0 + 0.0 = 0 \\
& ) \vee ( \\
& 0.7 + 0.0 + 0.0 + 0.2 + 0.1 = 0 \wedge \\
& 0.0 + 0.4 + 0.6 + 0.0 + 0.0 > 0 \wedge \\
& )) \wedge \\
& 1 \leq d_{21} ) \\
& ) \\
& ( \\
& ( \\
& 0.7 + 0.0 + 0.0 + 0.2 + 0.1 > 0 \wedge \\
& 0.7 + 0.0 + 0.0 + 0.2 + 0.1 > 0 \wedge \\
& \exists (\mu_{11}\mu_{12}\mu_{13}\mu_{14}\mu_{15}\mu_{21}\mu_{22}\mu_{23}\mu_{24}\mu_{25}\mu_{31}\mu_{32}\mu_{33}\mu_{34}\mu_{35}\mu_{41}\mu_{42}\mu_{43}\mu_{44}\mu_{45}\mu_{51}\mu_{52}\mu_{53}\mu_{54}\mu_{55}) \{ \\
& \mu_{11} \geq 0 \wedge \mu_{11} \leq 1 \wedge \mu_{12} \geq 0 \wedge \mu_{12} \leq 1 \wedge \mu_{13} \geq 0 \wedge \mu_{13} \leq 1 \wedge \mu_{14} \geq 0 \wedge \mu_{14} \leq 1 \wedge \mu_{15} \geq 0 \wedge \mu_{15} \leq 1 \wedge \\
& \mu_{21} \geq 0 \wedge \mu_{21} \leq 1 \wedge \mu_{22} \geq 0 \wedge \mu_{22} \leq 1 \wedge \mu_{23} \geq 0 \wedge \mu_{23} \leq 1 \wedge \mu_{24} \geq 0 \wedge \mu_{24} \leq 1 \wedge \mu_{25} \geq 0 \wedge \mu_{25} \leq 1 \wedge \\
& \mu_{31} \geq 0 \wedge \mu_{31} \leq 1 \wedge \mu_{32} \geq 0 \wedge \mu_{32} \leq 1 \wedge \mu_{33} \geq 0 \wedge \mu_{33} \leq 1 \wedge \mu_{34} \geq 0 \wedge \mu_{34} \leq 1 \wedge \mu_{35} \geq 0 \wedge \mu_{35} \leq 1 \wedge \\
& \mu_{41} \geq 0 \wedge \mu_{41} \leq 1 \wedge \mu_{42} \geq 0 \wedge \mu_{42} \leq 1 \wedge \mu_{43} \geq 0 \wedge \mu_{43} \leq 1 \wedge \mu_{44} \geq 0 \wedge \mu_{44} \leq 1 \wedge \mu_{45} \geq 0 \wedge \mu_{45} \leq 1 \wedge \\
& \mu_{51} \geq 0 \wedge \mu_{51} \leq 1 \wedge \mu_{52} \geq 0 \wedge \mu_{52} \leq 1 \wedge \mu_{53} \geq 0 \wedge \mu_{53} \leq 1 \wedge \mu_{54} \geq 0 \wedge \mu_{54} \leq 1 \wedge \mu_{55} \geq 0 \wedge \mu_{55} \leq 1 \wedge \\
& \mu_{11} + \mu_{21} + \mu_{31} + \mu_{41} + \mu_{51} = 0.7 \wedge \\
& \mu_{12} + \mu_{22} + \mu_{32} + \mu_{42} + \mu_{52} = 0.0 \wedge \\
& \mu_{13} + \mu_{23} + \mu_{33} + \mu_{43} + \mu_{53} = 0.0 \wedge \\
& \mu_{14} + \mu_{24} + \mu_{34} + \mu_{44} + \mu_{54} = 0.2 \wedge \\
& \mu_{15} + \mu_{25} + \mu_{35} + \mu_{45} + \mu_{55} = 0.1 \wedge \\
& \mu_{11} + \mu_{12} + \mu_{13} + \mu_{14} + \mu_{15} = 0.7 \wedge \\
& \mu_{21} + \mu_{22} + \mu_{23} + \mu_{24} + \mu_{25} = 0.0 \wedge \\
& \mu_{31} + \mu_{32} + \mu_{33} + \mu_{34} + \mu_{35} = 0.0 \wedge \\
& \mu_{41} + \mu_{42} + \mu_{43} + \mu_{44} + \mu_{45} = 0.2 \wedge \\
& \mu_{51} + \mu_{52} + \mu_{53} + \mu_{54} + \mu_{55} = 0.1 \wedge \\
& (d_{11}\mu_{11} + d_{12}\mu_{12} + d_{13}\mu_{13} + d_{14}\mu_{14} + d_{15}\mu_{15} + \\
& d_{21}\mu_{21} + d_{22}\mu_{22} + d_{23}\mu_{23} + d_{24}\mu_{24} + d_{25}\mu_{25} + \\
& d_{31}\mu_{31} + d_{32}\mu_{32} + d_{33}\mu_{33} + d_{34}\mu_{34} + d_{35}\mu_{35} + \\
& d_{41}\mu_{41} + d_{42}\mu_{42} + d_{43}\mu_{43} + d_{44}\mu_{44} + d_{45}\mu_{45} + \\
& d_{51}\mu_{51} + d_{52}\mu_{52} + d_{53}\mu_{53} + d_{54}\mu_{54} + d_{55}\mu_{55}) \leq d_{22} \\
& ) \\
& ) \\
& \vee \\
& ( \\
& 0.7 + 0.0 + 0.0 + 0.2 + 0.1 = 0 \wedge \\
& 0.7 + 0.0 + 0.0 + 0.2 + 0.1 = 0 \wedge \\
& 0 \geq d_{22} \\
& ) \\
& ) \\
& \vee \\
& ((( \\
& 0.7 + 0.0 + 0.0 + 0.2 + 0.1 > 0 \wedge \\
& 0.7 + 0.0 + 0.0 + 0.2 + 0.1 = 0 \\
& ) \vee ( \\
& 0.7 + 0.0 + 0.0 + 0.2 + 0.1 = 0 \wedge \\
& 0.7 + 0.0 + 0.0 + 0.2 + 0.1 > 0 \wedge \\
& )) \wedge
\end{aligned}$$





$$\begin{aligned} & ) \\ & \vee \\ & ( \\ & 0.7 + 0.0 + 0.0 + 0.2 + 0.1 = 0 \wedge \\ & 0.0 + 0.0 + 0.0 + 0.0 + 0.0 = 0 \wedge \\ & 0 \geq d_{24} \\ & ) \\ & \vee \\ & ((( \\ & 0.7 + 0.0 + 0.0 + 0.2 + 0.1 > 0 \wedge \\ & 0.0 + 0.0 + 0.0 + 0.0 + 0.0 = 0 \\ & ) \vee ( \\ & 0.7 + 0.0 + 0.0 + 0.2 + 0.1 = 0 \wedge \\ & 0.0 + 0.0 + 0.0 + 0.0 + 0.0 > 0 \wedge \\ & )) \wedge \\ & 1 \leq d_{24} ) \\ & ) \\ & ( \\ & ( \\ & 0.7 + 0.0 + 0.0 + 0.2 + 0.1 > 0 \wedge \\ & 0.0 + 0.0 + 0.0 + 0.0 + 1.0 > 0 \wedge \\ & \exists (\mu_{11} \mu_{12} \mu_{13} \mu_{14} \mu_{15} \mu_{21} \mu_{22} \mu_{23} \mu_{24} \mu_{25} \mu_{31} \mu_{32} \mu_{33} \mu_{34} \mu_{35} \mu_{41} \mu_{42} \mu_{43} \mu_{44} \mu_{45} \mu_{51} \mu_{52} \mu_{53} \mu_{54} \mu_{55}) \{ \\ & \quad \mu_{11} \geq 0 \wedge \mu_{11} \leq 1 \wedge \mu_{12} \geq 0 \wedge \mu_{12} \leq 1 \wedge \mu_{13} \geq 0 \wedge \mu_{13} \leq 1 \wedge \mu_{14} \geq 0 \wedge \mu_{14} \leq 1 \wedge \mu_{15} \geq 0 \wedge \mu_{15} \leq 1 \wedge \\ & \quad \mu_{21} \geq 0 \wedge \mu_{21} \leq 1 \wedge \mu_{22} \geq 0 \wedge \mu_{22} \leq 1 \wedge \mu_{23} \geq 0 \wedge \mu_{23} \leq 1 \wedge \mu_{24} \geq 0 \wedge \mu_{24} \leq 1 \wedge \mu_{25} \geq 0 \wedge \mu_{25} \leq 1 \wedge \\ & \quad \mu_{31} \geq 0 \wedge \mu_{31} \leq 1 \wedge \mu_{32} \geq 0 \wedge \mu_{32} \leq 1 \wedge \mu_{33} \geq 0 \wedge \mu_{33} \leq 1 \wedge \mu_{34} \geq 0 \wedge \mu_{34} \leq 1 \wedge \mu_{35} \geq 0 \wedge \mu_{35} \leq 1 \wedge \\ & \quad \mu_{41} \geq 