Continuous Probability and Nondeterminism in Labeled Transition Systems

Thesis for the degree of Ph.D. in Computer Science
FaMAF – Universidad Nacional de Córdoba

Nicolás Wolovick
Advisor: Pedro R. D'Argenio

$$T_a : S \to \Delta(\Sigma)$$

$$\Sigma \leftarrow H(\Delta(\Sigma)) : T_a^{-1}$$

Committee: V. Braberman, P. Sánchez-Terraf, I. Viglizzo.

Córdoba, Argentina, March 15, 2012.

A Sole, Lucas y Lola,
sin ellos, *nada, nada, nada*[1]
hubiera sido posible.

[1]Leer esto con la entonación de M. E. Walsh en "*Historia de una Princesa, su papá y el Príncipe Kinoto Fukasuka*".

Agradecimientos

Desde Marzo 2004 cuando cursé la materia Funciones Reales a cargo de Linda Saal, hasta la defensa en Marzo de 2012, un montón de personas colaboraron, apoyaron, dieron aliento y hasta soplaron para que esta fase de mi formación concluya. A todos ellos, gracias.

A mi director, Pedro R. D'Argenio, que confió en mí para llevar adelante este tema. Sin su gran tozudez, que se balancea perfectamente con sus capacidades cognitivas, jamás hubiéramos llegado a $T_a : S \to \Delta(\Sigma)$. A Pedro Sánchez-Terraf que se incorporó luego al equipito y le puso mucho empuje, teoría, y Kechris a todo esto.

A los miembros del Tribunal de Tesis, por su cuidadosa lectura del manuscrito final. La versión final de esta Tesis mejoró gracias a sus aportes.

A Joost-Pieter Katoen y a Mariëlle Stoelinga del Grupo de Métodos Formales y Herramientas de la Universidad de Twente. A Holger Hermanns y todo su equipo del Grupo Sistemas y Software Dependibles de la Universidad de Saarland: Reza Pulungan, Sven Johr y Lijun Zhang. Los días en Saarbrücken fueron muy productivos en muchos aspectos de mi formación.

Al Programa Alßan[2] que en 2004 me permitió trabajar en la Universidad de Twente y en 2005-2006 en la Universidad de Saarland. A la SeCyT-UNC, que me otrogó una beca[3] en el tramo final de mi Doctorado.

[2]Financiado por el Programa Alßan, el Programa de la Unión Europea de Becas de Alto Nivel para América Latina. Beca E04D030410AR.

[3]Financiado por el programa Finalización de Maestrías y Doctorados de la SeCyT-UNC, desde Septiembre de 2010 a Febrero 2012.

Abstract

We define a model for interacting systems involving continuous probabilistic and nondeterministic choices over continuous state spaces. Our model is a generalization of labeled Markov process, a well established formalism that capture interacting probabilistic continuous systems with strong basis in Measure Theory, but lacking internal nondeterminism. We define the extension to continuous internal nondeterminism in such a way that it allows for quantification of nondeterminism through schedulers, and it also allows for a sound definition of an existential modal operator. The model is used to capture the semantics of a probabilistic timed stochastic process algebra, a stochastic hybrid automata, and a probabilistic and nondeterministic programming language; they are continuous systems involving a nontrivial mix of probabilities and nondeterminism. We also compare our model with other know continuous probabilistic and nondeterministic labeled transition systems and the embedding is defined if applicable. Behavioral equivalence is captured by one notion of point-wise strong bisimulation, and two notions of event-wise strong bisimulation. We show that the three notions are different. We define a two-level infinitary modal logic that characterize the coarser behavioral event-wise bisimulation. Finally we show that a model and a randomized history-dependent scheduler resolving the nondeterminism renders a probabilistic trace semantics.

Resumen

En este trabajo definimos un modelo para sistemas interactivos con probabilidades y nodeterminismo continuo sobre un espacio de estados continuo. Nuestro modelo generaliza los procesos de Markov etiquetados, un formalismo que captura sistemas interactivos continuos con probabilismo sobre una base fuerte en Teoría de la Medida, sin embargo, dicho modelo carece de nodeterminismo interno. Nuestra extensión a nodeterminismo interno continuo permite por un lado cuantificar ese nodeterminismo a través de planificadores, y también permite definir de manera consistente un operador existencial de lógica modal. Usamos nuestro modelo para capturar la semántica de un álgebra de procesos estocástica con tiempo, de un autómata híbrido estocástico y de un lenguaje de programación probabilístico y nodeterminístico; todos sistemas continuos que mezclan probabilidades y nodeterminismo de manera no trivial. También comparamos nuestro formalismo con otros sistemas de transición etiquetados conocidos que incluyen la posibilidad de nodeterminismo y probabilidades continuas, y si resulta posible, damos la transformación a nuestros sistemas. La equivalencia de comportamiento de los sistemas la capturamos con una noción puntual de bisimulación fuerte y dos nociones de bisimulación fuerte pero basada en eventos. Mostramos a través de ejemplos que las tres nociones son diferentes. Definimos una lógica infinitaria de dos niveles que caracteriza la equivalencia de comportamientos más gruesa basada en eventos. Finalmente, mostramos que un modelo más un planificador al azar que depende de la historia y resuelve el nodeterminismo, genera una semántica de trazas probabilísticas.

Contents

Chapter 1

Introduction

1.1 Motivation and Context

The interplay of probabilistic and nondeterministic choice in systems that live in continuous state space is becoming more common. Nowadays software applications for mobile devices mix those ingredients. They have discrete state (memory hierarchy) as well as continuous state (position, orientation, acceleration, battery voltage, etc.). These continuous quantities are perturbed by the environment, and internally, many algorithms make use of discrete probabilities. Moreover, they operate in meshes of devices where the relative speeds of execution among them are not known in advance, therefore there is no information on how these devices interleave their operations in the timeline. Observations of discrete values like enabled or disabled buttons, and also observations of continuous values like displayed roll angle in a cell phone, are part of these systems.

The massive amount of production of those mobile devices implies that the costs incurred in software faults are not minor. Methodologies, techniques and formalisms to tackle the complexity of this continuous probabilistic and nondeterministic systems, become an important part in the production and maintenance cycles as well as in their related costs

This thesis addresses the problem of defining a model that captures labeled transition systems in continuous state space, continuous nondeterminism, continuous probabilistic choices and continuous labels.

The work is centered in a sound mathematical definition of the model, behavioral equivalences, modal logic and schedulers. We give enough evidence that the definition is adequate, either by means of examples, as well as nontrivial results relating behavioral equivalences on the model and a logic.

We also capture the semantics of previously defined models in the fields of stochastic timed automata and stochastic hybrid systems.

We can classify the current modeling tools in the form of labeled transition systems that include nondeterministic and probabilistic choice together. This classification is by the type of state space (discrete or continuous) and the possibility to include, beside a probabilistic choice, external as well as internal nondeterminism. Table 1.1 summarizes our selection of models, and the position that our *nondeterministic labeled Markov processes* (NLMPs) model occupy.

Nondeterminism/State Space	Discrete	Continuous
None	MC	MP
External	PLTS	LMP
Internal & External	PA	**NLMP**

Table 1.1: Taxonomy of some probabilistic systems divided by kind of nondeterminism.

Markov chains (MC) [54] and Markov processes (MP) [27] are the classical models capturing discrete and continuous probabilities. If we augment MC with observable labels, we obtain probabilistic labeled transition systems (PLTS) [41]. This kind of systems are usually called *reactive models*. Labeled Markov processes (LMPs) [20] provide a continuous counterpart. If internal nondeterministic choice is added to the already present external nondeterminism of PLTS, we obtain probabilistic automata (PA) [58].

Our NLMPs are both an extension of PA to give them a measure theoretic sound basis to continuous state space, and an extension of LMPs to include internal nondeterministic choices.

One possible use of the nondeterminism is subspecification of systems. Suppose we want to model a coin that can behave biased towards heads or tails, but there is no information on how it is biased except that the probabilities of each side cannot be lower than 1/4. Observe that from certain perspective it represents a continuous span of probabilities, since every possible probabilistic behavior assigning at least 1/4 of chance to each side is possible.

There are in the literature two ways to model this kind of continuous probabilistic subspecification. They are represented in Figure 1.1.

On the left-hand side, there is a PA using discrete nondeterminism to model a subspecified coin. With the usual resolution of nondeterminism

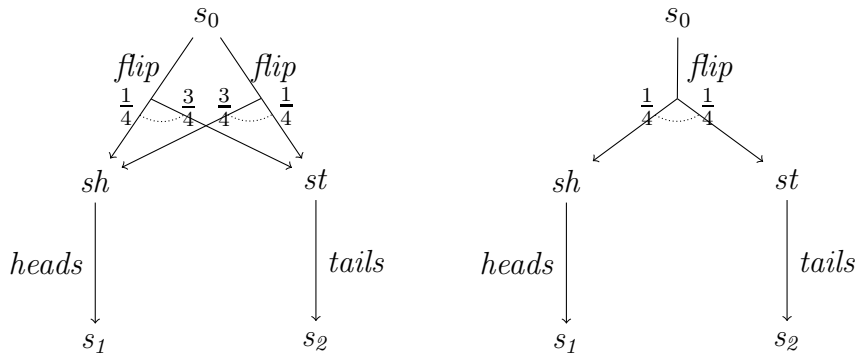


Figure 1.1: Two ways of modeling continuous subspecification of probabilistic choice.

provided by the schedulers, all possible intermediate probability distributions (convex combinations) can also be attained by involving a *probabilistic choice* over the possible actions. On the right-hand side, a subprobability distribution is used to encode partial information. This makes reference to all possible completions of this subprobability to a probability measure.

The NLMP model attacks this problem in a different way. It defines a *transition function* whose target is a (possibly) continuous *set of probabilities*. We can encode the nondeterminism of the probabilistic choice of Figure 1.1 in a concise way:

$$\begin{aligned} T_{\mathit{flip}}(s_0) &= \Delta^{=1}(\{sh, st\}) \cap \Delta^{\geq 1/4}(\{sh\}) \cap \Delta^{\geq 1/4}(\{st\}) \\ T_{\mathit{heads}}(sh) &= \Delta^{=1}(\{s_1\}) \\ T_{\mathit{tails}}(st) &= \Delta^{=1}(\{s_2\}) \end{aligned}$$

The first line says that from s_0 there is a transition with label *flip* to the set of all probabilities such that:

- the probability is concentrated in states sh and st: $\Delta^{=1}(\{sh, st\})$;
- there is at least $1/4$ of chance to move to state sh: $\Delta^{\geq 1/4}(\{sh\})$; and also
- there is at least $1/4$ of chance to move to state st: $\Delta^{\geq 1/4}(\{st\})$.

It is interesting to note that if we move from discrete to continuous state space, similarly simple expressions capture continuous nondeterminism. For example if we want to specify a system in which from s_0 there is a transition with label a such that:

- it can reach the $[0, 1]$ real interval with any probability measure; and
- there is at least $1/4$ of the probability in the point $\{0\}$, and at least $1/4$ in the point $\{1\}$;
- the rest of the probability is in the interval $[1/4, 3/4]$, but it is no greater than $1/2$.

This can be done with the following transition function of an NLMP:

$$T_a(s_0) = \Delta^{=1}([0,1]) \cap \Delta^{\geq 1/4}(\{0\}) \cap \Delta^{\geq 1/4}(\{1\}) \cap \Delta^{\leq 1/2}([1/4, 3/4])$$

Here the expression $\Delta^{\leq 1/2}([1/4, 3/4])$ refers to the set of all measures such that it quantifies at most $1/2$ in the interval $[1/4, 3/4]$, i.e.

$$\Delta^{\leq 1/2}([1/4, 3/4]) = \{\mu \mid \mu([1/4, 3/4]) \leq 1/2\}$$

Nondeterminism is more than subspecification, it also captures interleaving of concurrent systems, design choices and abstractions. This is why in this thesis we will develop NLMPs, a theory for nondeterministic and probabilistic transition functions in the setting of general state spaces.

1.2 Our Contribution

We obtained a measure theoretical sound definition of a transition function targeting nondeterministic probabilistic choices over topology-free state spaces. This allowed us to:

- define various notions of bisimulations, ranging from traditional to measure theoretic views,
- define a modal logic and prove that it characterizes some of the bisimulations, and also
- resolve the nondeterminism by means of a scheduler.

We give examples showing the strong capabilities of the model for probabilistic subspecification. We also show that our model captures the semantics of various nontrivial probabilistic and nondeterministic systems.

1.3 Thesis Outline

Besides this introduction, this thesis is composed of the following chapters:

Chapter 2: We briefly review the main concepts of labeled transitions systems (bisimulation, logical characterization, scheduler) that we are going to generalize to continuous state spaces.

Chapter 3: We present a strong background on Measure Theory, with special emphasis on the tools that we are going to use.

Chapter 4: This chapter introduces NLMPs, the model of measurable transition functions capturing probabilistic and nondeterministic choices.

Chapter 5: We present the σ-algebra of measures $\Delta(\Sigma)$ as a specification language. We also show how NLMPs captures the semantics of higher level models including probabilistic and nondeterministic choice. We compare similar models.

Chapter 6: We give three different notions of bisimulation for NLMPs, and we show the relations among them.

Chapter 7: We define schedulers as a way to resolve the nondeterminism. Using the schedulers we construct the probabilistic trace semantics.

Chapter 8: We summarize our work and achievements and give future directions of research.

1.4 Origin of the Thesis

Many results presented in this thesis appeared before in some of our papers. The definition of NLMPs (Chapter 4) and bisimulations (Chapter 6) are from [17, 18]. The non-probabilistic NLMPs of Section 4.4 as well as the examples in Section 6.2 separating the notions of bisimulation are from [17]. The semantics of SHA of Section 5.2 was presented in [28], but its technical details are new in this thesis. Chapter 7 is a generalization of [67].

Chapter 2

Background

Program semantics can be given as a transfer function from initial to final states, and it is usually called denotational semantics. Program meaning can also be captured by the so called small-step semantics where all the intermediate steps of the computation are explicit, therefore two different programs having the same in/out function can be distinguished. In the latter case the evolution of a program is given by a state relation $s \to s'$ giving rise to the so called transition systems. It could be deterministic (one successor state) or nondeterministic (many successor states). Nondeterministic constructs arise in Computer Science as a way to subspecify a program, later to be specialized, or as the result of the independent execution of actions in a parallel composition of interacting processes.