0 \wedge \mu_{41} \leq 1 \wedge \mu_{42} \geq 0 \wedge \mu_{42} \leq 1 \wedge \mu_{43} \geq 0 \wedge \mu_{43} \leq 1 \wedge \mu_{44} \geq 0 \wedge \mu_{44} \leq 1 \wedge \mu_{45} \geq 0 \wedge \mu_{45} \leq 1 \wedge \\ & \quad \mu_{51} \geq 0 \wedge \mu_{51} \leq 1 \wedge \mu_{52} \geq 0 \wedge \mu_{52} \leq 1 \wedge \mu_{53} \geq 0 \wedge \mu_{53} \leq 1 \wedge \mu_{54} \geq 0 \wedge \mu_{54} \leq 1 \wedge \mu_{55} \geq 0 \wedge \mu_{55} \leq 1 \wedge \\ & \quad \mu_{11} + \mu_{21} + \mu_{31} + \mu_{41} + \mu_{51} = 0.7 \wedge \\ & \quad \mu_{12} + \mu_{22} + \mu_{32} + \mu_{42} + \mu_{52} = 0.0 \wedge \\ & \quad \mu_{13} + \mu_{23} + \mu_{33} + \mu_{43} + \mu_{53} = 0.0 \wedge \\ & \quad \mu_{14} + \mu_{24} + \mu_{34} + \mu_{44} + \mu_{54} = 0.2 \wedge \\ & \quad \mu_{15} + \mu_{25} + \mu_{35} + \mu_{45} + \mu_{55} = 0.1 \wedge \\ & \quad \mu_{11} + \mu_{12} + \mu_{13} + \mu_{14} + \mu_{15} = 0.0 \wedge \\ & \quad \mu_{21} + \mu_{22} + \mu_{23} + \mu_{24} + \mu_{25} = 0.0 \wedge \\ & \quad \mu_{31} + \mu_{32} + \mu_{33} + \mu_{34} + \mu_{35} = 0.0 \wedge \\ & \quad \mu_{41} + \mu_{42} + \mu_{43} + \mu_{44} + \mu_{45} = 0.0 \wedge \\ & \quad \mu_{51} + \mu_{52} + \mu_{53} + \mu_{54} + \mu_{55} = 1.0 \wedge \\ & \quad (d_{11}\mu_{11} + d_{12}\mu_{12} + d_{13}\mu_{13} + d_{14}\mu_{14} + d_{15}\mu_{15} + \\ & \quad d_{21}\mu_{21} + d_{22}\mu_{22} + d_{23}\mu_{23} + d_{24}\mu_{24} + d_{25}\mu_{25} + \\ & \quad d_{31}\mu_{31} + d_{32}\mu_{32} + d_{33}\mu_{33} + d_{34}\mu_{34} + d_{35}\mu_{35} + \\ & \quad d_{41}\mu_{41} + d_{42}\mu_{42} + d_{43}\mu_{43} + d_{44}\mu_{44} + d_{45}\mu_{45} + \\ & \quad d_{51}\mu_{51} + d_{52}\mu_{52} + d_{53}\mu_{53} + d_{54}\mu_{54} + d_{55}\mu_{55}) \leq d_{25} \\ & \} \\ & ) \\ & \vee \\ & ( \\ & 0.7 + 0.0 + 0.0 + 0.2 + 0.1 = 0 \wedge \\ & 0.0 + 0.0 + 0.0 + 0.0 + 1.0 = 0 \wedge \\ & 0 \geq d_{25} \\ & ) \\ & \vee \\ & ((( \\ & 0.7 + 0.0 + 0.0 + 0.2 + 0.1 > 0 \wedge \\ & 0.0 + 0.0 + 0.0 + 0.0 + 1.0 = 0 \\ & ) \vee ( \\ & 0.7 + 0.0 + 0.0 + 0.2 + 0.1 = 0 \wedge \\ & 0.0 + 0.0 + 0.0 + 0.0 + 1.0 > 0 \wedge \\ & )) \wedge \\ & 1 \leq d_{25} ) \\ & ) \\ & ( \\ & ( \\ & 0.0 + 0.0 + 1.0 + 0.0 + 0.0 > 0 \wedge \\ & 0.0 + 0.4 + 0.6 + 0.0 + 0.0 > 0 \wedge \\ & \exists (\mu_{11} \mu_{12} \mu_{13} \mu_{14} \mu_{15} \mu_{21} \mu_{22} \mu_{23} \mu_{24} \mu_{25} \mu_{31} \mu_{32} \mu_{33} \mu_{34} \mu_{35} \mu_{41} \mu_{42} \mu_{43} \mu_{44} \mu_{45} \mu_{51} \mu_{52} \mu_{53} \mu_{54} \mu_{55}) \{ \\ & \quad \mu_{11} \geq 0 \wedge \mu_{11} \leq 1 \wedge \mu_{12} \geq 0 \wedge \mu_{12} \leq 1 \wedge \mu_{13} \geq 0 \wedge \mu_{13} \leq 1 \wedge \mu_{14} \geq 0 \wedge \mu_{14} \leq 1 \wedge \mu_{15} \geq 0 \wedge \mu_{15} \leq 1 \wedge \\ & \quad \mu_{21} \geq 0 \wedge \mu_{21} \leq 1 \wedge \mu_{22} \geq 0 \wedge \mu_{22} \leq 1 \wedge \mu_{23} \geq 0 \wedge \mu_{23} \leq 1 \wedge \mu_{24} \geq 0 \wedge \mu_{24} \leq 1 \wedge \mu_{25} \geq 0 \wedge \mu_{25} \leq 1 \wedge \\ & \quad \mu_{31} \geq 0 \wedge \mu_{31} \leq 1 \wedge \mu_{32} \geq 0 \wedge \mu_{32} \leq 1 \wedge \mu_{33} \geq 0 \wedge \mu_{33} \leq 1 \wedge \mu_{34} \geq 0 \wedge \mu_{34} \leq 1 \wedge \mu_{35} \geq 0 \wedge \mu_{35} \leq 1 \wedge \\ & \quad \mu_{41} \geq 0 \wedge \mu_{41} \leq 1 \wedge \mu_{42} \geq 0 \wedge \mu_{42} \leq 1 \wedge \mu_{43} \geq 0 \wedge \mu_{43} \leq 1 \wedge \mu_{44} \geq 0 \wedge \mu_{44} \leq 1 \wedge \mu_{45} \geq 0 \wedge \mu_{45} \leq 1 \wedge \\ & \quad \mu_{51} \geq 0 \wedge \mu_{51} \leq 1 \wedge \mu_{52} \geq 0 \wedge \mu_{52} \leq 1 \wedge \mu_{53} \geq 0 \wedge \mu_{53} \leq 1 \wedge \mu_{54} \geq 0 \wedge \mu_{54} \leq 1 \wedge \mu_{55} \geq 0 \wedge \mu_{55} \leq 1 \wedge \\ & \quad \mu_{11} + \mu_{21} + \mu_{31} + \mu_{41} + \mu_{51} = 0.0 \wedge \\ & \} \\ & ) \end{aligned}$$

$$\begin{aligned}
& \mu_{12} + \mu_{22} + \mu_{32} + \mu_{42} + \mu_{52} = 0.0 \wedge \\
& \mu_{13} + \mu_{23} + \mu_{33} + \mu_{43} + \mu_{53} = 1.0 \wedge \\
& \mu_{14} + \mu_{24} + \mu_{34} + \mu_{44} + \mu_{54} = 0.0 \wedge \\
& \mu_{15} + \mu_{25} + \mu_{35} + \mu_{45} + \mu_{55} = 0.0 \wedge \\
& \mu_{11} + \mu_{12} + \mu_{13} + \mu_{14} + \mu_{15} = 0.0 \wedge \\
& \mu_{21} + \mu_{22} + \mu_{23} + \mu_{24} + \mu_{25} = 0.4 \wedge \\
& \mu_{31} + \mu_{32} + \mu_{33} + \mu_{34} + \mu_{35} = 0.6 \wedge \\
& \mu_{41} + \mu_{42} + \mu_{43} + \mu_{44} + \mu_{45} = 0.0 \wedge \\
& \mu_{51} + \mu_{52} + \mu_{53} + \mu_{54} + \mu_{55} = 0.0 \wedge \\
& (d_{11}\mu_{11} + d_{12}\mu_{12} + d_{13}\mu_{13} + d_{14}\mu_{14} + d_{15}\mu_{15} + \\
& d_{21}\mu_{21} + d_{22}\mu_{22} + d_{23}\mu_{23} + d_{24}\mu_{24} + d_{25}\mu_{25} + \\
& d_{31}\mu_{31} + d_{32}\mu_{32} + d_{33}\mu_{33} + d_{34}\mu_{34} + d_{35}\mu_{35} + \\
& d_{41}\mu_{41} + d_{42}\mu_{42} + d_{43}\mu_{43} + d_{44}\mu_{44} + d_{45}\mu_{45} + \\
& d_{51}\mu_{51} + d_{52}\mu_{52} + d_{53}\mu_{53} + d_{54}\mu_{54} + d_{55}\mu_{55}) \leq d_{31} \\
& ) \\
& ) \\
& \vee \\
& ( \\
& 0.0 + 0.0 + 1.0 + 0.0 + 0.0 = 0 \wedge \\
& 0.0 + 0.4 + 0.6 + 0.0 + 0.0 = 0 \wedge \\
& 0 \geq d_{31} \\
& ) \\
& \vee \\
& ((( \\
& 0.0 + 0.0 + 1.0 + 0.0 + 0.0 > 0 \wedge \\
& 0.0 + 0.4 + 0.6 + 0.0 + 0.0 = 0 \\
& ) \vee ( \\
& 0.0 + 0.0 + 1.0 + 0.0 + 0.0 = 0 \wedge \\
& 0.0 + 0.4 + 0.6 + 0.0 + 0.0 > 0 \wedge \\
& )) \wedge \\
& 1 \leq d_{31} ) \\
& ) \\
& ( \\
& ( \\
& 0.0 + 0.0 + 1.0 + 0.0 + 0.0 > 0 \wedge \\
& 0.7 + 0.0 + 0.0 + 0.2 + 0.1 > 0 \wedge \\
& \exists (\mu_{11}\mu_{12}\mu_{13}\mu_{14}\mu_{15}\mu_{21}\mu_{22}\mu_{23}\mu_{24}\mu_{25}\mu_{31}\mu_{32}\mu_{33}\mu_{34}\mu_{35}\mu_{41}\mu_{42}\mu_{43}\mu_{44}\mu_{45}\mu_{51}\mu_{52}\mu_{53}\mu_{54}\mu_{55}) \{ \\
& \mu_{11} \geq 0 \wedge \mu_{11} \leq 1 \wedge \mu_{12} \geq 0 \wedge \mu_{12} \leq 1 \wedge \mu_{13} \geq 0 \wedge \mu_{13} \leq 1 \wedge \mu_{14} \geq 0 \wedge \mu_{14} \leq 1 \wedge \mu_{15} \geq 0 \wedge \mu_{15} \leq 1 \wedge \\
& \mu_{21} \geq 0 \wedge \mu_{21} \leq 1 \wedge \mu_{22} \geq 0 \wedge \mu_{22} \leq 1 \wedge \mu_{23} \geq 0 \wedge \mu_{23} \leq 1 \wedge \mu_{24} \geq 0 \wedge \mu_{24} \leq 1 \wedge \mu_{25} \geq 0 \wedge \mu_{25} \leq 1 \wedge \\
& \mu_{31} \geq 0 \wedge \mu_{31} \leq 1 \wedge \mu_{32} \geq 0 \wedge \mu_{32} \leq 1 \wedge \mu_{33} \geq 0 \wedge \mu_{33} \leq 1 \wedge \mu_{34} \geq 0 \wedge \mu_{34} \leq 1 \wedge \mu_{35} \geq 0 \wedge \mu_{35} \leq 1 \wedge \\
& \mu_{41} \geq 0 \wedge \mu_{41} \leq 1 \wedge \mu_{42} \geq 0 \wedge \mu_{42} \leq 1 \wedge \mu_{43} \geq 0 \wedge \mu_{43} \leq 1 \wedge \mu_{44} \geq 0 \wedge \mu_{44} \leq 1 \wedge \mu_{45} \geq 0 \wedge \mu_{45} \leq 1 \wedge \\
& \mu_{51} \geq 0 \wedge \mu_{51} \leq 1 \wedge \mu_{52} \geq 0 \wedge \mu_{52} \leq 1 \wedge \mu_{53} \geq 0 \wedge \mu_{53} \leq 1 \wedge \mu_{54} \geq 0 \wedge \mu_{54} \leq 1 \wedge \mu_{55} \geq 0 \wedge \mu_{55} \leq 1 \wedge \\
& \mu_{11} + \mu_{21} + \mu_{31} + \mu_{41} + \mu_{51} = 0.0 \wedge \\
& \mu_{12} + \mu_{22} + \mu_{32} + \mu_{42} + \mu_{52} = 0.0 \wedge \\
& \mu_{13} + \mu_{23} + \mu_{33} + \mu_{43} + \mu_{53} = 1.0 \wedge \\
& \mu_{14} + \mu_{24} + \mu_{34} + \mu_{44} + \mu_{54} = 0.0 \wedge \\
& \mu_{15} + \mu_{25} + \mu_{35} + \mu_{45} + \mu_{55} = 0.0 \wedge \\
& \mu_{11} + \mu_{12} + \mu_{13} + \mu_{14} + \mu_{15} = 0.7 \wedge \\
& \mu_{21} + \mu_{22} + \mu_{23} + \mu_{24} + \mu_{25} = 0.0 \wedge \\
& \mu_{31} + \mu_{32} + \mu_{33} + \mu_{34} + \mu_{35} = 0.0 \wedge \\
& \mu_{41} + \mu_{42} + \mu_{43} + \mu_{44} + \mu_{45} = 0.2 \wedge \\
& \mu_{51} + \mu_{52} + \mu_{53} + \mu_{54} + \mu_{55} = 0.1 \wedge \\
& (d_{11}\mu_{11} + d_{12}\mu_{12} + d_{13}\mu_{13} + d_{14}\mu_{14} + d_{15}\mu_{15} + \\
& d_{21}\mu_{21} + d_{22}\mu_{22} + d_{23}\mu_{23} + d_{24}\mu_{24} + d_{25}\mu_{25} + \\
& d_{31}\mu_{31} + d_{32}\mu_{32} + d_{33}\mu_{33} + d_{34}\mu_{34} + d_{35}\mu_{35} + \\
& d_{41}\mu_{41} + d_{42}\mu_{42} + d_{43}\mu_{43} + d_{44}\mu_{44} + d_{45}\mu_{45} + \\
& d_{51}\mu_{51} + d_{52}\mu_{52} + d_{53}\mu_{53} + d_{54}\mu_{54} + d_{55}\mu_{55}) \leq d_{32} \\
& ) \\
& ) \\
& \vee \\
& ( \\
& 0.0 + 0.0 + 1.0 + 0.0 + 0.0 = 0 \wedge \\
& 0.7 + 0.0 + 0.0 + 0.2 + 0.1 = 0 \wedge \\
& 0 \geq d_{32} \\
& ) \\
& \vee \\
& ((( \\
& 0.0 + 0.0 + 1.0 + 0.0 + 0.0 > 0 \wedge \\
& 0.7 + 0.0 + 0.0 + 0.2 + 0.1 = 0 \\
& ) \vee ( \\
& 0.0 + 0.0 + 1.0 + 0.0 + 0.0 = 0 \wedge
\end{aligned}$$





$$\begin{aligned}
& \mu_{51} \geq 0 \wedge \mu_{51} \leq 1 \wedge \mu_{52} \geq 0 \wedge \mu_{52} \leq 1 \wedge \mu_{53} \geq 0 \wedge \mu_{53} \leq 1 \wedge \mu_{54} \geq 0 \wedge \mu_{54} \leq 1 \wedge \mu_{55} \geq 0 \wedge \mu_{55} \leq 1 \wedge \\
& \mu_{11} + \mu_{21} + \mu_{31} + \mu_{41} + \mu_{51} = 0.0 \wedge \\
& \mu_{12} + \mu_{22} + \mu_{32} + \mu_{42} + \mu_{52} = 0.0 \wedge \\
& \mu_{13} + \mu_{23} + \mu_{33} + \mu_{43} + \mu_{53} = 0.0 \wedge \\
& \mu_{14} + \mu_{24} + \mu_{34} + \mu_{44} + \mu_{54} = 0.0 \wedge \\
& \mu_{15} + \mu_{25} + \mu_{35} + \mu_{45} + \mu_{55} = 0.0 \wedge \\
& \mu_{11} + \mu_{12} + \mu_{13} + \mu_{14} + \mu_{15} = 0.0 \wedge \\
& \mu_{21} + \mu_{22} + \mu_{23} + \mu_{24} + \mu_{25} = 0.4 \wedge \\
& \mu_{31} + \mu_{32} + \mu_{33} + \mu_{34} + \mu_{35} = 0.6 \wedge \\
& \mu_{41} + \mu_{42} + \mu_{43} + \mu_{44} + \mu_{45} = 0.0 \wedge \\
& \mu_{51} + \mu_{52} + \mu_{53} + \mu_{54} + \mu_{55} = 0.0 \wedge \\
& (d_{11}\mu_{11} + d_{12}\mu_{12} + d_{13}\mu_{13} + d_{14}\mu_{14} + d_{15}\mu_{15} + \\
& d_{21}\mu_{21} + d_{22}\mu_{22} + d_{23}\mu_{23} + d_{24}\mu_{24} + d_{25}\mu_{25} + \\