The parallel components evolve asynchronously but also synchronize through some form of communication. The synchronization mechanism is modeled in transition systems using labels. Those labels are a new component of the transition relation, so transitions become $s \xrightarrow{a} s'$. At each state there is a subset of enabled labels, the set of current interaction choices given by the process and observed by the environment.

If the programming language enriches its semantics with a probabilistic choice as in pGCL [46] ($P_1 \oplus_{\frac{1}{4}} P_2$), or random assignment ($x := \mathit{uniform}([0,1])$), then the labeled transition system should change to a relation where given a current state and a label, the next state is quantified by a probability distribution.

Probabilistic and nondeterministic choice can coexist and this gives rise to a very rich model where all the information levels are present: certainty (deterministic), quantified uncertainty (probabilistic) and pure uncertainty (nondeterministic). Moreover nondeterministic and probabilistic choices can be present at the same time, giving rise to what we call probabilistic subspecification.

The aim of the current section is to briefly review what we consider the three most prominent labeled transition systems over discrete state spaces that embody nondeterministic and probabilistic choices, and their combination. They are labeled transition systems (LTS) [37], probabilistic labeled transition systems (PLTS) [41] and probabilistic automata (PA) [59]. In this reviewing process we single out which are the main concepts we are going to extend in our work, where the most expressive version, namely the PA, is taken to the realms of continuous state space, continuous labels, continuous nondeterminism and continuous probabilitistic choice.

For more information on discrete systems including probabilistic and nondeterministic choices, we recommend [61, Chapter 2] and [63].

2.1 Labeled Transition Systems

Labeled transition systems (LTS) were introduced in [37] as a way to model concurrent program semantics. Its definition is similar to nondeterministic automata.

Definition 2.1 (LTS)**.** A *labeled transition system* is a tuple $(S, L, \rightarrow)$, where S is a countable *set of states*, L is a countable set of *labels* or *actions*, and $\rightarrow \subseteq S \times L \times S$ is a *transition relation*, where $(s, a, s') \in \rightarrow$ is written $s \xrightarrow{a} s'$.

Our concern is on interactive behavior, where the labeled transition systems conceptually depart from nondeterministic automaton. Finite automata is an accepting mechanism for words, whereas in LTS the focus is in the actions that are enabled at each step. The LTS mechanism is usually depicted as a black box holding the transition relation, exposing enabled/disabled buttons that the environment can observe and press, usually called external nondeterminism (Figure 2.1).

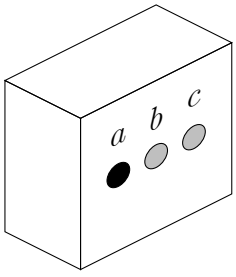


Figure 2.1: Reactive behavior of an LTS with labels $L = \{a, b, c\}$, where only a is enabled.

The following classical example [45] shows the difference between nondeterministic automata and LTS (see Figure 2.2).

Example 2.2. Here we show two language equivalent LTS (they both generate $\{ab, ac\}$), that are not behaviorally equivalent. In the left system after label a is chosen, b or c could be chosen; whereas in the right system, given that a was chosen, then either b or c may be refused, depending on which a was first executed.

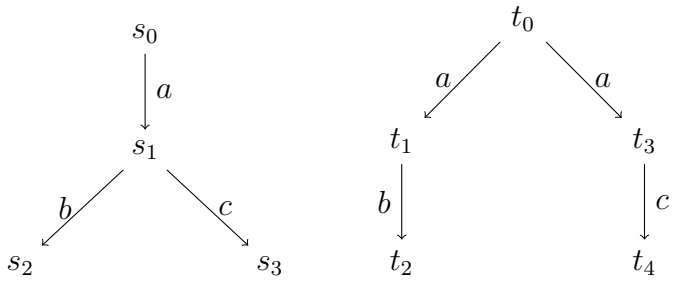


Figure 2.2: Language equivalent LTS that have different interactive behaviors.

The distributivity of sequential composition with respect to choice composition is lost: $a(b + c) \neq ab + ac$. One of the many concept that captures this behavioral equivalence is (strong) bisimulation [43, 51].

Definition 2.3 (Bisimulation)**.** A relation $R \subseteq S \times S$ is a *bisimulation* if for all $s, t \in S$ such that $s\ R\ t$, for all $a \in L$, $s \xrightarrow{a} s'$ implies that there is t' such that $t \xrightarrow{a} t'$ and $s'\ R\ t'$; and vice versa, that is, $t \xrightarrow{a} t'$ implies there is s' such that $s \xrightarrow{a} s'$ and $s'\ R\ t'$.

In Example 2.2 there is no bisimulation relation R between s_0 and t_0 in the left and the right LTS respectively. If we relate $s_0\ R\ t_0$ then s_1 should be related with both t_1, t_3, and $s_1\ R\ t_1$ is not possible since there is no outgoing transition from t_1 labeled with c.

Consider the lifting of a relation R to sets encoding the universal-existential quantification in both directions:

$$A\ R\ B \doteq (\forall a \in A, \exists b \in B, a\ R\ b) \wedge (\forall b \in B, \exists a \in A, a\ R\ b) \tag{2.1}$$

then bisimulation can be compactly recast using this lifting:

$$s\ R\ t \Rightarrow \forall a \in L, (s \xrightarrow{a})\ R\ (t \xrightarrow{a}) \tag{2.2}$$

The largest bisimulation is called bisimilarity and it is an equivalence relation.

Definition 2.4 (Bisimilarity)**.** The union of all bisimulations is a bisimulation and also the largest one. It is called *bisimilarity* $\sim$ and it is also an equivalence relation. The definition is as follows:

$$\sim \doteq \bigcup\{R \mid R \text{ is bisimulation}\}$$

Bisimilar states can be interchanged keeping the behavior of the whole system, and this is fundamental in order to provide different implementations for parts of a system.

Bisimilarity has a logical characterization in terms of what is called Hennessy-Milner logic [45, 60], a very terse logic with modalities.

Definition 2.5 (Hennessy-Milner logic)**.** The *Hennessy-Milner logic* syntax is defined inductively by:

$$\phi ::= \top \mid \phi_1 \wedge \phi_2 \mid \neg\phi \mid \langle a\rangle\phi$$

Its semantics is interpreted over LTS as the set of states where the formula is valid, i.e. $[\![\phi]\!] = \{s \mid s \models \phi\}$. Its definition is:

$$[\![\top]\!] = S \qquad [\![\phi_1 \wedge \phi_2]\!] = [\![\phi_1]\!] \cap [\![\phi_2]\!] \qquad [\![\neg\phi]\!] = [\![\phi]\!]^{\mathsf{c}}$$
$$[\![\langle a\rangle\phi]\!] = \{s \mid s \xrightarrow{a} s', s' \in [\![\phi]\!]\}$$

The next theorem states that the logic characterizes the bisimulation for finite LTS.

Theorem 2.1. *Two states of an LTS are bisimilar iff they satisfy exactly the same formulas of the Hennessy-Milner logic. i.e.* $s \sim t$ *iff* $\forall\phi, s \in [\![\phi]\!] \Leftrightarrow t \in [\![\phi]\!]$.

Using Theorem 2.1 in Example 2.2, we can compactly say $s_0 \nsim t_0$, since the semantics of the formula $\langle a\rangle\neg\langle c\rangle\top$ includes t_0 and excludes s_0.

There are also many notions of behavioral equivalence [32], in which Definition 2.3 (strong bisimulation) distinguishes more (the finer). At the other end of the spectrum, the coarse trace equivalence is very close to language equivalence of finite automata. There are intermediate notions suitable for abstraction [31], where some labels are for internal silent steps.

2.2 Probabilistic Labeled Transition Systems

In the seminal work [41], deterministic LTS are extended with probability distributions over discrete state spaces. Instead of giving a set of successor

states with no information about the likelihood among them, probabilistic labeled transition systems (PLTS) change the target of the transition relation, from sets of states to probabilities over states. This quantified uncertainty allows PLTS to represent small step semantics of probabilistic languages like pGCL.

A PLTS modeling a fair coin is given in Figure 2.3.

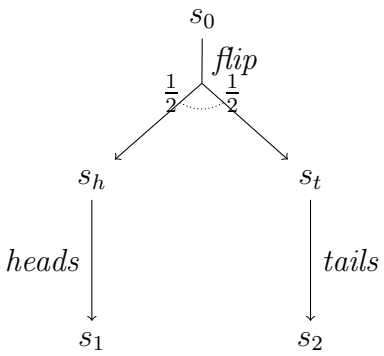


Figure 2.3: A fair coin modeled with a probabilistic labeled transition system.

The formal definition of PLTS is given in the following.

Definition 2.6 (PLTS)**.** A *probabilistic labeled transition system* is a tuple $(S, L, \rightarrow)$ where S is a countable *set of states*, L is a countable set of *labels*, and $\rightarrow$ is a *transition function* $\rightarrow : S \times L \rightarrow \Delta(S)$, where $\rightarrow ((s, a)) = \mu$ is written $s \xrightarrow{a} \mu$, and $\Delta(S)$ is the set of discrete probability distributions over S.

As in the nonprobabilistic case, probabilistic bisimulation is a relation on the state space that equates the stepwise behavior of the system. The idea to generalize bisimulation to probabilistic systems is not new, a similar notion called *lumpability* for Markov chains [38] have been used in Queuing Theory for decades. The key point is how to lift relation R to the target probabilities. We will motivate it using the example shown in Figure 2.4.

The states r_0 and s_0 could be distinguished by an external observer by repeating the experiment of pushing a and counting how many times b or c are enabled. In the long run, the observer will be able to distinguish that the system starting at r_0 is biased towards enabling b, while the other is biased towards c. These two states should not be related in a probabilistic version of bisimilarity. On the contrary, states s_0 and t_0 are not distinguishable by experimentation on the observable events, even though after pressing a, the

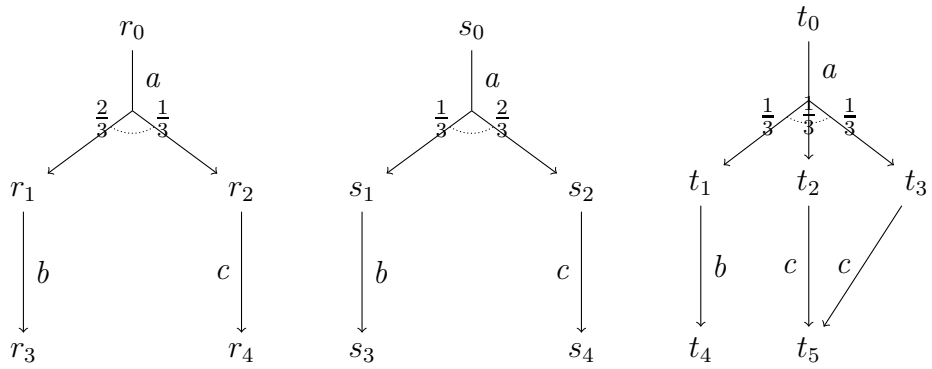


Figure 2.4: Nonbisimilar and bisimilar PLTSs.

rightmost system diverge to t_2, t_3 in a probabilistic way. The accumulated probability of t_2 and t_3 are the same as s_2 and this is the main idea. The probabilities should be equal for all the sets that are closed under the relation R, where a set Q of states is R-closed if there is no state in Q that has an outside R-related state. In symbols:

$$R\text{-closed}(Q) \doteq R(Q) \subseteq Q, \quad \text{where } R(Q) \doteq \{t \mid \exists s \in Q, s \; R \; t\} \tag{2.3}$$

Clearly if we want to capture observational equivalence the set $\{s_2, t_2, t_3\}$ should be R-closed. The relation R can be lifted from states to probabilities using R-closed sets as the events that should quantify the same. We write:

$$\mu \; R \; \mu' \doteq \forall Q, R\text{-closed}(Q), \mu(Q) = \mu'(Q) \tag{2.4}$$

Using (2.4), we can provide a definition for bisimulation on PLTS that resembles the definition of bisimulation for LTS given in (2.2).

Definition 2.7 (Bisimulation for PLTS)**.** A relation $R \subseteq S \times S$ is a *bisimulation* for probabilistic labeled transition system $(S, L, \rightarrow)$ if,

$$s \; R \; t \Rightarrow \forall a \in L, (s \xrightarrow{a}) \; R \; (t \xrightarrow{a})$$

The logic that characterizes bisimulation for PLTS was first given in [41]. The modal operator is augmented with probabilities and typically takes the form $\langle a \rangle_q \phi$, with semantics:

$$[\![\langle a \rangle_q \phi]\!] = \{s \mid s \xrightarrow{a} \mu, q < \mu([\![\phi]\!])\}$$

That is, the set of states such that there is an a transition to a measure that when applied to the event ϕ quantifies greater than the rational q. In [20] it

is shown that negation is not needed and it also covers the continuous state space case. The reason is that two different probability distributions can be distinguished by measuring a particular set[1]. One probability will be strictly less than q, while the other will be strictly greater than q.

Theorem 2.2. *The probabilistic Hennessy-Milner logic* $\phi ::= \top \mid \phi_1 \wedge \phi_2 \mid \langle a \rangle_q \phi$ *characterizes probabilistic bisimulation for PLTS.*

It is worth noting that in [41], the logical characterization included negation.

2.3 Probabilistic Automata

This modeling formalism was introduced by Segala & Lynch [58], and extends PLTS with internal nondeterminism, i.e. a nondeterminism that cannot be resolved externally by composition.

Suppose we have a coin that is not completely specified (subspecified), it can be fair or biased towards heads, and it can be modeled as in Figure 2.5. Even though we conduct an enormous amount of *flip* experiments, the best we can get is that the observed *head*s ratio is greater than 1/2 and lower than 2/3, while the *tail*s ratio is between 1/3 to 1/2.

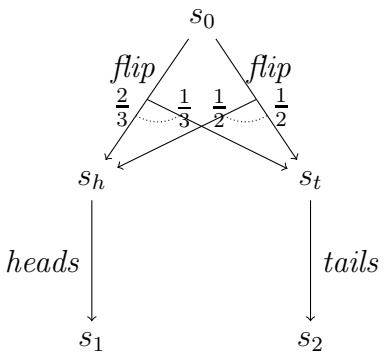


Figure 2.5: Nondeterministic unfair/fair coin modeled with probabilistic automata.