& d_{31}\mu_{31} + d_{32}\mu_{32} + d_{33}\mu_{33} + d_{34}\mu_{34} + d_{35}\mu_{35} + \\
& d_{41}\mu_{41} + d_{42}\mu_{42} + d_{43}\mu_{43} + d_{44}\mu_{44} + d_{45}\mu_{45} + \\
& d_{51}\mu_{51} + d_{52}\mu_{52} + d_{53}\mu_{53} + d_{54}\mu_{54} + d_{55}\mu_{55}) \leq d_{41} \\
& ) \\
& ) \\
& \vee \\
& ( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 0.0 = 0 \wedge \\
& 0.0 + 0.4 + 0.6 + 0.0 + 0.0 = 0 \wedge \\
& 0 \geq d_{41} \\
& ) \\
& \vee \\
& ((( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 0.0 > 0 \wedge \\
& 0.0 + 0.4 + 0.6 + 0.0 + 0.0 = 0 \\
& ) \vee ( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 0.0 = 0 \wedge \\
& 0.0 + 0.4 + 0.6 + 0.0 + 0.0 > 0 \wedge \\
& )) \wedge \\
& 1 \leq d_{41} ) \\
& ) \\
& ( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 0.0 > 0 \wedge \\
& 0.7 + 0.0 + 0.0 + 0.2 + 0.1 > 0 \wedge \\
& \exists (\mu_{11}\mu_{12}\mu_{13}\mu_{14}\mu_{15}\mu_{21}\mu_{22}\mu_{23}\mu_{24}\mu_{25}\mu_{31}\mu_{32}\mu_{33}\mu_{34}\mu_{35}\mu_{41}\mu_{42}\mu_{43}\mu_{44}\mu_{45}\mu_{51}\mu_{52}\mu_{53}\mu_{54}\mu_{55}) \{ \\
& \mu_{11} \geq 0 \wedge \mu_{11} \leq 1 \wedge \mu_{12} \geq 0 \wedge \mu_{12} \leq 1 \wedge \mu_{13} \geq 0 \wedge \mu_{13} \leq 1 \wedge \mu_{14} \geq 0 \wedge \mu_{14} \leq 1 \wedge \mu_{15} \geq 0 \wedge \mu_{15} \leq 1 \wedge \\
& \mu_{21} \geq 0 \wedge \mu_{21} \leq 1 \wedge \mu_{22} \geq 0 \wedge \mu_{22} \leq 1 \wedge \mu_{23} \geq 0 \wedge \mu_{23} \leq 1 \wedge \mu_{24} \geq 0 \wedge \mu_{24} \leq 1 \wedge \mu_{25} \geq 0 \wedge \mu_{25} \leq 1 \wedge \\
& \mu_{31} \geq 0 \wedge \mu_{31} \leq 1 \wedge \mu_{32} \geq 0 \wedge \mu_{32} \leq 1 \wedge \mu_{33} \geq 0 \wedge \mu_{33} \leq 1 \wedge \mu_{34} \geq 0 \wedge \mu_{34} \leq 1 \wedge \mu_{35} \geq 0 \wedge \mu_{35} \leq 1 \wedge \\
& \mu_{41} \geq 0 \wedge \mu_{41} \leq 1 \wedge \mu_{42} \geq 0 \wedge \mu_{42} \leq 1 \wedge \mu_{43} \geq 0 \wedge \mu_{43} \leq 1 \wedge \mu_{44} \geq 0 \wedge \mu_{44} \leq 1 \wedge \mu_{45} \geq 0 \wedge \mu_{45} \leq 1 \wedge \\
& \mu_{51} \geq 0 \wedge \mu_{51} \leq 1 \wedge \mu_{52} \geq 0 \wedge \mu_{52} \leq 1 \wedge \mu_{53} \geq 0 \wedge \mu_{53} \leq 1 \wedge \mu_{54} \geq 0 \wedge \mu_{54} \leq 1 \wedge \mu_{55} \geq 0 \wedge \mu_{55} \leq 1 \wedge \\
& \mu_{11} + \mu_{21} + \mu_{31} + \mu_{41} + \mu_{51} = 0.0 \wedge \\
& \mu_{12} + \mu_{22} + \mu_{32} + \mu_{42} + \mu_{52} = 0.0 \wedge \\
& \mu_{13} + \mu_{23} + \mu_{33} + \mu_{43} + \mu_{53} = 0.0 \wedge \\
& \mu_{14} + \mu_{24} + \mu_{34} + \mu_{44} + \mu_{54} = 0.0 \wedge \\
& \mu_{15} + \mu_{25} + \mu_{35} + \mu_{45} + \mu_{55} = 0.0 \wedge \\
& \mu_{11} + \mu_{12} + \mu_{13} + \mu_{14} + \mu_{15} = 0.7 \wedge \\
& \mu_{21} + \mu_{22} + \mu_{23} + \mu_{24} + \mu_{25} = 0.0 \wedge \\
& \mu_{31} + \mu_{32} + \mu_{33} + \mu_{34} + \mu_{35} = 0.0 \wedge \\
& \mu_{41} + \mu_{42} + \mu_{43} + \mu_{44} + \mu_{45} = 0.2 \wedge \\
& \mu_{51} + \mu_{52} + \mu_{53} + \mu_{54} + \mu_{55} = 0.1 \wedge \\
& (d_{11}\mu_{11} + d_{12}\mu_{12} + d_{13}\mu_{13} + d_{14}\mu_{14} + d_{15}\mu_{15} + \\
& d_{21}\mu_{21} + d_{22}\mu_{22} + d_{23}\mu_{23} + d_{24}\mu_{24} + d_{25}\mu_{25} + \\
& d_{31}\mu_{31} + d_{32}\mu_{32} + d_{33}\mu_{33} + d_{34}\mu_{34} + d_{35}\mu_{35} + \\
& d_{41}\mu_{41} + d_{42}\mu_{42} + d_{43}\mu_{43} + d_{44}\mu_{44} + d_{45}\mu_{45} + \\
& d_{51}\mu_{51} + d_{52}\mu_{52} + d_{53}\mu_{53} + d_{54}\mu_{54} + d_{55}\mu_{55}) \leq d_{42} \\
& ) \\
& ) \\
& \vee \\
& ( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 0.0 = 0 \wedge \\
& 0.7 + 0.0 + 0.0 + 0.2 + 0.1 = 0 \wedge \\
& 0 \geq d_{42} \\
& ) \\
& \vee \\
& ((( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 0.0 > 0 \wedge \\
& 0.7 + 0.0 + 0.0 + 0.2 + 0.1 = 0
\end{aligned}$$



$$\begin{aligned}
& d_{31}\mu_{31} + d_{32}\mu_{32} + d_{33}\mu_{33} + d_{34}\mu_{34} + d_{35}\mu_{35} + \\
& d_{41}\mu_{41} + d_{42}\mu_{42} + d_{43}\mu_{43} + d_{44}\mu_{44} + d_{45}\mu_{45} + \\
& d_{51}\mu_{51} + d_{52}\mu_{52} + d_{53}\mu_{53} + d_{54}\mu_{54} + d_{55}\mu_{55} \leq d_{44} \\
& ) \\
& ) \\
& \vee \\
& ( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 0.0 = 0 \wedge \\
& 0.0 + 0.0 + 0.0 + 0.0 + 0.0 = 0 \wedge \\
& 0 \geq d_{44} \\
& ) \\
& \vee \\
& ((( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 0.0 > 0 \wedge \\
& 0.0 + 0.0 + 0.0 + 0.0 + 0.0 = 0 \\
& ) \vee ( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 0.0 = 0 \wedge \\
& 0.0 + 0.0 + 0.0 + 0.0 + 0.0 > 0 \wedge \\
& )) \wedge \\
& 1 \leq d_{44} ) \\
& ) \\
& ( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 0.0 > 0 \wedge \\
& 0.0 + 0.0 + 0.0 + 0.0 + 1.0 > 0 \wedge \\
& \exists (\mu_{11}\mu_{12}\mu_{13}\mu_{14}\mu_{15}\mu_{21}\mu_{22}\mu_{23}\mu_{24}\mu_{25}\mu_{31}\mu_{32}\mu_{33}\mu_{34}\mu_{35}\mu_{41}\mu_{42}\mu_{43}\mu_{44}\mu_{45}\mu_{51}\mu_{52}\mu_{53}\mu_{54}\mu_{55}) \{ \\
& \quad \mu_{11} \geq 0 \wedge \mu_{11} \leq 1 \wedge \mu_{12} \geq 0 \wedge \mu_{12} \leq 1 \wedge \mu_{13} \geq 0 \wedge \mu_{13} \leq 1 \wedge \mu_{14} \geq 0 \wedge \mu_{14} \leq 1 \wedge \mu_{15} \geq 0 \wedge \mu_{15} \leq 1 \wedge \\
& \quad \mu_{21} \geq 0 \wedge \mu_{21} \leq 1 \wedge \mu_{22} \geq 0 \wedge \mu_{22} \leq 1 \wedge \mu_{23} \geq 0 \wedge \mu_{23} \leq 1 \wedge \mu_{24} \geq 0 \wedge \mu_{24} \leq 1 \wedge \mu_{25} \geq 0 \wedge \mu_{25} \leq 1 \wedge \\
& \quad \mu_{31} \geq 0 \wedge \mu_{31} \leq 1 \wedge \mu_{32} \geq 0 \wedge \mu_{32} \leq 1 \wedge \mu_{33} \geq 0 \wedge \mu_{33} \leq 1 \wedge \mu_{34} \geq 0 \wedge \mu_{34} \leq 1 \wedge \mu_{35} \geq 0 \wedge \mu_{35} \leq 1 \wedge \\
& \quad \mu_{41} \geq 0 \wedge \mu_{41} \leq 1 \wedge \mu_{42} \geq 0 \wedge \mu_{42} \leq 1 \wedge \mu_{43} \geq 0 \wedge \mu_{43} \leq 1 \wedge \mu_{44} \geq 0 \wedge \mu_{44} \leq 1 \wedge \mu_{45} \geq 0 \wedge \mu_{45} \leq 1 \wedge \\
& \quad \mu_{51} \geq 0 \wedge \mu_{51} \leq 1 \wedge \mu_{52} \geq 0 \wedge \mu_{52} \leq 1 \wedge \mu_{53} \geq 0 \wedge \mu_{53} \leq 1 \wedge \mu_{54} \geq 0 \wedge \mu_{54} \leq 1 \wedge \mu_{55} \geq 0 \wedge \mu_{55} \leq 1 \wedge \\
& \quad \mu_{11} + \mu_{21} + \mu_{31} + \mu_{41} + \mu_{51} = 0.0 \wedge \\
& \quad \mu_{12} + \mu_{22} + \mu_{32} + \mu_{42} + \mu_{52} = 0.0 \wedge \\
& \quad \mu_{13} + \mu_{23} + \mu_{33} + \mu_{43} + \mu_{53} = 0.0 \wedge \\
& \quad \mu_{14} + \mu_{24} + \mu_{34} + \mu_{44} + \mu_{54} = 0.0 \wedge \\
& \quad \mu_{15} + \mu_{25} + \mu_{35} + \mu_{45} + \mu_{55} = 0.0 \wedge \\
& \quad \mu_{11} + \mu_{12} + \mu_{13} + \mu_{14} + \mu_{15} = 0.0 \wedge \\
& \quad \mu_{21} + \mu_{22} + \mu_{23} + \mu_{24} + \mu_{25} = 0.0 \wedge \\
& \quad \mu_{31} + \mu_{32} + \mu_{33} + \mu_{34} + \mu_{35} = 0.0 \wedge \\
& \quad \mu_{41} + \mu_{42} + \mu_{43} + \mu_{44} + \mu_{45} = 0.0 \wedge \\
& \quad \mu_{51} + \mu_{52} + \mu_{53} + \mu_{54} + \mu_{55} = 1.0 \wedge \\
& \quad (d_{11}\mu_{11} + d_{12}\mu_{12} + d_{13}\mu_{13} + d_{14}\mu_{14} + d_{15}\mu_{15} + \\
& \quad d_{21}\mu_{21} + d_{22}\mu_{22} + d_{23}\mu_{23} + d_{24}\mu_{24} + d_{25}\mu_{25} + \\
& \quad d_{31}\mu_{31} + d_{32}\mu_{32} + d_{33}\mu_{33} + d_{34}\mu_{34} + d_{35}\mu_{35} + \\
& \quad d_{41}\mu_{41} + d_{42}\mu_{42} + d_{43}\mu_{43} + d_{44}\mu_{44} + d_{45}\mu_{45} + \\
& \quad d_{51}\mu_{51} + d_{52}\mu_{52} + d_{53}\mu_{53} + d_{54}\mu_{54} + d_{55}\mu_{55}) \leq d_{45} \\
& ) \\
& ) \\
& \vee \\
& ( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 0.0 = 0 \wedge \\
& 0.0 + 0.0 + 0.0 + 0.0 + 1.0 = 0 \wedge \\
& 0 \geq d_{45} \\
& ) \\
& \vee \\
& ((( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 0.0 > 0 \wedge \\
& 0.0 + 0.0 + 0.0 + 0.0 + 1.0 = 0 \\
& ) \vee ( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 0.0 = 0 \wedge \\
& 0.0 + 0.0 + 0.0 + 0.0 + 1.0 > 0 \wedge \\
& )) \wedge \\
& 1 \leq d_{45} ) \\
& ) \\
& ( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 1.0 > 0 \wedge \\
& 0.0 + 0.4 + 0.6 + 0.0 + 0.0 > 0 \wedge \\
& \exists (\mu_{11}\mu_{12}\mu_{13}\mu_{14}\mu_{15}\mu_{21}\mu_{22}\mu_{23}\mu_{24}\mu_{25}\mu_{31}\mu_{32}\mu_{33}\mu_{34}\mu_{35}\mu_{41}\mu_{42}\mu_{43}\mu_{44}\mu_{45}\mu_{51}\mu_{52}\mu_{53}\mu_{54}\mu_{55}) \{ \\
& \quad \mu_{11} \geq 0 \wedge \mu_{11} \leq 1 \wedge \mu_{12} \geq 0 \wedge \mu_{12} \leq 1 \wedge \mu_{13} \geq 0 \wedge \mu_{13} \leq 1 \wedge \mu_{14} \geq 0 \wedge \mu_{14} \leq 1 \wedge \mu_{15} \geq 0 \wedge \mu_{15} \leq 1 \wedge \\
& \quad \mu_{21} \geq 0 \wedge \mu_{21} \leq 1 \wedge \mu_{22} \geq 0 \wedge \mu_{22} \leq 1 \wedge \mu_{23} \geq 0 \wedge \mu_{23} \leq 1 \wedge \mu_{24} \geq 0 \wedge \mu_{24} \leq 1 \wedge \mu_{25} \geq 0 \wedge \mu_{25} \leq 1 \wedge
\end{aligned}$$

$$\begin{aligned}
& \mu_{31} \geq 0 \wedge \mu_{31} \leq 1 \wedge \mu_{32} \geq 0 \wedge \mu_{32} \leq 1 \wedge \mu_{33} \geq 0 \wedge \mu_{33} \leq 1 \wedge \mu_{34} \geq 0 \wedge \mu_{34} \leq 1 \wedge \mu_{35} \geq 0 \wedge \mu_{35} \leq 1 \wedge \\
& \mu_{41} \geq 0 \wedge \mu_{41} \leq 1 \wedge \mu_{42} \geq 0 \wedge \mu_{42} \leq 1 \wedge \mu_{43} \geq 0 \wedge \mu_{43} \leq 1 \wedge \mu_{44} \geq 0 \wedge \mu_{44} \leq 1 \wedge \mu_{45} \geq 0 \wedge \mu_{45} \leq 1 \wedge \\
& \mu_{51} \geq 0 \wedge \mu_{51} \leq 1 \wedge \mu_{52} \geq 0 \wedge \mu_{52} \leq 1 \wedge \mu_{53} \geq 0 \wedge \mu_{53} \leq 1 \wedge \mu_{54} \geq 0 \wedge \mu_{54} \leq 1 \wedge \mu_{55} \geq 0 \wedge \mu_{55} \leq 1 \wedge \\
& \mu_{11} + \mu_{21} + \mu_{31} + \mu_{41} + \mu_{51} = 0.0 \wedge \\