In order to quantify this system, the nondeterminism need to be resolved by a scheduler or policy [54, 64] that chooses randomly among them. The

[1]See Proposition 3.40.

underspecification is captured by the scheduler function that can deterministically choose the fair coin, the unfair, or any intermediate probability. Therefore this discrete system is in fact representing a continuous spectrum of probabilistic choices. This is consistent with other formalism's semantics, like the convex closure requirement for pGCL semantics [46].

We give a formal definition of probabilistic automata.

Definition 2.8 (PA)**.** A *probabilistic automata* is a tuple $(S, L, \to)$, where S is a countable *set of states*, L is a countable set of *labels*, and $\to \subseteq S \times L \times \Delta(S)$ is a countable *transition relation*, where $(s, a, \mu) \in \to$ is written $s \xrightarrow{a} \mu$, and $\Delta(S)$ is the set of discrete probability distributions over S.

Notice that an LTS can be embedded in a PA using Dirac delta probability distributions δ_s, where:

$$\delta_s(\{s'\}) = \begin{cases} 1 & \text{if } s = s' \\ 0 & \text{otherwise} \end{cases}$$

The encoding is such that each $s \xrightarrow{a} s'$ corresponds to $s \xrightarrow{a} \delta_{s'}$ and vice versa.

Bisimulation for PA was defined in [58]. We give its definition using the same expression as before but lifting R twice through (2.4) for probabilities first and (2.1) for nondeterminism afterward. Notice that this twice-lifted relation involves a logic expression having alternating nested quantifiers of depth three, exposing the complexity of the model. The complete expression is:

$$\begin{aligned} &\Big(\forall s \xrightarrow{a} \mu, \exists t \xrightarrow{a} \mu', \forall R\text{-closed}(Q), \mu(Q) = \mu'(Q)\Big) \\ \wedge &\Big(\forall t \xrightarrow{a} \mu', \exists s \xrightarrow{a} \mu, \forall R\text{-closed}(Q), \mu(Q) = \mu'(Q)\Big) \end{aligned}$$

Definition 2.9 (Bisimulation for PA)**.** A relation $R \subseteq S \times S$ is a *bisimulation* for a probabilistic automata $(S, L, \to)$ if,

$$s \; R \; t \Rightarrow \forall a \in L, (s \xrightarrow{a}) \; R \; (t \xrightarrow{a})$$

The logic characterizing bisimulation for PA is relatively new, it was introduced first by [13] and latter by [52] for finite nondeterminism. The continuous nondeterministic case was introduced by us in [18]. We take this later form of presentation since we consider it neater.

Theorem 2.3. *The two-level logic characterizes probabilistic bisimulation for PA. Its syntax is defined:*

$$\begin{aligned} \phi &::= \top \mid \phi_1 \wedge \phi_2 \mid \langle a \rangle \psi \\ \psi &::= \textstyle\bigvee_i \psi_i \mid \neg\psi \mid [\phi]_q, \end{aligned}$$

where the disjunction is denumerable, and q is a positive rational number not greater than one. The first level semantics is given as the set of states that verify the formula:

$$[\![\top]\!] = S \qquad [\![\phi_1 \wedge \phi_2]\!] = [\![\phi_1]\!] \cap [\![\phi_2]\!]$$
$$[\![\langle a \rangle \psi]\!] = \{s \mid s \xrightarrow{a} \mu, \mu \in [\![\psi]\!]\}$$

while the second level encodes sets of probabilities:

$$[\![\textstyle\bigvee_i \psi_i]\!] = \bigcup_i [\![\psi_i]\!] \qquad [\![\neg\psi]\!] = [\![\psi]\!]^{\mathsf{c}}$$
$$[\![[\phi]_q]\!] = \{\mu \mid q < \mu([\![\phi]\!])\}$$

The nondeterminism in Figure 2.5 can be resolved by an external agent usually called scheduler or policy. The scheduler defines a probability distribution of the possible labels and probabilities coming out of a state s. This probabilistic choice could depend not only on the current state, but also on previous history of states, labels and probabilities.

Definition 2.10 (Scheduler for PA)**.** A *scheduler* for probabilistic automata $(S, L, \rightarrow)$ is a function η from finite path traces to distributions over labels and probabilities,

$$\eta : S \times (L \times \Delta(S) \times S)^\star \rightarrow \Delta(L \times \Delta(S))$$

where for every *finite path trace* $\alpha = s_1\ a_2\mu_2 s_2\ \ldots\ a_n\mu_n s_n \in S \times (L \times \Delta(S) \times S)^\star$, the *last state* is denoted by $last(\alpha) = s_n$, and the probability of $\eta(\alpha)$ is *concentrated* on $\{(a, \mu) \mid last(\alpha) \xrightarrow{a} \mu\}$, (i.e. $\eta(\alpha)(a, \mu) = 0$ if $last(\alpha) \not\xrightarrow{a} \mu$) the set of outgoing labels and probabilities of the path last state.

The scheduler can be combined with the PA, resolving the nondeterminism and giving rise to a purely probabilistic system.

Definition 2.11 (Combined transition for PA)**.** Given PA $(S, L, \rightarrow)$ and scheduler η, the *combined transition* is a function:

$$\mu_{\eta(\cdot)} : S \times (L \times \Delta(S) \times S)^\star \rightarrow \Delta(L \times \Delta(S) \times S)$$

defined as follows:

$$\mu_{\eta(\alpha)}(a, \mu, s) = \mu(s) \cdot \eta(\alpha)(a, \mu)$$

Notice that if a or μ do not belong to the outgoing transitions, by definition of the scheduler, its probability is zero. Given an initial state distribution, and a scheduler, the trace semantics for PA can be readily given as a set of probabilistic words, or more precisely by the set of trace probabilities.

Definition 2.12 (Trace probability for PA)**.** Given a PA $(S, L, \to)$, a scheduler η on it and an *initial probability over states* ν, the *trace probability* is defined as follows:

$$P_{\nu,\eta}(s_1) = \nu(s_1)$$
$$P_{\nu,\eta}(\alpha\ a_n\mu_n s_n) = P_{\nu,\eta}(\alpha) \cdot \mu_{\eta(\alpha)}(a_n\mu_n s_n)$$

2.4 Concluding Remarks

We gave a summary of labeled transition systems including nondeterministic and probabilistic choices over discrete state space.

If we move to continuous states, labels or nondeterminism, there are known mathematical artifacts that could endanger the soundness of the definition. Namely there are subsets that are not measurable (Vitali sets), and this implies that core definitions like the semantics of the modal operator or the scheduler cannot be given.

Also this uncountability leads to another problem. The quantification of uncountable sets usually implies that single elements like in Definition 2.11 and 2.12, are assigned zero probability.

The mathematical structure that is tailored to the quantification of uncountable sets is the measure space [2], This structure is studied in the mathematical field of Measure Theory, the main topic of the next chapter.

Chapter 3

Measure Theory

This chapter presents a revision of Measure Theory and Descriptive Set Theory. It contains what we consider a good selection, order and emphasis of these mathematical fields, where the rest of the work is rooted. The material is collected from [1, 2, 4, 9, 36]. Since none of the results are new, we do not exhaustively recreate the proofs unless we consider it convenient.

The chapter is organized as follows. In Section 3.1, we introduce σ-algebras and present three proof strategies to show properties on measurable sets. In Section 3.2 we show how to build σ-algebras from families of functions, and use this to define product σ-algebras and function space σ-algebras. In Section 3.3, we introduce the notion of measure, and in Section 3.4, we introduce Lebesgue integration for general measures. We also show how to integrate new measures. We devoted Section 3.5 to the σ-algebra of measures $\Delta(\Sigma)$, a key mathematical object for the rest of the work. The chapter ends in Section 3.6 with standard tools to endow the product σ-algebra with a measure.

Most of the set theoretic operations in this work are denumerable, therefore, to simplify the notation we write $\bigcup_i A_i$ instead of $\bigcup_{i\in\mathbb{N}} A_i$. Usually we will omit all references to the index set $\mathbb{N}$ as in the denumerable family $\{A_i\}_i$, or the denumerable sequence of functions $(f_i)_i$. Arbitrary sets are denoted by I, so an arbitrary intersection is $\bigcap_{i\in I} A_i$. A disjoint union $\biguplus_i A_i$ denotes the union of the pair-wise disjoint family $\{A_i\}_i$. The powerset of a set S is denoted by $2^S \doteq \{A \mid A \subseteq S\}$. Given a set A the characteristic function χ_A is defined by $\chi_A(s) = 1$ if $s \in A, 0$ otherwise. We will use point-free operations on functions like $f = \chi_A + \chi_B$, meaning point-to-point equality $f(x) = \chi_A(x) + \chi_B(x)$.

We suggest the reader that is acquainted with Measure Theory to quickly go through this chapter, in order to get used to our notation and the main

results we need in the forthcoming chapters. For other summaries of Measure Theory related to Concurrency Theory in Computer Science, please refer to [25, 49, 50]

3.1 σ-algebra

Measure theory deals with events and their quantification. It is a generalization of common concepts like length, area and probability. The events are the *measurable* objects (segments) quantified by a *measure* (length). Given a *base set* S we want to define a *family of events* Σ that are measurable by the *measure set function* $\mu : \Sigma \to \mathbb{R}^+$. We think of the events as possible outcomes of an experiment we want to quantify with the measure. We formalize the family of events.

Definition 3.1 (σ-algebra)**.** A *σ-algebra* Σ on a set S is a nonempty family of subsets of S such that it is closed under complement and denumerable union:

$$A \in \Sigma \Rightarrow A^{\mathsf{c}} \in \Sigma \tag{3.1}$$

$$\{A_i\}_i \subseteq \Sigma \Rightarrow \bigcup_i A_i \in \Sigma \tag{3.2}$$

The elements $A \in \Sigma$ are called *measurable sets*, and the pair (S, Σ) is a *measurable space*.

Note that σ-algebras are also closed under finite union, as well as countable and finite intersection. Given that the family Σ is nonempty, by closure properties it always includes $\emptyset$ and S. The smallest σ-algebra is $\{\emptyset, S\}$, and the largest is the powerset 2^S. There are cases where 2^S is too fine for our measuring purposes (think of a dice used as a coin), or it simply introduces mathematical artifacts (like non-measurable sets). For discrete sets S, the usual measurable space is $(S, 2^S)$.

Definition 3.1 captures what is expected for events later to be measured. If an event is measurable, its non-occurrence should also be measurable (3.1); if a sequence of events is measurable, its aggregation should also be (3.2). Perhaps what is not intuitive is why the union closure needs to hold for the *denumerable* case, instead the more natural finite one. Citing [2]:

> Closure under countable union and intersection is difficult to justify physically, and perhaps the most convincing reason for requiring it is that a richer mathematical theory is obtained. Specifically, we are able to assert that the limit of a sequence of events is an event.

If the structure imposed on the events only count for *finite* union, the family is called an algebra of sets.

Definition 3.2 (Algebra of sets)**.** An *algebra* Γ of a set S is a nonempty family of subsets of S such that it is closed under complement and finite union:

$$A \in \Gamma \Rightarrow A^{\mathsf{c}} \in \Gamma$$
$$\{A_i\}_{i=1}^n \subseteq \Gamma \Rightarrow \bigcup_{i=1}^n A_i \in \Gamma$$

We obtain an equivalent definition if we replace closure under finite union with closure under binary union.

Example 3.3. It can be checked that the family Γ given by all the finite disjoint unions of left-closed, right-open rational intervals $[p, q)$ forms an algebra. Notice that even though $[0, 1 + 1/i) \in \Gamma$ for all i, its denumerable intersection $\bigcap_i [0, 1 + 1/i) = [0, 1]$ does not belong to Γ.

Previous example shows the kind of denumerable closure properties that do not hold for algebras and are essential in the development of Measure Theory. The following example from [4] shows a non-trivial σ-algebra.

Example 3.4. Let Σ be a family of subsets of S that are either countable or its complement is countable (i.e. it is cocountable). Notice that if S is uncountable, then there is a set X such that both X and its complement are uncountable hence $X \notin \Sigma$. For $S = \mathbb{R}$ this set could be $X = (1/2, 1]$. Note that $X = \bigcup_{1/2<x\leq 1} \{x\}$, which shows that a σ-algebra might not be closed under *arbitrary* union.

This example also gives a reason why σ-algebras are *not closed under arbitrary union*. It would trivialize every σ-algebra having measurable singletons. The intersection of an arbitrary family of σ-algebras results in a new σ-algebra. This calls for the idea of minimal σ-algebra containing a given family, the set of *generators*.

Definition 3.5 (Generated σ-algebra)**.** Given the generator family $\mathcal{A} \subseteq 2^S$, the *generated σ-algebra* $\sigma(\mathcal{A})$ is the intersection of all σ-algebras containing $\mathcal{A}$, that is:

$$\sigma(\mathcal{A}) = \bigcap \{\Sigma \subseteq 2^S \mid \Sigma \supseteq \mathcal{A}\}$$

It can be seen that this generated σ-algebra contains $\mathcal{A}$ and is minimal. Now some basic facts about generated σ-algebras, that although simple, represent useful tools.

Proposition 3.1. *i. If $\mathcal{A} \subseteq \mathcal{A}'$, then $\sigma(\mathcal{A}) \subseteq \sigma(\mathcal{A}')$.*

ii. If $\mathcal{A} \subseteq \mathcal{A}' \subseteq \sigma(\mathcal{A})$, then $\sigma(\mathcal{A}) = \sigma(\mathcal{A}')$.

A well-know generated σ-algebra is the Borel σ-algebra.

Definition 3.6 (Borel σ-algebra)**.** The *Borel σ-algebra* on $\mathbb{R}$ is the σ-algebra generated by the left-closed, right-open rational intervals, that is $\mathcal{B}(\mathbb{R}) = \sigma(\{[p, q) \mid p, q \in \mathbb{Q}\})$.

This definition is one of many possible candidates since any kind of rational interval would define the same set. Take for example these two generator sets, $\mathcal{A} = \{(p, q) \mid p, q \in \mathbb{Q}\}$, $\mathcal{A}' = \{[p, q) \mid p, q \in \mathbb{Q}\}$. The equalities $(p, q) = \bigcup_{p' \in \mathbb{Q}, p < p'} [p', q)$, and $[p, q) = \bigcap_{p' \in \mathbb{Q}, p' < p} (p', q)$, imply that $\mathcal{A} \subseteq \sigma(\mathcal{A}')$ and $\mathcal{A}' \subseteq \sigma(\mathcal{A})$. By idempotence of the σ operator, $\mathcal{A}$ and $\mathcal{A}'$ generate the same σ-algebra. We also prove that Borel measurable sets generated by real-valued intervals are equivalent to the Borel measurable sets generated by rational-valued intervals.

Proposition 3.2. *The Borel σ-algebra over the reals are either generated by rational or real endpoint intervals.*

Proof. Let $\mathcal{A} = \{[p, q) \mid p, q \in \mathbb{Q}\}$ and $\mathcal{A}' = \{[a, b) \mid a, b \in \mathbb{R}\}$, then $\mathcal{A} \subset \mathcal{A}'$. Using equalities $[a, q) = \bigcap\{(p, q) \mid p < a, p \in \mathbb{Q}\}$, $[a, b) = \bigcup\{[a, q) \mid q < b, q \in \mathbb{Q}\}$ is clear that $\mathcal{A}' \subset \sigma(\mathcal{A})$. By Proposition 3.1, $\sigma(\mathcal{A}) = \sigma(\mathcal{A}') = \mathcal{B}(\mathbb{R})$. Observe that both set operations are denumerable. □

Given that $\mathbb{Q}$, and therefore $\mathbb{Q} \times \mathbb{Q}$ are countable sets, the Borel σ-algebra is generated by a denumerable family.

Definition 3.7 (Countably generated σ-algebra)**.** A σ-algebra Σ is *countably generated* or *separable* if it is generated by a countable class of sets.

A σ-algebra that is countably generated deserves a special status since it is the main ingredient of many useful results throughout this thesis. Although trivial from certain perspective, the next proposition proves useful to clean out hypothesis of theorems related to countably generated σ-algebras.

Proposition 3.3. *Given a countable family $\mathcal{C}$, the closure of $\mathcal{C}$ with respect to binary intersection, given by $\{\bigcap_{i=1}^{n} A_i \mid \{A_i\}_{i=1}^{n} \subseteq \mathcal{C}, n \in \mathbb{N}\}$, is also countable.*

Proof. The set of finite subsets of a countable set is also countable [1], and thus the result holds. □

A simple corollary is that closing a countable set by binary/finite unions and complements (i.e. obtaining an algebra out of a countable set) remains countable.

It is relevant to notice that open and closed sets in the usual topology of the reals are measurable. We include the proof given the unusual way it uses union of a denumerable number of sets. Sometimes this result is taken as definition of the Borel σ-algebra.

Proposition 3.4. *The σ-algebra $\mathcal{B}(\mathbb{R})$ contains all open and closed sets.*

Proof. Let G be open in $\mathbb{R}$ and $x \in G$, then there exists an open interval with rational endpoints (p_x, q_x) such that $x \in (p_x, q_x) \subseteq G$. Then $G = \bigcup_{x \in G}(p_x, q_x)$. Since there are countably many open intervals with rational endpoints, we conclude G is measurable. For a closed set F it also follows, since it is the complement of an open set. □

Another important class of σ-algebras are the σ-algebras that can separate points.

Definition 3.8. A measurable space (S, Σ) *separates points* if for all $s, s' \in S$ such that $s \neq s'$, there is a measurable set $A \in \Sigma$ with $s \notin A \ni s'$.

Note that if all singletons are measurable ($\forall s \in S, \{s\} \in \Sigma$), Σ separates points. The following proposition gives more conditions for σ-algebras that separates points.

Proposition 3.5. *Given $\mathcal{A} \subseteq 2^S$, $\mathcal{A}$ separates points iff $\sigma(\mathcal{A})$ does.*

Proof. The if part is trivial since generators are included in the generated σ-algebra. For the other implication suppose towards a contradiction that given s, s', there is a $Q \in \sigma(\mathcal{A})$ separating them $s \notin Q \ni s'$, but none of the generators can separate, i.e. $\forall A \in \mathcal{A}, s \in A \Leftrightarrow s' \in A$. Being the last property stable by complements and denumerable unions, we conclude $s \in Q \Leftrightarrow s' \in Q$ contradicting the assumption. □

The following σ-algebra is countably generated but it does not separate points nor its singletons are measurable.

Example 3.9 (**Q-coQ**)**.** The σ-algebra **Q-coQ** $\doteq 2^{\mathbb{Q}} \cup \{\mathbb{R} \setminus Q \mid Q \in 2^{\mathbb{Q}}\}$ is generated by the denumerable family $\{\{q\} \mid q \in \mathbb{Q}\} \uplus \{\emptyset\}$. Notice that **Q-coQ** cannot separate one irrational from another (let alone asking for all singletons being measurable).

On the contrary the Borel σ-algebra has many pleasant properties.

Example 3.10. The Borel σ-algebra $\mathcal{B}(\mathbb{R})$ is countably generated, separates points and has measurable singletons. All these properties are consequence of the density of the rationals over the reals.

Given a σ-algebra, smaller ones can be defined.

Definition 3.11 (Sub-σ-algebra)**.** Given σ-algebra Σ, the σ-algebra Λ is a *sub-σ-algebra* of Σ if $\Lambda \subseteq \Sigma$.

Given a σ-algebra and an arbitrary set, we can define a new σ-algebra.

Definition 3.12 (Relative σ-algebra)**.** Given the measurable space (S, Σ), and the (arbitrary) set $X \subseteq S$, then the *relative σ-algebra* is defined by $\Sigma|X = \{A \cap X \mid A \in \Sigma\}$. If $\Sigma = \sigma(\mathcal{A})$ then $\Sigma|X = \sigma(\mathcal{A}|X)$. Moreover, if $X \in \Sigma$, then $\Sigma|X = \{A \in \Sigma \mid A \subseteq X\}$.

Using previous definition we can write $\mathcal{B}([0,1])$ for the relative σ-algebra $\mathcal{B}(\mathbb{R})|[0,1] = \{A \in \mathcal{B}(\mathbb{R}) \mid A \subseteq [0,1]\}$, that is the Borel σ-algebra generated by all rational endpoints intervals in $[0,1]$.

We now give the first closure result for σ-algebras, namely that it includes the limits of events, a property that algebras lack (see Example 3.3).

Definition 3.13 (Increasing/decreasing sequence limits)**.** Given an increasing (resp. decreasing) sequence $(A_i)_i$, A is the limit of A_i denoted $A_i \nearrow A$ (resp. $A_i \searrow A$), if $A = \bigcup_i A_i$ (resp. $A = \bigcap_i A_i$).

A family of sets that is closed under limits is said to be monotone.

Definition 3.14 (Monotone family)**.** A family $\mathcal{A} \subseteq 2^S$ is *monotone* if $\{A_i\}_i \subseteq \mathcal{A}$, then if either $A_i \nearrow A$ or $A_i \searrow A$, then $A \in \mathcal{A}$.

Proposition 3.6. *A σ-algebra is a monotone family.*

The task of proving that properties holding on generators extend to the whole σ-algebra is not direct. Special proof strategies exist. Now we describe three different techniques used to show that a σ-algebra inherits properties of its generators.

Good Sets Principle. The first technique requires no assumptions on the family of generators $\mathcal{A}$, but it needs the strongest property of the three techniques for the so-called *good sets* $\mathcal{G}$.

Proposition 3.7 (Good sets principle)**.** *Let $\mathcal{A}$ be the family of generators and $\mathcal{G}$ the family of good sets. If $\mathcal{A} \subseteq \mathcal{G}$ and $\mathcal{G}$ is a σ-algebra, then $\sigma(\mathcal{A}) \subseteq \mathcal{G}$. Namely, if the* good sets *form a σ-algebra, and it contains the generators, then the generated σ-algebra consists of good sets.*

Example 3.15. Let B be a Borel measurable set in $\mathcal{B}(\mathbb{R})$, then $a + B \doteq \{a + x \mid x \in B\}$ is also Borel measurable. That is, Borel measurable sets are measurable-invariant under translations.

The proof is as follows. Let the *good sets* family $\mathcal{G} = \{B \mid a + B \in \mathcal{B}(\mathbb{R})\}$. The rational endpoints intervals $\mathcal{A} = \{[p, q) \mid p, q \in \mathbb{Q}\}$ are good $\mathcal{A} \subseteq \mathcal{G}$. We need to show that $\mathcal{G}$ is a σ-algebra. $\mathcal{G}$ is nonempty: clearly $\emptyset \in \mathcal{G}$. Notice that $\mathcal{G}$ is closed under complements since $\{a+x \mid x \in B\}^{\mathsf{c}} = \{a+x \mid x \in B^{\mathsf{c}}\}$, and it is closed under denumerable unions since $\{a + x \mid x \in \bigcup_i B_i\} = \bigcup_i \{a + x \mid x \in B_i\}$. By the good sets principle the property follows.

We continue using $\mathcal{G}$ to denote the family of good sets in the next two techniques.

Dynkin's π-λ Lemma. The second tool trades a more restrictive property of the family of generators, it needs it to be a π-system, for a less restrictive condition on the family of good sets, it has to be a λ-system. First we define these two new properties of families of sets.

Definition 3.16 (π-system)**.** A nonempty family $\mathcal{P} \subseteq 2^S$ is a *π-system* if it is closed under binary (finite) intersection, that is,

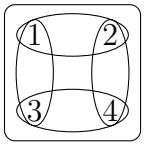
$$X_1, X_2 \in \mathcal{P} \Rightarrow X_1 \cap X_2 \in \mathcal{P}$$

Definition 3.17 (λ-system)**.** A nonempty family $\mathcal{L} \subseteq 2^S$ is a *λ-system* if it is closed under complements and countable *disjoint* union, that is:

$$X \in \mathcal{L} \Rightarrow X^{\mathsf{c}} \in \mathcal{L}$$
$$\{X_i\}_i \subseteq \mathcal{L} \Rightarrow \biguplus_i X_i \in \mathcal{L}$$

A λ-system is not a stronger notion than a σ-algebra, since every σ-algebra is also a λ-system. The inclusion is strict.

Example 3.18. The empty set together with the family $\mathcal{L}$ given by the figure is a λ-system.



However the union of $\{1, 2\}$ and $\{1, 3\}$ is not in $\mathcal{L}$. This is the smallest model that shows proper inclusion of σ-algebras in λ-systems.

Notice, however, that $\mathcal{L}$ is not a π-system, and this is due to the next lemma.

Lemma 3.8. *If $\mathcal{A}$ is a π-system and a λ-system, then $\mathcal{A}$ is a σ-algebra. The converse is direct.*

The next lemma is due to Eugene Dynkin, and it is the core of the second technique.

Lemma 3.9 (Dynkin's π-λ)**.** *If $\mathcal{A}$ is a π-system and $\mathcal{G}$ is a λ-system, then $\mathcal{A} \subseteq \mathcal{G}$ implies $\sigma(\mathcal{A}) \subseteq \mathcal{G}$.*

The usage example, which is canonical in Measure Theory, is the uniqueness of measures and we present it in Theorem 3.20.

Monotone Family Theorem. The third technique is the next theorem that asks the strongest premise over the generators (they should form an algebra) and the weakest over the good sets, they only need to be a monotone family.

Theorem 3.10 (Monotone family)**.** *Let $\mathcal{A}$ be an algebra and $\mathcal{G}$ be a monotone family, then $\mathcal{A} \subseteq \mathcal{G}$ implies $\sigma(\mathcal{A}) \subseteq \mathcal{G}$.*

The typical usage of Monotone Family Theorem is in the proof of Carathéodory extension Theorem 3.21 and Fubini Theorem 3.44.

The importance of these techniques resides in the fact that there are Borel sets that cannot be arrived from the intervals by any finite or countable sequence of set operations, where each of these operations are finite or countable [4, p.26]. This result implicitly disregards natural induction as a tool. Table 3.1 presents a summary of the techniques.

Name	Generators	Good sets
Good Sets Principle	any	σ-algebra
Dynkin's Lemma	π-system	λ-system
Monotone Family Theorem	algebra	monotone family

Table 3.1: The three proof techniques for generated σ-algebras ordered by the strength of the hypothesis over the generators.

3.2 Measurable Functions

Functions play an important role in measure theory. The key property is that the inverse of a function $f : S \to S'$ preserves complements and arbitrary (in particular countable) unions,

$$f^{-1}(X^{\mathsf{c}}) = (f^{-1}(X))^{\mathsf{c}}$$
$$f^{-1}(\textstyle\bigcup_i X_i) = \bigcup_i f^{-1}(X_i)$$

that is $f^{-1} : 2^{S'} \to 2^S$ preserves all set operations in S'. A function whose inverse preserves measurable sets is called measurable.

Definition 3.19 (Measurable function)**.** Given a function $f : S \to S'$ and measurable spaces (S, Σ), (S', Σ'), the *function is measurable* if $f^{-1}(A') \in \Sigma$ for all $A' \in \Sigma'$. In this case we write $f : (S, \Sigma) \to (S', \Sigma')$.

If the domain and image of a function is the same measurable space (S, Σ), then f is measurable if Σ is closed under f^{-1}. Since denumerable union and complement commute with the inverse of f, we have the following proposition.

Proposition 3.11. *Let (S, Σ) and (S', Σ') be two measurable spaces and $\mathcal{A}' \subseteq 2^{S'}$ such that $\Sigma' = \sigma(\mathcal{A}')$, then $f : S \to S'$ is measurable iff for all $A' \in \mathcal{A}'$, $f^{-1}(A') \in \Sigma$.*

Now we show how to construct a σ-algebra on a domain set S from functions $S \to S_i$, where each S_i has a σ-algebra attached [36].

Definition 3.20 (σ-algebra generated by a family of functions)**.** Given set S, an (arbitrary) family of measurable spaces $((S_i, \Sigma_i))_{i \in I}$, and functions $f_i : S \to S_i$, the *σ-algebra generated by $(f_i)_{i \in I}$*, is defined by the family of generators $\{f_i^{-1}(A_i) \mid A_i \in \Sigma_i, i \in I\}$. If Σ_i is generated by $\mathcal{A}_i$, then $\{f_i^{-1}(A_i) \mid A_i \in \mathcal{A}_i, i \in I\}$ also generates Σ.