& \mu_{12} + \mu_{22} + \mu_{32} + \mu_{42} + \mu_{52} = 0.0 \wedge \\
& \mu_{13} + \mu_{23} + \mu_{33} + \mu_{43} + \mu_{53} = 0.0 \wedge \\
& \mu_{14} + \mu_{24} + \mu_{34} + \mu_{44} + \mu_{54} = 0.0 \wedge \\
& \mu_{15} + \mu_{25} + \mu_{35} + \mu_{45} + \mu_{55} = 1.0 \wedge \\
& \mu_{11} + \mu_{12} + \mu_{13} + \mu_{14} + \mu_{15} = 0.0 \wedge \\
& \mu_{21} + \mu_{22} + \mu_{23} + \mu_{24} + \mu_{25} = 0.4 \wedge \\
& \mu_{31} + \mu_{32} + \mu_{33} + \mu_{34} + \mu_{35} = 0.6 \wedge \\
& \mu_{41} + \mu_{42} + \mu_{43} + \mu_{44} + \mu_{45} = 0.0 \wedge \\
& \mu_{51} + \mu_{52} + \mu_{53} + \mu_{54} + \mu_{55} = 0.0 \wedge \\
& (d_{11}\mu_{11} + d_{12}\mu_{12} + d_{13}\mu_{13} + d_{14}\mu_{14} + d_{15}\mu_{15} + \\
& d_{21}\mu_{21} + d_{22}\mu_{22} + d_{23}\mu_{23} + d_{24}\mu_{24} + d_{25}\mu_{25} + \\
& d_{31}\mu_{31} + d_{32}\mu_{32} + d_{33}\mu_{33} + d_{34}\mu_{34} + d_{35}\mu_{35} + \\
& d_{41}\mu_{41} + d_{42}\mu_{42} + d_{43}\mu_{43} + d_{44}\mu_{44} + d_{45}\mu_{45} + \\
& d_{51}\mu_{51} + d_{52}\mu_{52} + d_{53}\mu_{53} + d_{54}\mu_{54} + d_{55}\mu_{55}) \leq d_{51} \\
& ) \\
& ) \\
& \vee \\
& ( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 1.0 = 0 \wedge \\
& 0.0 + 0.4 + 0.6 + 0.0 + 0.0 = 0 \wedge \\
& 0 \geq d_{51} \\
& ) \\
& \vee \\
& ((( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 1.0 > 0 \wedge \\
& 0.0 + 0.4 + 0.6 + 0.0 + 0.0 = 0 \\
& ) \vee ( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 1.0 = 0 \wedge \\
& 0.0 + 0.4 + 0.6 + 0.0 + 0.0 > 0 \wedge \\
& )) \wedge \\
& 1 \leq d_{51} ) \\
& ) \\
& ( \\
& ( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 1.0 > 0 \wedge \\
& 0.7 + 0.0 + 0.0 + 0.2 + 0.1 > 0 \wedge \\
& \exists (\mu_{11}\mu_{12}\mu_{13}\mu_{14}\mu_{15}\mu_{21}\mu_{22}\mu_{23}\mu_{24}\mu_{25}\mu_{31}\mu_{32}\mu_{33}\mu_{34}\mu_{35}\mu_{41}\mu_{42}\mu_{43}\mu_{44}\mu_{45}\mu_{51}\mu_{52}\mu_{53}\mu_{54}\mu_{55}) \{ \\
& \mu_{11} \geq 0 \wedge \mu_{11} \leq 1 \wedge \mu_{12} \geq 0 \wedge \mu_{12} \leq 1 \wedge \mu_{13} \geq 0 \wedge \mu_{13} \leq 1 \wedge \mu_{14} \geq 0 \wedge \mu_{14} \leq 1 \wedge \mu_{15} \geq 0 \wedge \mu_{15} \leq 1 \wedge \\
& \mu_{21} \geq 0 \wedge \mu_{21} \leq 1 \wedge \mu_{22} \geq 0 \wedge \mu_{22} \leq 1 \wedge \mu_{23} \geq 0 \wedge \mu_{23} \leq 1 \wedge \mu_{24} \geq 0 \wedge \mu_{24} \leq 1 \wedge \mu_{25} \geq 0 \wedge \mu_{25} \leq 1 \wedge \\
& \mu_{31} \geq 0 \wedge \mu_{31} \leq 1 \wedge \mu_{32} \geq 0 \wedge \mu_{32} \leq 1 \wedge \mu_{33} \geq 0 \wedge \mu_{33} \leq 1 \wedge \mu_{34} \geq 0 \wedge \mu_{34} \leq 1 \wedge \mu_{35} \geq 0 \wedge \mu_{35} \leq 1 \wedge \\
& \mu_{41} \geq 0 \wedge \mu_{41} \leq 1 \wedge \mu_{42} \geq 0 \wedge \mu_{42} \leq 1 \wedge \mu_{43} \geq 0 \wedge \mu_{43} \leq 1 \wedge \mu_{44} \geq 0 \wedge \mu_{44} \leq 1 \wedge \mu_{45} \geq 0 \wedge \mu_{45} \leq 1 \wedge \\
& \mu_{51} \geq 0 \wedge \mu_{51} \leq 1 \wedge \mu_{52} \geq 0 \wedge \mu_{52} \leq 1 \wedge \mu_{53} \geq 0 \wedge \mu_{53} \leq 1 \wedge \mu_{54} \geq 0 \wedge \mu_{54} \leq 1 \wedge \mu_{55} \geq 0 \wedge \mu_{55} \leq 1 \wedge \\
& \mu_{11} + \mu_{21} + \mu_{31} + \mu_{41} + \mu_{51} = 0.0 \wedge \\
& \mu_{12} + \mu_{22} + \mu_{32} + \mu_{42} + \mu_{52} = 0.0 \wedge \\
& \mu_{13} + \mu_{23} + \mu_{33} + \mu_{43} + \mu_{53} = 0.0 \wedge \\
& \mu_{14} + \mu_{24} + \mu_{34} + \mu_{44} + \mu_{54} = 0.0 \wedge \\
& \mu_{15} + \mu_{25} + \mu_{35} + \mu_{45} + \mu_{55} = 1.0 \wedge \\
& \mu_{11} + \mu_{12} + \mu_{13} + \mu_{14} + \mu_{15} = 0.7 \wedge \\
& \mu_{21} + \mu_{22} + \mu_{23} + \mu_{24} + \mu_{25} = 0.0 \wedge \\
& \mu_{31} + \mu_{32} + \mu_{33} + \mu_{34} + \mu_{35} = 0.0 \wedge \\
& \mu_{41} + \mu_{42} + \mu_{43} + \mu_{44} + \mu_{45} = 0.2 \wedge \\
& \mu_{51} + \mu_{52} + \mu_{53} + \mu_{54} + \mu_{55} = 0.1 \wedge \\
& (d_{11}\mu_{11} + d_{12}\mu_{12} + d_{13}\mu_{13} + d_{14}\mu_{14} + d_{15}\mu_{15} + \\
& d_{21}\mu_{21} + d_{22}\mu_{22} + d_{23}\mu_{23} + d_{24}\mu_{24} + d_{25}\mu_{25} + \\
& d_{31}\mu_{31} + d_{32}\mu_{32} + d_{33}\mu_{33} + d_{34}\mu_{34} + d_{35}\mu_{35} + \\
& d_{41}\mu_{41} + d_{42}\mu_{42} + d_{43}\mu_{43} + d_{44}\mu_{44} + d_{45}\mu_{45} + \\
& d_{51}\mu_{51} + d_{52}\mu_{52} + d_{53}\mu_{53} + d_{54}\mu_{54} + d_{55}\mu_{55}) \leq d_{52} \\
& ) \\
& ) \\
& \vee \\
& ( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 1.0 = 0 \wedge \\
& 0.7 + 0.0 + 0.0 + 0.2 + 0.1 = 0 \wedge \\
& 0 \geq d_{52} \\
& ) \\
& \vee \\
& (((
\end{aligned}$$



$0.0 + 0.0 + 0.0 + 0.0 + 1.0 > 0 \wedge$   
 $0.7 + 0.0 + 0.0 + 0.2 + 0.1 = 0$   
 $) \vee ($   
 $0.0 + 0.0 + 0.0 + 0.0 + 1.0 = 0 \wedge$   
 $0.7 + 0.0 + 0.0 + 0.2 + 0.1 > 0 \wedge$   
 $)) \wedge$   
 $1 \leq d_{52}$   
 $)$   
 $($   
 $($   
 $0.0 + 0.0 + 0.0 + 0.0 + 1.0 > 0 \wedge$   
 $0.0 + 0.0 + 1.0 + 0.0 + 0.0 > 0 \wedge$   
 $\exists (\mu_{11} \mu_{12} \mu_{13} \mu_{14} \mu_{15} \mu_{21} \mu_{22} \mu_{23} \mu_{24} \mu_{25} \mu_{31} \mu_{32} \mu_{33} \mu_{34} \mu_{35} \mu_{41} \mu_{42} \mu_{43} \mu_{44} \mu_{45} \mu_{51} \mu_{52} \mu_{53} \mu_{54} \mu_{55}) \{$   
 $\mu_{11} \geq 0 \wedge \mu_{11} \leq 1 \wedge \mu_{12} \geq 0 \wedge \mu_{12} \leq 1 \wedge \mu_{13} \geq 0 \wedge \mu_{13} \leq 1 \wedge \mu_{14} \geq 0 \wedge \mu_{14} \leq 1 \wedge \mu_{15} \geq 0 \wedge \mu_{15} \leq 1 \wedge$   
 $\mu_{21} \geq 0 \wedge \mu_{21} \leq 1 \wedge \mu_{22} \geq 0 \wedge \mu_{22} \leq 1 \wedge \mu_{23} \geq 0 \wedge \mu_{23} \leq 1 \wedge \mu_{24} \geq 0 \wedge \mu_{24} \leq 1 \wedge \mu_{25} \geq 0 \wedge \mu_{25} \leq 1 \wedge$   
 $\mu_{31} \geq 0 \wedge \mu_{31} \leq 1 \wedge \mu_{32} \geq 0 \wedge \mu_{32} \leq 1 \wedge \mu_{33} \geq 0 \wedge \mu_{33} \leq 1 \wedge \mu_{34} \geq 0 \wedge \mu_{34} \leq 1 \wedge \mu_{35} \geq 0 \wedge \mu_{35} \leq 1 \wedge$   
 $\mu_{41} \geq 0 \wedge \mu_{41} \leq 1 \wedge \mu_{42} \geq 0 \wedge \mu_{42} \leq 1 \wedge \mu_{43} \geq 0 \wedge \mu_{43} \leq 1 \wedge \mu_{44} \geq 0 \wedge \mu_{44} \leq 1 \wedge \mu_{45} \geq 0 \wedge \mu_{45} \leq 1 \wedge$   
 $\mu_{51} \geq 0 \wedge \mu_{51} \leq 1 \wedge \mu_{52} \geq 0 \wedge \mu_{52} \leq 1 \wedge \mu_{53} \geq 0 \wedge \mu_{53} \leq 1 \wedge \mu_{54} \geq 0 \wedge \mu_{54} \leq 1 \wedge \mu_{55} \geq 0 \wedge \mu_{55} \leq 1 \wedge$   
 $\mu_{11} + \mu_{21} + \mu_{31} + \mu_{41} + \mu_{51} = 0.0 \wedge$   
 $\mu_{12} + \mu_{22} + \mu_{32} + \mu_{42} + \mu_{52} = 0.0 \wedge$   
 $\mu_{13} + \mu_{23} + \mu_{33} + \mu_{43} + \mu_{53} = 0.0 \wedge$   
 $\mu_{14} + \mu_{24} + \mu_{34} + \mu_{44} + \mu_{54} = 0.0 \wedge$   
 $\mu_{15} + \mu_{25} + \mu_{35} + \mu_{45} + \mu_{55} = 1.0 \wedge$   
 $\mu_{11} + \mu_{12} + \mu_{13} + \mu_{14} + \mu_{15} = 0.0 \wedge$   
 $\mu_{21} + \mu_{22} + \mu_{23} + \mu_{24} + \mu_{25} = 0.0 \wedge$   
 $\mu_{31} + \mu_{32} + \mu_{33} + \mu_{34} + \mu_{35} = 1.0 \wedge$   
 $\mu_{41} + \mu_{42} + \mu_{43} + \mu_{44} + \mu_{45} = 0.0 \wedge$   
 $\mu_{51} + \mu_{52} + \mu_{53} + \mu_{54} + \mu_{55} = 0.0 \wedge$   
 $(d_{11}\mu_{11} + d_{12}\mu_{12} + d_{13}\mu_{13} + d_{14}\mu_{14} + d_{15}\mu_{15} +$   
 $d_{21}\mu_{21} + d_{22}\mu_{22} + d_{23}\mu_{23} + d_{24}\mu_{24} + d_{25}\mu_{25} +$   
 $d_{31}\mu_{31} + d_{32}\mu_{32} + d_{33}\mu_{33} + d_{34}\mu_{34} + d_{35}\mu_{35} +$   
 $d_{41}\mu_{41} + d_{42}\mu_{42} + d_{43}\mu_{43} + d_{44}\mu_{44} + d_{45}\mu_{45} +$   
 $d_{51}\mu_{51} + d_{52}\mu_{52} + d_{53}\mu_{53} + d_{54}\mu_{54} + d_{55}\mu_{55}) \leq d_{53}$   
 $\}$   
 $)$   
 $\vee$   
 $($   
 $0.0 + 0.0 + 0.0 + 0.0 + 1.0 = 0 \wedge$   
 $0.0 + 0.0 + 1.0 + 0.0 + 0.0 = 0 \wedge$   
 $0 \geq d_{53}$   
 $)$   
 $\vee$   
 $(($   
 $0.0 + 0.0 + 0.0 + 0.0 + 1.0 > 0 \wedge$   
 $0.0 + 0.0 + 1.0 + 0.0 + 0.0 = 0$   
 $) \vee ($   
 $0.0 + 0.0 + 0.0 + 0.0 + 1.0 = 0 \wedge$   
 $0.0 + 0.0 + 1.0 + 0.0 + 0.0 > 0 \wedge$   
 $)) \wedge$   
 $1 \leq d_{53}$   
 $)$   
 $($   
 $($   
 $0.0 + 0.0 + 0.0 + 0.0 + 1.0 > 0 \wedge$   
 $0.0 + 0.0 + 0.0 + 0.0 + 0.0 > 0 \wedge$   
 $\exists (\mu_{11} \mu_{12} \mu_{13} \mu_{14} \mu_{15} \mu_{21} \mu_{22} \mu_{23} \mu_{24} \mu_{25} \mu_{31} \mu_{32} \mu_{33} \mu_{34} \mu_{35} \mu_{41} \mu_{42} \mu_{43} \mu_{44} \mu_{45} \mu_{51} \mu_{52} \mu_{53} \mu_{54} \mu_{55}) \{$   
 $\mu_{11} \geq 0 \wedge \mu_{11} \leq 1 \wedge \mu_{12} \geq 0 \wedge \mu_{12} \leq 1 \wedge \mu_{13} \geq 0 \wedge \mu_{13} \leq 1 \wedge \mu_{14} \geq 0 \wedge \mu_{14} \leq 1 \wedge \mu_{15} \geq 0 \wedge \mu_{15} \leq 1 \wedge$   
 $\mu_{21} \geq 0 \wedge \mu_{21} \leq 1 \wedge \mu_{22} \geq 0 \wedge \mu_{22} \leq 1 \wedge \mu_{23} \geq 0 \wedge \mu_{23} \leq 1 \wedge \mu_{24} \geq 0 \wedge \mu_{24} \leq 1 \wedge \mu_{25} \geq 0 \wedge \mu_{25} \leq 1 \wedge$   
 $\mu_{31} \geq 0 \wedge \mu_{31} \leq 1 \wedge \mu_{32} \geq 0 \wedge \mu_{32} \leq 1 \wedge \mu_{33} \geq 0 \wedge \mu_{33} \leq 1 \wedge \mu_{34} \geq 0 \wedge \mu_{34} \leq 1 \wedge \mu_{35} \geq 0 \wedge \mu_{35} \leq 1 \wedge$   
 $\mu_{41} \geq 0 \wedge \mu_{41} \leq 1 \wedge \mu_{42} \geq 0 \wedge \mu_{42} \leq 1 \wedge \mu_{43} \geq 0 \wedge \mu_{43} \leq 1 \wedge \mu_{44} \geq 0 \wedge \mu_{44} \leq 1 \wedge \mu_{45} \geq 0 \wedge \mu_{45} \leq 1 \wedge$   
 $\mu_{51} \geq 0 \wedge \mu_{51} \leq 1 \wedge \mu_{52} \geq 0 \wedge \mu_{52} \leq 1 \wedge \mu_{53} \geq 0 \wedge \mu_{53} \leq 1 \wedge \mu_{54} \geq 0 \wedge \mu_{54} \leq 1 \wedge \mu_{55} \geq 0 \wedge \mu_{55} \leq 1 \wedge$   
 $\mu_{11} + \mu_{21} + \mu_{31} + \mu_{41} + \mu_{51} = 0.0 \wedge$   