This σ-algebra Σ is the smallest σ-algebra on S that makes all maps $f_i : (S, \Sigma) \to (S_i, \Sigma_i)$ measurable. Using this definition we now show how to construct finite and denumerable product σ-algebras, and their generalization: σ-algebras over function spaces. Later on, using σ-algebras over function spaces, we will construct $\Delta(\Sigma)$, the σ-algebra of measures that has a prominent role in this work.

Products. If a σ-algebra represents the possible outcomes of an experiment, the product σ-algebra captures the possible outcomes of *repeating an experiment* a finite or denumerable number of times. In order to generate a *product σ-algebra* on $\prod_{i\in I} S_i$, we use projections $\pi_j : \prod_{i\in I} S_i \to S_j$ as the generating function family, so that, by construction, all projections are measurable functions.

Definition 3.21 (General product σ-algebra)**.** The *product σ-algebra* on measurable spaces $((S_i, \Sigma_i))_{i\in I}$ is generated by the sets

$$\{\pi_i^{-1}(A_i) \mid A_i \in \Sigma_i, i \in I\}$$

Observe that the generators are exactly $\prod_{i\in I} A_i$ where for all $i \in I$, $A_i \in \Sigma_i$, and $|\{A_i \mid A_i \neq S_i, i \in I\}| \leq 1$. The definition is equivalent if we allow $|\{A_i \mid A_i \neq S_i, i \in I\}| \leq n$ thanks to Proposition 3.1.

The standard definitions of finite product σ-algebra given in the literature can be straightforwardly deduced from Definition 3.21.

Proposition 3.12 (Finite product σ-algebra)**.** *Given measurable spaces $((S_i, \Sigma_i))_{i=1}^n$, the* finite product σ-algebra *denoted by $\otimes_{i=1}^n \Sigma_i$ is generated by the* measurable rectangles $\prod_{i=1}^n A_i$ *with $A_i \in \Sigma_i$.*

Proposition 3.13 (Denumerable product σ-algebra)**.** *Given measurable spaces $((S_i, \Sigma_i))_i$, the* denumerable product σ-algebra *denoted by $\bigotimes_i \Sigma_i$ is generated by the* measurable rectangles $(\prod_{i=1}^n A_i) \times \left(\prod_{n<i} S_i\right)$ *with $A_i \in \Sigma_i$, also called* measurable rectangles with base $A_1, \ldots, A_n$.

Proposition 3.12 and Proposition 3.13 follows by the observation below Definition 3.21. Some texts give the definition of denumerable product σ-algebra using *cylinders*.

Proposition 3.14 (Denumerable product σ-algebra using cylinders)**.** *Given measurable spaces $((S_i, \Sigma_i))_i$, the* denumerable product σ-algebra *denoted by $\bigotimes_i \Sigma_i$ is generated by B_n, the* cylinders with base B^n, *where $B^n \in \bigotimes_{i=1}^n \Sigma_i$ is a measurable set in the n-dimensional product space, and $B_n \doteq B^n \times \left(\prod_{n<i} S_i\right)$.*

Proof. Note that $\mathcal{A} = \{(\prod_{i=1}^{n} A_i) \times \left(\prod_{n<i} S_i\right)\} \subseteq \mathcal{A}' = \{B^n \times \left(\prod_{n<i} S_i\right)\}$, and by Proposition 3.1, $\sigma(\mathcal{A}) \subseteq \sigma(\mathcal{A}')$. The other inclusion, $\mathcal{A}' \subseteq \sigma(\mathcal{A})$, follows using the good sets principle. □

Product σ-algebras of the same measurable space (S, Σ) are denoted by $\Sigma^n \doteq \bigotimes_{i=1}^{n} \Sigma$ for the finite case, and $\Sigma^\omega \doteq \bigotimes_i \Sigma$ for the countable case. The property of being countably generated is inherited from its components.

Proposition 3.15. *If Σ is countably generated, then Σ^n and Σ^ω are also countably generated.*

If we observe the natural isomorphism between $S_1 \times (S_2 \times S_3)$ and $(S_1 \times S_2) \times S_3$, then the product operator on σ-algebras is associative, parenthesis can be omitted and we can write $(\Sigma_1 \otimes \Sigma_2) \otimes \Sigma_3 = \Sigma_1 \otimes (\Sigma_2 \otimes \Sigma_3) = \Sigma_1 \otimes \Sigma_2 \otimes \Sigma_3$.

A simple but useful result regarding product space Borel σ-algebras is a direct consequence of Proposition 3.1.

Proposition 3.16. $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^2)$.

This can be generalized to k-dimensional products, and as a consequence $\mathcal{B}(\mathbb{R})^k$ is generated by the rational endpoint rectangles $\prod_{i=1}^{k}[p_i, q_i]$.

Functions. Observe that there is an isomorphism between tuples in S^n and functions from $[1..n] \to S$. If $(s_1, \ldots, s_n) \in S^n$, we define $f : [1..n] \to S$ by $f(i) = s_i$ for all $1 \leq i \leq n$. Conversely given a function $f : [1..n] \to S$, the related tuple is $(f(1), \ldots, f(n))$. The projections in the tuple space play the role of function evaluation in the function space.

In the case of a product measurable space (S^n, Σ^n), the σ-algebra constructed in Definition 3.21 guarantees that projections or *evaluations* $(\lambda f : f(i)) : ([1..n] \to S) \to S$ *are measurable for every i.* The same argument is valid for $\mathbb{N}$ used as index set (functions $\mathbb{N} \to S$), $\mathbb{R}$ used as index set (functions $\mathbb{R} \to S$), or even arbitrary domains I (functions $I \to S$). Therefore definitions of *function space σ-algebra* are also a particular case of Definition 3.21.

Definition 3.22 (Function space σ-algebra)**.** Given a measurable space (S, Σ) and function space $I \to S$ (or equivalently S^I) the *function space σ-algebra* is the one generated by $f^A(i) = (\lambda f : f(i))^{-1}(A)$, with $A \in \Sigma$ and $i \in I$. If $\Sigma = \sigma(\mathcal{A})$ then $\{f^A(i) \mid A \in \mathcal{A}, i \in I\}$ also generates the σ-algebra.

Notice that a generator $f^A(i) = \{f \in I \to S \mid f(i) \in A\}$ contains all the functions such that their value at i belongs to the measurable set A.

There is an operator over binary products of measurable spaces that enjoy pleasant properties. It *slices* a measurable set in a binary product σ-algebra at any given value of the first (second) measurable space, giving a measurable set in the second (first) measurable space.

Definition 3.23. Given a measurable set A in the product measurable space $(S_1 \times S_2, \Sigma_1 \otimes \Sigma_2)$, the *section of A at s_1* is $A_{|s_1} = \{s_2 \in S_2 \mid (s_1, s_2) \in A\}$.

Proposition 3.17. *Given a product measurable space $(S_1 \times S_2, \Sigma_1 \otimes \Sigma_2)$, for all $A \in \Sigma_1 \otimes \Sigma_2$, $s_1 \in S_1$, it is $A_{|s_1} \in \Sigma_2$.*

Proof. It follows using Proposition 3.7 and good sets $\mathcal{G} = \{A \mid A_{|s_1} \in \Sigma\}$. □

We point out that previous definition and proposition can be generalized to finite and denumerable product spaces. The converse of Proposition 3.17 is not valid.

Example 3.24. Let $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be the Borel measurable space. We define the subset $Vd = \{(x, x) \mid x \in V\} \subseteq \mathbb{R}^2$, where V is a nonmeasurable Vitali set[1]. Every section $Vd_{|x}$ in both coordinates are either $\{x\}$ or $\emptyset$, therefore measurable. However Vd is not measurable in $\mathcal{B}(\mathbb{R}^2)$. Let $d(x) = (x, x)$ be the diagonal function. The function d is measurable since for every rational-endpoint square, the inverse image of d is a closed interval, $d^{-1}([p_1, q_1] \times [p_2, q_2]) = \{x \mid (x, x) \in [p_1, q_1] \times [p_2, q_2]\} \in \mathcal{B}(\mathbb{R})$. If we suppose Vd measurable, using the measurable d function, $d^{-1}(Vd) = V$ would be measurable, contradicting the nonmeasurability of V.

Notice this is *not a projection*, it is a slice of a bidimensional set. We remark there is not a standard notation to indicate the coordinate where the sectioning is being done. It will be explicitly stated if needed.

Sums. Although it is not a particular case of Definition 3.20, we complete this section with sum σ-algebras. Instead of building the product of measurable spaces, it is possible to obtain the sum σ-algebra by taking the disjoint union of measurable sets.

Definition 3.25. Given measurable disjoint spaces $(S_i, \Sigma_i)_{i \in I}$, the *sum measurable space* $(\bigoplus_{i \in I} S_i, \bigoplus_{i \in I} \Sigma_i)$ is defined by the disjoint union of the spaces $\bigoplus_{i \in I} S_i \doteq \biguplus_{i \in I} S_i$, and the σ-algebras $\bigoplus_{i \in I} \Sigma_i \doteq \biguplus_{i \in I} \Sigma_i$.

Observe that if the sets are not disjoint we can replace them by an isomorphic disjoint copy, and use the above definition.

[1] $V \in 2^{\mathbb{R}}$, but $V \notin \mathcal{B}(\mathbb{R})$. It will be defined later on.

3.3 Measures

Events defined by a σ-algebra can be quantified or *measured*. A measure is the generalization of many quantifying functions like length, area, volume, cost, energy, probability, etc.

Definition 3.26 (Measure)**.** A *measure* on the measurable space (S, Σ) is a function $\mu : \Sigma \to \mathbb{R}^+ \cup \{\infty\}$ such that it is *strict* and σ*-additive* for pair-wise disjoint measurable sets $\{A_i\}_i$, that is:

$$\mu(\emptyset) = 0$$
$$\mu(\biguplus_i A_i) = \sum_i \mu(A_i)$$

A measure is called σ*-finite* if for some $\{A_i\}_i \subseteq \Sigma$, $S = \biguplus_i A_i$ and $\mu(A_i) < \infty$ for all i. It is *finite* if $\mu(S) < \infty$. If $\mu(S) = 1$ then it is a *probability measure*, if $\mu(S) \leq 1$ it is called *subprobability measure*. The triple (S, Σ, μ) is a *measure space*.

The difference between general, σ-finite, finite, probability and subprobability measures is not minor. Many results strongly depend on the type of bounding that the measure has. We will emphasize this fact whenever needed. Notice we could have completely avoided this problem living in the setting of probability measures, where all properties hold. We rather pay the price of generality where it can be obtained in order to capture not only measures of probability but also other interesting measures like subprobabilities and measures similar to length, volume and cost.

Example 3.27 (Counting measure)**.** Given measurable space $(S, 2^S)$, the *counting measure* μ is defined by $\mu(A) = |A|$ for finite sets, otherwise $\mu(A) = \infty$. Clearly it is $\mu(\emptyset) = 0$ and it is σ-additive. If S is uncountable, countable or finite then μ is a general measure, a σ-finite measure, or a finite measure respectively.

Measures enjoy monotonicity and limit-preserving properties.

Theorem 3.18 (Monotonicity and continuity of measures)**.** *Let μ be a measure on (S, Σ), and a sequence $(A_i)_i$ of measurable sets, then:*

i. Inclusion-Exclusion principle. $\mu(A_1 \cup A_2) + \mu(A_1 \cap A_2) = \mu(A_1) + \mu(A_2)$.

ii. Monotonicity. $A_1 \subseteq A_2 \Rightarrow \mu(A_1) \leq \mu(A_2)$.

iii. Countable subadditivity. $\mu(\bigcup_i A_i) \leq \sum_i \mu(A_i)$.

iv. Continuity from below. If $A_i \nearrow A$, *then* $\mu(A_i) \nearrow \mu(A)$.

v. *Continuity from above. If* $A_i \searrow A$ *and* $\mu(A_1) < \infty$*, then* $\mu(A_i) \searrow \mu(A)$.

A measurable function induces measures in the target σ-algebra from a measure in the source σ-algebra.

Proposition 3.19. *Let* (S_1, Σ_1) *and* (S_2, Σ_2) *be two measurable spaces. Let* $f : (S_1, \Sigma_1) \to (S_2, \Sigma_2)$ *be a measurable function. If* μ_1 *is a measure on* (S_1, Σ_1)*, then* $\mu_2 \doteq \mu_1 \circ f^{-1}$ *is an* induced measure *on* (S_2, Σ_2).

Proof. The defined measure μ_2 is strict since $\mu_2(\emptyset) = \mu_1(f^{-1}(\emptyset)) = \mu_1(\emptyset) = 0$. For σ-additivity we calculate with $\{A_i\}_i \subseteq \Sigma_2$: $\mu_2(\biguplus_i A_i) = \mu_1(f^{-1}(\biguplus_i A_i)) = \mu_1(\biguplus_i f^{-1}(A_i)) = \sum_i \mu_1(f^{-1}(A_i)) = \sum_i \mu_2(A_i)$. □

If a σ-algebra is generated by a π-system $\mathcal{P}$, the measure on $\sigma(\mathcal{P})$ is uniquely defined by its values on the events of $\mathcal{P}$. This fact is stated in the following theorem. We also include the proof since it is the canonical use of Lemma 3.9.

Theorem 3.20 (Measure uniqueness)**.** *Suppose* μ_1 *and* μ_2 *are* finite *measures on* $\sigma(\mathcal{P})$ *agreeing on* S*,* $\mu_1(S) = \mu_2(S) < \infty$*, where* $\mathcal{P}$ *is a* π*-system. If* μ_1 *and* μ_2 *agree on* $\mathcal{P}$*, then they agree on* $\sigma(\mathcal{P})$.