 $\mu_{12} + \mu_{22} + \mu_{32} + \mu_{42} + \mu_{52} = 0.0 \wedge$   
 $\mu_{13} + \mu_{23} + \mu_{33} + \mu_{43} + \mu_{53} = 0.0 \wedge$   
 $\mu_{14} + \mu_{24} + \mu_{34} + \mu_{44} + \mu_{54} = 0.0 \wedge$   
 $\mu_{15} + \mu_{25} + \mu_{35} + \mu_{45} + \mu_{55} = 1.0 \wedge$   
 $\mu_{11} + \mu_{12} + \mu_{13} + \mu_{14} + \mu_{15} = 0.0 \wedge$   
 $\mu_{21} + \mu_{22} + \mu_{23} + \mu_{24} + \mu_{25} = 0.0 \wedge$   
 $\mu_{31} + \mu_{32} + \mu_{33} + \mu_{34} + \mu_{35} = 0.0 \wedge$   
 $\mu_{41} + \mu_{42} + \mu_{43} + \mu_{44} + \mu_{45} = 0.0 \wedge$   
 $\mu_{51} + \mu_{52} + \mu_{53} + \mu_{54} + \mu_{55} = 0.0 \wedge$

$$\begin{aligned}
& (d_{11}\mu_{11} + d_{12}\mu_{12} + d_{13}\mu_{13} + d_{14}\mu_{14} + d_{15}\mu_{15} + \\
& d_{21}\mu_{21} + d_{22}\mu_{22} + d_{23}\mu_{23} + d_{24}\mu_{24} + d_{25}\mu_{25} + \\
& d_{31}\mu_{31} + d_{32}\mu_{32} + d_{33}\mu_{33} + d_{34}\mu_{34} + d_{35}\mu_{35} + \\
& d_{41}\mu_{41} + d_{42}\mu_{42} + d_{43}\mu_{43} + d_{44}\mu_{44} + d_{45}\mu_{45} + \\
& d_{51}\mu_{51} + d_{52}\mu_{52} + d_{53}\mu_{53} + d_{54}\mu_{54} + d_{55}\mu_{55}) \leq d_{54} \\
& ) \\
& ) \\
& \vee \\
& ( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 1.0 = 0 \wedge \\
& 0.0 + 0.0 + 0.0 + 0.0 + 0.0 = 0 \wedge \\
& 0 \geq d_{54} \\
& ) \\
& \vee \\
& ((( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 1.0 > 0 \wedge \\
& 0.0 + 0.0 + 0.0 + 0.0 + 0.0 = 0 \\
& ) \vee ( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 1.0 = 0 \wedge \\
& 0.0 + 0.0 + 0.0 + 0.0 + 0.0 > 0 \wedge \\
& )) \wedge \\
& 1 \leq d_{54} ) \\
& ) \\
& ( \\
& ( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 1.0 > 0 \wedge \\
& 0.0 + 0.0 + 0.0 + 0.0 + 1.0 > 0 \wedge \\
& \exists (\mu_{11}\mu_{12}\mu_{13}\mu_{14}\mu_{15}\mu_{21}\mu_{22}\mu_{23}\mu_{24}\mu_{25}\mu_{31}\mu_{32}\mu_{33}\mu_{34}\mu_{35}\mu_{41}\mu_{42}\mu_{43}\mu_{44}\mu_{45}\mu_{51}\mu_{52}\mu_{53}\mu_{54}\mu_{55}) \{ \\
& \mu_{11} \geq 0 \wedge \mu_{11} \leq 1 \wedge \mu_{12} \geq 0 \wedge \mu_{12} \leq 1 \wedge \mu_{13} \geq 0 \wedge \mu_{13} \leq 1 \wedge \mu_{14} \geq 0 \wedge \mu_{14} \leq 1 \wedge \mu_{15} \geq 0 \wedge \mu_{15} \leq 1 \wedge \\
& \mu_{21} \geq 0 \wedge \mu_{21} \leq 1 \wedge \mu_{22} \geq 0 \wedge \mu_{22} \leq 1 \wedge \mu_{23} \geq 0 \wedge \mu_{23} \leq 1 \wedge \mu_{24} \geq 0 \wedge \mu_{24} \leq 1 \wedge \mu_{25} \geq 0 \wedge \mu_{25} \leq 1 \wedge \\
& \mu_{31} \geq 0 \wedge \mu_{31} \leq 1 \wedge \mu_{32} \geq 0 \wedge \mu_{32} \leq 1 \wedge \mu_{33} \geq 0 \wedge \mu_{33} \leq 1 \wedge \mu_{34} \geq 0 \wedge \mu_{34} \leq 1 \wedge \mu_{35} \geq 0 \wedge \mu_{35} \leq 1 \wedge \\
& \mu_{41} \geq 0 \wedge \mu_{41} \leq 1 \wedge \mu_{42} \geq 0 \wedge \mu_{42} \leq 1 \wedge \mu_{43} \geq 0 \wedge \mu_{43} \leq 1 \wedge \mu_{44} \geq 0 \wedge \mu_{44} \leq 1 \wedge \mu_{45} \geq 0 \wedge \mu_{45} \leq 1 \wedge \\
& \mu_{51} \geq 0 \wedge \mu_{51} \leq 1 \wedge \mu_{52} \geq 0 \wedge \mu_{52} \leq 1 \wedge \mu_{53} \geq 0 \wedge \mu_{53} \leq 1 \wedge \mu_{54} \geq 0 \wedge \mu_{54} \leq 1 \wedge \mu_{55} \geq 0 \wedge \mu_{55} \leq 1 \wedge \\
& \mu_{11} + \mu_{21} + \mu_{31} + \mu_{41} + \mu_{51} = 0.0 \wedge \\
& \mu_{12} + \mu_{22} + \mu_{32} + \mu_{42} + \mu_{52} = 0.0 \wedge \\
& \mu_{13} + \mu_{23} + \mu_{33} + \mu_{43} + \mu_{53} = 0.0 \wedge \\
& \mu_{14} + \mu_{24} + \mu_{34} + \mu_{44} + \mu_{54} = 0.0 \wedge \\
& \mu_{15} + \mu_{25} + \mu_{35} + \mu_{45} + \mu_{55} = 1.0 \wedge \\
& \mu_{11} + \mu_{12} + \mu_{13} + \mu_{14} + \mu_{15} = 0.0 \wedge \\
& \mu_{21} + \mu_{22} + \mu_{23} + \mu_{24} + \mu_{25} = 0.0 \wedge \\
& \mu_{31} + \mu_{32} + \mu_{33} + \mu_{34} + \mu_{35} = 0.0 \wedge \\
& \mu_{41} + \mu_{42} + \mu_{43} + \mu_{44} + \mu_{45} = 0.0 \wedge \\
& \mu_{51} + \mu_{52} + \mu_{53} + \mu_{54} + \mu_{55} = 1.0 \wedge \\
& (d_{11}\mu_{11} + d_{12}\mu_{12} + d_{13}\mu_{13} + d_{14}\mu_{14} + d_{15}\mu_{15} + \\
& d_{21}\mu_{21} + d_{22}\mu_{22} + d_{23}\mu_{23} + d_{24}\mu_{24} + d_{25}\mu_{25} + \\
& d_{31}\mu_{31} + d_{32}\mu_{32} + d_{33}\mu_{33} + d_{34}\mu_{34} + d_{35}\mu_{35} + \\
& d_{41}\mu_{41} + d_{42}\mu_{42} + d_{43}\mu_{43} + d_{44}\mu_{44} + d_{45}\mu_{45} + \\
& d_{51}\mu_{51} + d_{52}\mu_{52} + d_{53}\mu_{53} + d_{54}\mu_{54} + d_{55}\mu_{55}) \leq d_{55} \\
& ) \\
& ) \\
& \vee \\
& ( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 1.0 = 0 \wedge \\
& 0.0 + 0.0 + 0.0 + 0.0 + 1.0 = 0 \wedge \\
& 0 \geq d_{55} \\
& ) \\
& \vee \\
& ((( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 1.0 > 0 \wedge \\
& 0.0 + 0.0 + 0.0 + 0.0 + 1.0 = 0 \\
& ) \vee ( \\
& 0.0 + 0.0 + 0.0 + 0.0 + 1.0 = 0 \wedge \\
& 0.0 + 0.0 + 0.0 + 0.0 + 1.0 > 0 \wedge \\
& )) \wedge \\
& 1 \leq d_{55} ) \\
& )
\end{aligned}$$