Proof. Let $\mathcal{G} = \{A \in \sigma(\mathcal{P}) \mid \mu_1(A) = \mu_2(A)\}$ be the good sets, namely the measurable sets where the two measures agree. We will see $\mathcal{G}$ is a λ-system. First notice it is nonempty since S belongs. Using $\mu_1(S) = \mu_2(S) < \infty$, closure under complement follows by $\mu_1(A^\mathsf{c}) = \mu_1(S) - \mu_1(A) = \mu_2(S) - \mu_2(A) = \mu_2(A^\mathsf{c})$. Closure under denumerable disjoint union is: $\mu_1(\biguplus_i A_i) = \sum_i \mu_1(A_i) = \sum_i \mu_2(A_i) = \mu_1(\biguplus_i A_i)$. Given that $\mathcal{P} \subseteq \mathcal{G}$ then by Lemma 3.9, $\sigma(\mathcal{P}) \subseteq \mathcal{G}$, that is μ_1 and μ_2 agree in the generated σ-algebra. □

A similar result trades more restrictive families where the measures coincide, for more general measures. It is also stronger in the sense it shows the *existence* of the measure.

Theorem 3.21 (Carathéodory extension)**.** *Let* μ *be a measure on the algebra* Γ *of subsets of* S*, and assume that* μ *is* σ-finite *on* Γ*. Then* μ *has a unique extension to the measure on* $\sigma(\Gamma)$.

Notice that in particular if two σ-finite measures coincide on an algebra, they coincide in the generated σ-algebra. Another important result in this line is that measures can be *approximated* by the family of the generators given that they form an algebra [4].

Corollary 3.22. *Let Γ be an algebra, and let μ a* finite *measure on $\sigma(\Gamma)$. Then, for each $A \in \sigma(\Gamma)$ and $0 < \varepsilon$, there is a set $B \in \Gamma$ such that $\mu(A \triangle B) < \varepsilon$, where $A \triangle B \doteq (A \setminus B) \cup (B \setminus A)$ is the symmetric difference.*

Previous results are important since they characterize and approximate the measure in terms of families smaller than the whole σ-algebra. This will be fundamental in the proof of many forthcoming results. For example, if we define a measure giving a value to each possible rational-endpoint generator $[p, q)$ of the Borel σ-algebra, there is a unique measure in the generated $\mathcal{B}(\mathbb{R})$, because rational endpoint intervals form a π-system. The standard σ-finite measure for the Borel measurable space is the Lebesgue measure.

Definition 3.28 (Lebesgue measure)**.** For the Borel measurable space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ the σ-finite measure defined on the generators by $\lambda([p, q)) = q - p$ is *the Lebesgue measure.*

The following example shows a measurable set A that is dense in $[0, 1]$, it has no interval included, and its Lebesgue measure is strictly between 0 and 1[2].

Example 3.29. Let $d_n(x)$ to be the n-th digit of the binary representation of $x \in (0, 1]$; therefore each x can be thought as an infinite sequence of coin tosses. We define the event $A_n = \{x \in (0, 1] \mid d_i(x) = d_{i+n}(x) = d_{i+2n}(x), 1 \leq i \leq n\}$, that is the set of all sequences of coin tosses such that the initial n digits are immediately repeated two more times. For example, $A_3 = \{.\underline{010}\overline{010}\underline{010}0101\ldots,\ .\underline{111}\overline{111}\underline{111}111\ldots,\ .\underline{001}\overline{001}\underline{001}111\ldots,\ \ldots\}$. Let $A = \bigcup_i A_i$ be the event of having an infinite sequence of tosses of a coin, where some finite initial segment is repeated twice over. The Lebesgue measure of A_n is the probability of flipping a fair coin $3n$ times and getting the first third equal to the other two, that is $\lambda(A_n) = 2^n/2^{3n} = 1/2^{2n}$, with $n > 0$. Therefore, by countable subadditivity (Theorem 3.18), $0 < \lambda(A) = \lambda(\bigcup_{0<n} A_n) \leq \sum_{0<n} 1/2^{2n} = 1/3$. What is remarkable is that the event A is dense in $[0, 1]$, no interval is included ($\forall p < q, [p, q) \not\subseteq A$), and it has a positive measure that is strictly lower than 1. The example can be generalized to k repetitions for large k in order that A measures less than an arbitrary $\varepsilon > 0$.

One important result in Measure Theory is the existence of sets in the real line that are not quantifiable by the Lebesgue measure. The theorem by Vitali [4, p.41] [66, Theorem 3.38] says *there are nonmeasurable sets in the measure space* $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ These are called *Vitali sets.* Although this result depends on the Axiom of Choice (AC), and therefore constructive

[2]Compare this with $[0, 1] \setminus \mathbb{Q}$ that is dense, has no interval included, but it measures 1.

mathematicians find it unacceptable, it is one of the most important reasons to construct sound results based in Measure Theory.

If a measure is concentrated in a countable set, it is called *discrete*.

Definition 3.30 (Discrete measure)**.** Given a measure space (S, Σ, μ), μ is called *discrete* if for some countable $A \subseteq S$, $\mu(S \setminus A) = 0$.

Measures defined by $\mu(A) = \sum_i c_i \chi_A(s_i)$, where $c_i \in \mathbb{R}^+$ and $\{s_i\}_i \in \Sigma$, are discrete. They are also denoted $\{s_i \mapsto c_i\}_i$. The Dirac delta is the simplest discrete measure.

Example 3.31 (Dirac delta)**.** The *Dirac delta* function $\delta_s(A) = \chi_A(s)$ is a discrete probability measure. Moreover, any discrete probability measure is a denumerable convex combination of Dirac deltas $\mu(A) = \sum_i c_i \delta_{s_i}(A)$ with $\sum_i c_i = 1$.

The discrete part of a measure is contained in a countable set.

Proposition 3.23. *For a σ-finite measure, the set of measurable singletons having positive measure, $\{s \mid \{s\} \in \Sigma, 0 < \mu(\{s\})\}$ is at most countable.*

In contrast to discrete measures we define the continuous measures.

Definition 3.32 (Continuous measure)**.** A measure μ in the measurable space (S, Σ) is called *continuous* if $\mu(\{s\}) = 0$ for all $\{s\} \in \Sigma$, or equivalently for all denumerable $A \subseteq S$, $\mu(A) = 0$.

Properties that fail only in a null measure set, are said to be *valid almost everywhere* with respect to the measure.

Definition 3.33 (μ-a.e.)**.** Given a measure space (S, Σ, μ), we say that property $P \in \Sigma$ is *μ-almost everywhere valid* if $\mu(S \setminus P) = 0$. We write P μ-a.e.

A support set of a measure is a measurable set that concentrates all measure mass. Even though it is widely used, its definition in full generality has some drawbacks. The support may not exist [1, Example 12.15], or if it does it may not be uniquely defined since C can be added a zero-measure set Z still retaining all the mass. For measurable spaces coming from topological spaces ($\mathcal{B}(\mathbb{R}^k)$ for example), the definition is more precise and gains the status of a function: given a σ-algebra coming from a topological space $\mathcal{T}$, then there exists a *unique minimal closed set* C_0 such that $\mu(S \setminus C_0) = 0$ [4, Exercise 12.9].

Definition 3.34 (Support)**.** A set C is called *support* of μ, denoted $supp(\mu) = C$, if $\mu(S \setminus C) = 0$.

We end this section defining the measure for sums of measure spaces.

Definition 3.35. Given measure spaces $(S_i, \Sigma_i, \mu_i)_{i \in I}$, and the sum measurable space $(\bigoplus_{i \in I} S_i, \bigoplus_{i \in I} \Sigma_i)$, the sum measure μ is defined:

$$\mu(A) \doteq \sum_{i \in I} \mu_i(A \cap S_i)$$

3.4 Integration

First we give some closure results for measurable functions. The basic fact is that measurable functions are closed under composition.

Lemma 3.24. *Given that $f : (S, \Sigma) \to (S', \Sigma')$ and $g : (S', \Sigma') \to (S'', \Sigma'')$ are measurable, then $g \circ f : (S, \Sigma) \to (S'', \Sigma'')$ is also measurable.*

A function with n-dimensional image is measurable exactly when each of its components are.

Proposition 3.25. *$f : (S, \Sigma) \to (S', \Sigma')^n$ is measurable iff $f_i = \pi_i \circ f$ are measurable for every $1 \leq i \leq n$.*

Proof. The right to left implication is direct from the fact that measurable rectangles $\prod_{i=1}^n A_i$ generate the product measurable space, and that $f^{-1}(\prod_{i=1}^n A_i) = \bigcap_{i=1}^n f_i^{-1}(A_i)$. The converse is also direct since projections are measurable functions (Definition 3.21) and composition is a closed operator in the set of measurable functions (Lemma 3.24). □

Measurable functions of the form $f : (S, \Sigma) \to (\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$ are closed under the usual operators on the real numbers, as well as limits of functions.

Theorem 3.26. *Given measurable functions $f, g : (S, \Sigma) \to (\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$, and $c \in \mathbb{R}^+$, then*

$$cf \qquad f + g \qquad f - g \qquad fg \qquad f/g \qquad f \max g \qquad f \min g$$

are measurable. Also if $(f_i)_i$ is a sequence of measurable functions then

$$\lim_i f_i \qquad \limsup_i f_i \qquad \liminf_i f_i$$

are also measurable.

Using Theorem 3.26 we can easily build new and useful results.

Corollary 3.27. *Given a sequence $(f_i)_i$ of measurable functions $f_i : (S, \Sigma) \to (\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$, $\sum_i f_i$ is also measurable.*

Simple functions are the basis of measurable positive functions.

Definition 3.36 (Simple function)**.** A function $f : S \to \mathbb{R}^+$ is *simple* for measurable space (S, Σ), if it can be written as $f = \sum_{i=1}^{n} c_i \chi_{A_i}$ for disjoint $\{A_i\}_{i=1}^{n} \subseteq \Sigma$.

Observe that simple functions are measurable and take finitely many values. Although there are many ways of presenting the Lebesgue integral, we prefer to center its definition around the next result. It states that every measurable function has an increasing sequence of simple functions converging point-wise to it.

Theorem 3.28. *Given $f : (S, \Sigma) \to (\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$, there is a sequence of simple functions $(f_n)_n$ such that $f_n \nearrow f$.*

Proof. We fix n, and split the function range in two: $[0, n)$ and $[n, \infty)$. The first part is divided in 2^n intervals $[n\frac{i}{2^n}, n\frac{i+1}{2^n})$ indexed by i, $0 \leq i < 2^n$; while the second is kept together. Given that the intervals are measurable in $\mathcal{B}(\mathbb{R}^+)$, $f^{-1}([p, q))$ is the measurable set of all domain values reaching the range $[p, q)$ through f. The f_n is constructed as follows:

$$f_n = \sum_{i=0}^{2^n-1} n\frac{i}{2^n} \chi_{f^{-1}([n\frac{i}{2^n}, n\frac{i+1}{2^n}))} + n\chi_{f^{-1}([n,\infty))} \tag{3.3}$$

□

Figure 3.1 shows $f(x) = 2+\sin(x)+\sin(2x)$ together with simple function f_2.

For simple functions the integral is a finite sum.

Definition 3.37 (Lebesgue integral of simple functions)**.** Given a measure space (S, Σ, μ) and simple function $f : (S, \Sigma) \to (\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$, $f = \sum_{i=1}^{n} c_i \chi_{A_i}$, the *Lebesgue integral* is defined by:

$$\int f d\mu \doteq \sum_{i=1}^{n} c_i \mu(A_i)$$

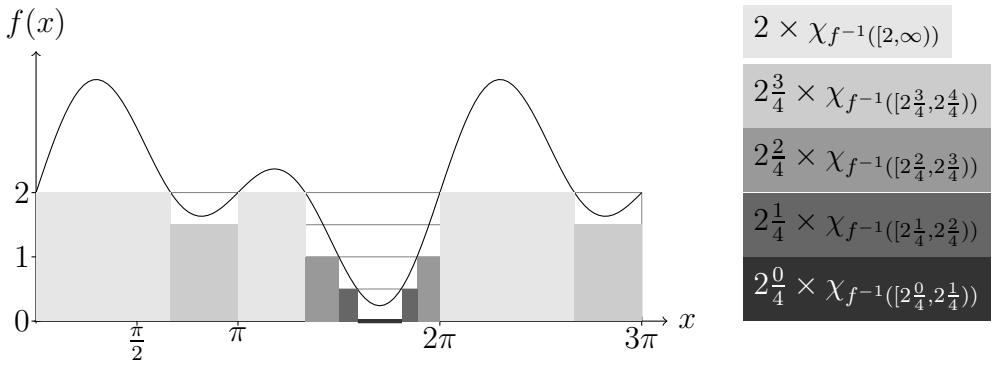


Figure 3.1: Simple function f_2 below $f(x) = 2 + \sin(x) + \sin(2x)$.

A problem of consistency could arise here, since there is not a unique way to write a simple function. Suppose $f = \sum_{i=1}^{n} a_i \chi_{A_i} = \sum_{j=1}^{m} b_j \chi_{B_j}$, then we can write $f = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i \chi_{A_i \cap B_j}$. We unfold the definition of the integral $\int f(x) d\mu(x) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i \mu(A_i \cap B_j)$, taking the constant terms out of the first sum $\sum_{i=1}^{n} a_i \sum_{j=1}^{m} \mu(A_i \cap B_i)$ and this is exactly $\sum_{i=1}^{n} a_i \mu(A_i)$. We can get $\sum_{j=1}^{m} b_j \chi_{B_j}$ similarly, therefore the definition is consistent.

Given that every measurable function is the limit of increasing simple functions, the integral of measurable functions can be readily defined.

Definition 3.38 (Lebesgue integral)**.** Let (S, Σ, μ) be a measure space and $f : (S, \Sigma) \to (\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$ be a measurable function. The *Lebesgue integral of f on measure μ* is defined as the right-hand side limit in case it is less than ∞. We write

$$\int f d\mu \doteq \lim_i \int f_i d\mu$$

where $f_i \nearrow f$ are the monotone converging simple functions of Theorem 3.28.

Sometimes we restrict the integration domain to a measurable set A, and we denote this

$$\int_A f d\mu \doteq \int f \chi_A d\mu$$

Whenever we write f *is μ integrable* or $\int f d\mu$, it means that f is measurable and the limit is a real number.

There are mainly three different ways of writing the Lebesgue integral

$$\int f d\mu \qquad \int f(x) d\mu(x) \qquad \int f(x) \mu(dx)$$

representing a trade-off between verbosity and precision. If we are integrating a one parameter function, we rather write $\int f d\mu$. If f has two or more parameters, we write $\int f(x_1, x_2) d\mu(x_1)$. Finally if the measure is defined in more than one dimension, we will write $\int f(x_2)\mu(x_1, dx_2)$.

Example 3.39. The typical example to show the virtues of Lebesgue integral is that it can integrate the everywhere non-continuous $f = \chi_{\mathbb{Q}\cap[0,1]}$, where $\int_{[0,1]} f d\lambda = 0$ for the Lebesgue measure λ. The Riemann integral is undefined for f since for every interval partition of the domain $[0, 1]$ the minimum and maximum value of the partition do not converge. It cannot be argued that Riemann integral fails because of null measure sets, since Example 3.29 defines a set that is similar to $\mathbb{Q} \cap [0, 1]$ but having a positive measure.

The Lebesgue integral has many pleasant properties, and that makes it easier to manipulate than the Riemann integral. Many of this properties are μ-a.e. invariant (Definition 3.33), since sets of μ-measure zero do not contribute to the value of the integral. The next proposition show how this μ-a.e. invariance works.

Proposition 3.29. *Let $0 \leq f$ be Borel measurable, $f = 0$ μ-a.e. iff $\int f d\mu = 0$.*

Proof. First the left to right implication. For simple functions $f = \sum_{i=1}^{n} c_i \chi_{A_i}$ with $0 \leq c_i$. If $0 < c_i$ then $\mu(A_i) = 0$ by hypothesis, therefore $\int f d\mu = \sum_{i=1}^{n} c_i \mu(A_i) = 0$. Let $f_i \nearrow f$ as in Theorem 3.28, then $0 \leq f_i \leq f$. Since $f = 0$ μ-a.e., f_i inherit the same property, then $\int f_i d\mu = 0$ for all i, concluding by the definition of the Lebesgue integral that $\int f d\mu = 0$.

For the converse notice that the predicate $f = 0$ μ-a.e. is by definition $\mu(\{x \mid 0 < f(x)\}) = 0$. Let $B = \{x \mid 0 < f(x)\}$ and $B_n = \{x \mid 1/n < f(x)\}$. By monotonicity of the integral, $0 \leq f\chi_{B_n} \leq f\chi_B = f$ implies $\int f\chi_{B_n} d\mu \leq \int f\chi_B d\mu = \int f d\mu = 0$. Given that $1/n\, \mu(B_n) \leq \int f\chi_{B_n} d\mu \leq 0$, then $\forall n, \mu(B_n) = 0$. Since $B_n \nearrow B$, by continuity from below we have $\mu(B) = 0$. □

We end this part by showing various results concerning integration. First we show results for the integral in terms of the function being integrated.

Theorem 3.30. *i. Monotonicity. $f \leq g$ μ-a.e. $\Rightarrow \int f d\mu \leq \int g d\mu$.*

ii. Linearity. $a, b \in \mathbb{R}^+ \Rightarrow \int (af + bg) d\mu = a \int f d\mu + b \int g d\mu$.

Monotonicity implies small corollaries like $f = 0$ μ-a.e. $\Rightarrow \int f d\mu = 0$ (already given in Proposition 3.29), and $f = g$ μ-a.e. $\Rightarrow \int f d\mu = \int g d\mu$. The following result is very important, since it allows to move limits out of the integral.

Theorem 3.31 (Monotone convergence)**.** *Given non-negative functions such that $f_i \nearrow f$, then $\int f_i d\mu \nearrow \int f d\mu$.*

The following corollary is a direct consequence of previous theorem and linearity.

Corollary 3.32. *If $(f_i)_i$ are μ integrable then $\int(\sum_i f_i)d\mu = \sum_i \int f_i d\mu$.*

If the family $\{f_i\}_i$ does not converge from below, but it is *dominated* by an integrable g, then the interchange of limit and integral is also valid[3].

Theorem 3.33 (Dominated convergence)**.** *Given non-negative and measurable $f, \{f_i\}_i, g$ such that for all i, $f_i \leq g$ μ-a.e., $f_i \to f$ μ-a.e., and g is μ integrable, then $\int f_i d\mu \to \int f d\mu$.*

Second, we show a property of the integral in terms of the measure.

Proposition 3.34. $\mu = \sum_i \mu_i \Rightarrow \int f d\mu = \sum_i \int f d\mu_i$.

The integral is also measure limit preserving. This is given by the Portmanteau Theorem [36]. It implies that if $\mu_i \to \mu$, then $\int f d\mu_i \to \int f d\mu_i$. However this is only valid for separable metrizable space generating a topology on the set of measures, and a continuous bounded function.

The next theorem states a kind of distributivity property of the integral with respect to the disjoint union of measurable sets where the integral takes place.

Theorem 3.35. *For pair-wise disjoint $\{A_i\}_i$ and non-negative measurable f, $\int_{(\biguplus_i A_i)} f d\mu = \sum_i \int_{A_i} f d\mu$.*

Proof.

$$\begin{aligned}\int_{(\biguplus_i A_i)} f d\mu &= \int \chi_{(\biguplus_i A_i)} f d\mu = \int (\textstyle\sum_i \chi_{A_i} f) d\mu && \text{(integral def. and disj. of } A_i\text{)} \\ &= \textstyle\sum_i \int \chi_{A_i} f d\mu = \sum_i \int_{A_i} f d\mu && \text{(Corollary 3.32 and def. again)}\end{aligned}$$

□

[3]There is one more classical result that lies between the monotone convergence theorem and dominated convergence theorem, it is called Fatou's lemma [2].

Integration of New Measures. The Lebesgue integral of a function with respect to a measure can be used to generate new measures. The integration of a so called density function induces a new measure. First the integration result.

Proposition 3.36. *Let (S, Σ, μ) be a measure space and $f : (S, \Sigma) \to (\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$ be a measurable function. Then $\nu(A) = \int_A f d\mu$ is a measure. In this case, it is said that ν* has density *f* with respect to *μ.*

Proof. Notice that ν is strict $\nu(\emptyset) = \int \chi_\emptyset f d\mu = \mu(\emptyset)$, by Definition 3.37, and it is also σ-additive: $\nu(\biguplus_i A_i) = \sum_i \nu(A_i)$ by Theorem 3.35. Therefore ν is a measure. □

Notice that whenever $\mu(A) = 0$ then $\nu(A) = 0$. The Radon-Nikodym theorem [2] goes in the opposite direction: if the last condition is fulfilled, that is, ν can be *disintegrated* or *derived* from μ, in symbols $f = \frac{d\nu}{d\mu}$. Observe that if we interpret the expression $\nu(A) = \int_A f d\mu$ of Proposition 3.36 without rigor, we could have written the equality $d\nu = f d\mu$, however it must be shown it is valid if used inside a Lebesgue integral.

Theorem 3.37. *If ν has density f with respect to μ, and f is non-negative, then $\int g d\nu = \int g f d\mu$ for all measurable function g mapping on the reals.*

Proof. First we prove it for a *characteristic function* $g = \chi_A$. Using Definition 3.37, Definition 3.38 and Proposition 3.36 we have:

$$\int g d\nu = \int \chi_A d\nu = \nu(A) = \int_A f d\mu = \int \chi_A f d\mu = \int g f d\mu$$

Now suppose $g = \sum_{i=1}^n c_i \chi_{A_i}$ is a *simple function.* Using linearity and our previous result, the next equality holds:

$$\begin{aligned}\int g d\nu = \int (\textstyle\sum_{i=1}^n c_i \chi_{A_i}) d\nu &= \textstyle\sum_{i=1}^n c_i \int \chi_{A_i} d\nu \\ &= \textstyle\sum_{i=1}^n c_i \int \chi_{A_i} f d\mu = \int (\sum_{i=1}^n c_i \chi_{A_i}) f d\mu = \int g f d\mu\end{aligned}$$

Finally the general case when g is a *non-negative measurable function.* By Theorem 3.28 there is a monotone sequence of simple functions such that $g_n \nearrow g$, therefore we can use the monotone convergence theorem (Theorem 3.31) and our previous result and write:

$$\begin{aligned}\int g d\nu = \int (\lim_i g_i) d\nu &= \lim_i \textstyle\int g_i d\nu \\ &= \lim_i \textstyle\int g_i f d\mu = \int (\lim_i g_i) f d\mu = \int g f d\mu\end{aligned}$$

□

Theorem 3.37 follows a typical proof strategy, that was also partially used in Proposition 3.29.

- First the result is proved for characteristic functions χ_A,
- then for simple functions $\sum_{i=1}^{n} c_i \chi_{A_i}$,
- finally for non-negative measurable functions.

3.5 The σ-algebra of Measures $\Delta(\Sigma)$

Given that measures on (S, Σ) are functions in $\Sigma \to \mathbb{R}^+$, using Definition 3.22 we can endow the space of measures with a σ-algebra. This structure is usually attributed in the literature to [30], but it can also be found in earlier works [26, 42].

Definition 3.40 (σ-algebra of measures)**.** Given measurable space (S, Σ), and the set of measures $\Delta(S)$, the σ-algebra $\Delta(\Sigma)$ is generated by Definition 3.20, with functions indexed by $Q \in \Sigma$

$$(\lambda\mu : \mu(Q)) : \Delta(S) \to \mathbb{R}^+$$

Usually the measures' σ-algebra is stated as in the following proposition.

Proposition 3.38. *The σ-algebra $\Delta(\Sigma)$ is generated by all sets of the form:*

$$\Delta^B(Q) \doteq (\lambda\mu : \mu(Q))^{-1}(B) = \{\mu \in \Delta(S) \mid \mu(Q) \in B\},$$

where $B \in \mathcal{B}(\mathbb{R}^+)$ and $Q \in \Sigma$.

We remark that by definition of $\Delta(\Sigma)$, the evaluation of measures $(\lambda\mu : \mu(Q)) : (\Delta(S), \Delta(\Sigma)) \to (\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$ is a measurable function for all $Q \in \Sigma$.

Mind the notation overloading for Δ, as it lifts the space S and the σ-algebra Σ to the measurable space of measures $(\Delta(S), \Delta(\Sigma))$, and it also denotes the generators of the σ-algebra[4]. We can write single bounds $\Delta^{\bowtie q}(Q) = \{\mu \mid \mu(Q) \bowtie q\}$, where $\bowtie \in \{>, <, \geq, \leq\}$, as well as intervals of measure values $\Delta^{[p,q)}(Q) = \{\mu \mid p \leq \mu(Q) < q\}$.

The σ-algebra of measures is denoted by $\Delta(\Sigma)$ and not the other way round: $\Sigma(\Delta)$ or Σ_Δ. We adhere to the fact that Δ is an endofunctor in the category Meas of measurable spaces (objects) and measurable functions (arrows), $\Delta : \mathsf{Meas} \to \mathsf{Meas}$ [47].

[4]There is also overloading for λ, as it was previously used for the Lebesgue measure. We allow the clash since both of them are common notation.

The notation $\Delta(\Sigma)$ does not specify the carrier set S, but it can be obtained from Σ. It also does not specify if the measures are general, σ-finite, finite, probabilities or subprobabilities, but it can be deduced from the context. We can take the general definition of measures and its σ-algebra and intersect it to obtain the relative σ-algebra (Definition 3.12) that we need. The restriction sets defining the relative σ-algebras are $\Delta(\Sigma)$ measurable. The following are particular cases that are important in this thesis:

$$\Delta^{\leq 1}(S) \quad \text{subprobability measures} \tag{3.4}$$
$$\Delta^{=1}(S) \quad \text{probability measures} \tag{3.5}$$
$$\bigcup_i \Delta^{<i}(S) \quad \text{finite measures} \tag{3.6}$$

For σ-finite measures it is not easy to describe it with $\Delta^{\bowtie q}(Q)$ generators, nevertheless the definition of the relative σ-algebra accepts arbitrary sets, so it is well defined.

Notice that by Definition 3.20 $\Delta^{>q}(Q) = \{\mu \in \Delta(S) \mid q < \mu(Q)\}$, where $q \in \mathbb{Q}^+$ also generates $\Delta(\Sigma)$ since $\{(q, \infty) \mid q \in \mathbb{Q}^+\}$ generates $\mathcal{B}(\mathbb{R}^+)$. Even though the bounds are now countable, if the σ-algebra Σ is not countable, $\Delta(\Sigma)$ is, in principle, not countably generated. The following result is from [65], and shows that if the underlying σ-algebra is generated by a π-system, then the related σ-algebra on probability measures is also generated by the same π-system.

Lemma 3.39. *Let (S, Σ) be a measurable space and let $\mathcal{P}$ be a π-system such that $\Sigma = \sigma(\mathcal{P})$. Then $\Delta(\Sigma) = \sigma(\{\Delta^{>q}(Q) \mid q \in \mathbb{Q} \cap [0,1], Q \in \mathcal{P}\})$.*

Observe that q is bounded to the interval $[0, 1]$. For subprobabilities and finite measures the proof can be easily adapted, however for σ-finite measures the lemma is not valid in general.

By Proposition 3.3 and the previous lemma, it is sufficient that Σ is countably generated for $\Delta(\Sigma)$ to be also countably generated. For example $\Delta(\mathcal{B}(\mathbb{R}^+))$ is countably generated for probability measures. Irrespective if the underlying σ-algebra separates points or not (Definition 3.8), $\Delta(\Sigma)$ always separates points.

Proposition 3.40. *$\Delta(\Sigma)$ separates points.*

Proof. Let $\mu \neq \mu'$, therefore there is $Q \in \Sigma$ and $q \in \mathbb{Q}^+$ such that, without loss of generality, $\mu(Q) < q < \mu'(Q)$. Therefore the generator $\Delta^{>q}(Q)$ separates μ and μ'. □

3.6 Transition Probabilities

Suppose we have a product space $(S_1 \times S_2, \Sigma_1 \otimes \Sigma_2)$, representing a two-stage experiment, and two measures μ_1, μ_2 respectively. If the experiments are independent, the joint measure of the outcome $A_1 \in \Sigma_1$ and then the outcome $A_2 \in \Sigma_2$, is the product of the measures $\mu(A_1 \times A_2) = \mu_1(A_1) \times \mu_2(A_2)$. However there could be dependencies in the outcome of the second experiment given the result of the first. Transition measures model the idea of *dependent quantification*, where the quantification of an experiment (measure) depends on previous experiment results.

In the next definition, given $s \in S$ (the result of the previous experiment) we can define the function $f(s, Q)$ that measures the next experiment Q given that the previous result was s. Also this conditional measure $f(s, Q)$ has to be a measurable function in the first parameter in order to integrate it with the measure of the first experiment.

Definition 3.41 (Conditional measure)**.** Given measurable spaces (S_1, Σ_1) and (S_2, Σ_2), a *transition measure* or *conditional measure* is a function $f : S_1 \times \Sigma_2 \to \mathbb{R}^+$, such that for all $s \in S_1$ $f(s, \cdot) : \Sigma_2 \to \mathbb{R}^+$ is a measure, and $f(\cdot, Q) : (S_1, \Sigma_1) \to (\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$ is measurable for all $Q \in \Sigma_2$. If $f(s, \cdot)$ is a probability measure it is called *conditional probability* or *Markov kernel*.

We give a simple but useful result that shows the connection between conditional measures and the σ-algebra of measures.

Lemma 3.41. *Function $f : S_1 \times \Sigma_2 \to \mathbb{R}^+$ is a conditional measure iff its curried version $f : S_1 \to \Delta(S_2)$ is measurable.*

Proof. It follows from the equalities

$$\begin{aligned} f(\cdot, Q)^{-1}(B) &= \{s \mid f(s, Q) \in B\} \\ &= \{s \mid f(s)(Q) \in B\} = \{s \mid f(s) \in \Delta^B(Q)\} = f^{-1}(\Delta^B(Q)) \end{aligned}$$

□

The main integration theorem says that a measure on Σ_1 and a conditional measure $S_1 \times \Sigma_2 \to \mathbb{R}^+$ can be integrated into a measure of the product space $\Sigma_1 \otimes \Sigma_2$. The proof is given for the finite measure case, since it shows measure theoretic techniques previously developed. The σ-finite version of the proof can be found, for example, in [2]; its proof requires different tools. We use Dynkin's Lemma twice instead of monotone convergence theorem and Carathéodory extension theorem.

Theorem 3.42 (Product measure)**.** *Let (S_1, Σ_1, μ_1) be a measure space with σ-finite μ_1, let (S_2, Σ_2) be a measurable space, and let $\mu_2 : S_1 \times \Sigma_2 \to \mathbb{R}^+$ be a conditional measure that is uniformly σ-finite (σ-finite independently of the first coordinate value). Then there is a unique* product measure *μ on $\Sigma_1 \otimes \Sigma_2$,*

$$\mu(A) = \int_{S_1} \mu_2(s_1, A_{|s_1}) d\mu_1(s_1),$$

where $A \in \Sigma_1 \otimes \Sigma_2$, such that for all $A_1 \in \Sigma_1$, and $A_2 \in \Sigma_2$,

$$\mu(A_1 \times A_2) = \int_{A_1} \mu_2(s_1, A_2) d\mu_1(s_1).$$

Proof. (finite measure case) First notice that by Proposition 3.17, $A_{|s_1}$ is a measurable set. Then in order to show that $\mu_2(s_1, A_{|s_1}) : S_1 \to \mathbb{R}^+$ is a measurable function, we will use π-λ (Lemma 3.9) with good sets $\mathcal{G} = \{A \in \Sigma_1 \otimes \Sigma_2 \mid \mu_2(s_1, A_{|s_1})$ is measurable$\}$. Let $\mathcal{A}$ be the generators (rectangles) of the product σ-algebra forming a π-system. For elements $A_1 \times A_2 \in \mathcal{A}$, we have $\mu_1(s_1, (A_1 \times A_2)_{|s_1}) = \mu_1(s_1, A_2)\chi_{A_1}(s_1)$. By Theorem 3.26 measurable functions are closed under binary products, therefore $\mathcal{A} \subseteq \mathcal{G}$. We will show that $\mathcal{G}$ is a λ-system therefore for every $A \in \Sigma_1 \otimes \Sigma_2$ the function $\mu_1(s_1, A_{|s_1})$ is measurable. Being the generators a subset of the good sets, we have that $S_1 \times S_2 \in \mathcal{G}$. Using finiteness of the measure, we write $\mu_1(s_1, A^{\mathsf{c}}_{|s_1}) = \mu_1(s_1, S_1 \times S_{2_{|s_1}}) - \mu_1(s_1, A_{|s_1})$, and again by Theorem 3.26, the good sets are closed under complements. For denumerable disjoint union of good sets, $\mu_1(s_1, (\biguplus_i A_i)_{|s_1}) = \sum_i \mu_1(s_1, A_{|s_1})$, therefore by Corollary 3.27 it is also good.

For the rectangle case it is direct, since:

$$\mu(A_1 \times A_2) = \textstyle\int_{S_1} \mu_2(s_1, A_2)\chi_{A_1}(s_1) d\mu_1(s_1) = \int_{A_1} \mu_2(s_1, A_2) d\mu_1(s_1).$$

Uniqueness follows from Theorem 3.20, since the generators form a π-system where the integrals should coincide. □

We emphasize again the importance of singling out the integration variable for this case, otherwise integrals like the ones appearing in the middle of previous proof should have been written in a less convenient way. For example instead of $\int_{S_1} \mu_2(s_1, A_2)\chi_{A_1}(s_1) d\mu_1(s_1)$ we could have written $\int_{S_1} \mu_2(\cdot, A_2)\chi_{A_1} d\mu_1$. The importance of this notation will become more evident in the generalization of this result to products of n spaces.

For a conditional measure that is independent of the first experiment, namely $\forall s_1, s_2 \in S, \mu_2(s_1, Q) = \mu_2(s_2, Q)$ (denoted $\mu_2(Q)$), we have the following classical result.

Corollary 3.43 (Classical product measure)**.** *Let* (S_1, Σ_1, μ_1) *and* (S_2, Σ_2, μ_2) *be measure spaces, where both* μ_1 *and* μ_2 *are* σ*-finite. The product measure space* $(S_1 \times S_2, \Sigma_1 \otimes \Sigma_2, \mu)$*, with*

$$\mu(A) = \int \mu_2(A_{|s_1}) d\mu_1(s_1) = \int \mu_1(A_{|s_2}) d\mu_2(s_2)$$

being the unique measure such that $\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$*. This measure is called* product measure *and it is written* $\mu = \mu_1 \times \mu_2$*.*

The particular case of Corollary 3.43 for $S_1 = S_2 = \mathbb{R}$, $\Sigma_1 = \Sigma_2 = \mathcal{B}(\mathbb{R})$, can be alternatively obtained in a few steps. By Proposition 3.16, $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^2)$. The finite measure $\mu = \mu_1 \times \mu_2$ is defined on the finite disjoint unions of intervals of $\mathbb{R}^2$, $\mu(\uplus_{i=1}^n([a_i, b_i) \times [a'_i, b'_i))) = \sum_{i=1}^n (b_i - a_i)(b'_i - a'_i)$. Since the finite disjoint unions of rectangles form an algebra, by Theorem 3.21 there is a unique extension of measure μ on the whole σ-algebra.

The product measure defined in Corollary 3.43 can be used to integrate a measurable function $f : S_1 \times S_2 \to \mathbb{R}^+$. This is done in Fubini's Theorem for dependent measures and in the classical Fubini theorem for the independent measures. The result is given since it is needed for the proof of Theorem 3.46.

Theorem 3.44 (Fubini)**.** *Assume the hypothesis of Theorem 3.42, and let* $f : S_1 \times S_2 \to \mathbb{R}^+$ *be measurable, then* $\int f(s_1, s_2)\mu_2(s_1, ds_2)$ *is well defined and measurable on* s_1*. Besides it holds that:*

$$\int f d\mu = \int \left(\int f(s_1, s_2)\mu_2(s_1, ds_2) \right) \mu_1(ds_1).$$

For the sake of completeness, given independent measures, we have the classical version of Theorem 3.44. It states that the integral of f is equal to the double iterated integral where the order among them is irrelevant.

Theorem 3.45 (Classical Fubini)**.** *Assume the hypothesis of Theorem 3.42, with independent measures,* $\mu = \mu_1 \times \mu_2$*, and let* $f : S_1 \times S_2 \to \mathbb{R}^+$ *be measurable, then*

$$\int f d\mu = \int \left(\int f(s_1, s_2) d\mu_1(s_1) \right) \mu_2(ds_2) = \int \left(\int f(s_1, s_2) d\mu_2(s_2) \right) \mu_1(ds_1).$$

The simultaneous generalization of both Theorem 3.42 and Theorem 3.44 gives the definition of the product measure for finite product measurable space, and how to iteratively integrate a measurable f. The conditional measure $\mu_2(s_1, Q_2)$ for a binary product space $S_1 \times S_2$ can be generalized straightforwardly to an n-stage experiment as $\mu_i(s_1, \ldots, s_{n-1}, A_n)$, a conditional measure for the product space $\prod_{i=1}^n S_i$.

Theorem 3.46 (Finite product measure)**.** *Let $((S_i, \Sigma_i))_{i=1}^n$ be measurable spaces. Let μ_1 be a σ-finite measure on Σ_1, and let $\mu_i(s_1, \ldots, s_{i-1}, A_i)$ be uniformly σ-finite conditional measures. There is a unique measure μ on $\bigotimes_{i=1}^n \Sigma_i$ such that for each measurable rectangle $\prod_{i=1}^n A_i$,*

$$\mu(\textstyle\prod_{i=1}^n A_i) = \int_{A_1} \mu_1(ds_1) \int_{A_2} \mu_2(s_1, ds_2) \ldots \int_{A_n} \mu_n(s_1, \ldots, s_{n-1}, ds_n).$$

For every measurable $f : \prod_{i=1}^n S_i \to \mathbb{R}^+$,

$$\int f d\mu = \int \mu_1(ds_1) \int \mu_2(s_1, ds_2) \ldots \int f(s_1, \ldots, s_n) \mu_n(s_1, \ldots, s_{n-1}, ds_n),$$

where the intermediate integrals are measurable functions $\prod_{i=1}^j S_i \to \mathbb{R}^+$, with $j < n$.

Proof. By induction on n. The base case $n = 2$ is valid by Theorem 3.42 and Theorem 3.44. Suppose it is valid for n, we prove it for $n + 1$.

Using Theorem 3.42 we decompose the $n + 1$ tuple in a pair of lengths n and 1.

$$\mu((\textstyle\prod_{i=1}^n A_i) \times A_{n+1}) = \int_{\prod_{i=1}^n A_i} \mu_{n+1}(s_1, \ldots, s_n, A_{n+1}) d\mu(s_1, \ldots s_n).$$

Notice we are overloading μ, it is the product measure for all finite dimensions. The right-hand side is an integral of a measurable function in an n-dimensional product measure space, therefore we apply induction hypothesis using the fact that:

$$\begin{aligned} &\int_{\prod_{i=1}^n A_i} f d\mu = \int \chi_{(\prod_{i=1}^n A_i)} f d\mu \\ &= \int \chi_{A_1}(s_1) \mu_1(ds_1) \int \ldots \int f(s_1, \ldots, s_n) \chi_{A_n}(s_n) \mu_n(s_1, \ldots, s_{n-1}, ds_n). \end{aligned}$$

Then

$$\begin{aligned} &\int_{A_1} \mu_1(ds_1) \int_{A_2} \mu_2(s_1, ds_2) \ldots \int_{A_n} \mu_{n+1}(s_1, \ldots, s_n, A_{n+1}) \mu_n(s_1, \ldots, s_{n-1}, ds_n) \\ &= \int_{A_1} \mu_1(ds_1) \int_{A_2} \mu_2(s_1, ds_2) \ldots \int_{A_n} \mu_n(s_1, \ldots, s_{n-1}, ds_n) \int_{A_{n+1}} \mu_{n+1}(s_1, \ldots, s_n, ds_{n+1}). \end{aligned}$$

The $\int f d\mu$ case is handled in a similar way. We split in two the $n+1$ product, apply the $n = 2$ case, and then the induction hypothesis. ∎

The previous theorem can be rearranged to obtain a recursive version of the iterated integral:

$$\mu(A_1) = \mu_1(A_1)$$
$$\mu(A \times A_{n+1}) = \int_A \mu_{n+1}(s_1, \ldots, s_n, A_{n+1}) d\mu(s_1, \ldots, s_n)$$

where $A \in \bigotimes_{i=1}^{n} \Sigma_i$. Again observe that μ is overloaded, or in mathematical terms, it is a measure in the sum measurable space $(\bigoplus_n (\prod_{i=1}^{n} S_i), \bigoplus_n (\bigotimes_{i=1}^{n} \Sigma_i))$.

Likewise, denumerable product σ-algebras are defined in terms of finite product σ-algebras; measures on denumerable product σ-algebras are defined in terms of the measures on finite product σ-algebras. However, for the unbounded denumerable product we can only consider *probability measures.*

Given the family of measurable spaces $((S_i, \Sigma_i))_i$, its denumerable product σ-algebra is built around measurable rectangles of the form $(\prod_{i=1}^{n} A_i) \times (\prod_{n<i} S_i)$. This introduces a possible inconsistency. In a denumerable product σ-algebra, the measurable rectangle with base $A_1 \times A_2$ is equal to the measurable rectangle with base $A_1 \times A_2 \times S_3$. Therefore, in order to obtain a well-defined measure on denumerable product σ-algebras, the two events $A_1 \times A_2$ and $A_1 \times A_2 \times S_3$ should be equally quantified. Therefore, the following theorem only applies to probability measures.

Theorem 3.47 (Denumerable product probability measure)**.** *Let $((S_i, \Sigma_i))_i$ be a denumerable sequence of measurable spaces. Let μ_1 be a probability measure on Σ_1, and let $\mu_i(s_1, \ldots, s_{i-1}, A_i)$ be a* transition probability *or Markov kernel. Then the following definition of probability measure over measurable rectangles:*

$$\mu((\prod_{i=1}^{n} A_i) \times (\prod_{n<i} S_i)) \doteq \mu(\prod_{i=1}^{n} A_i)$$

extends uniquely to $(\prod_i S_i, \bigotimes_i \Sigma_i)$, where $\mu(\prod_{i=1}^{n} A_i)$ is as in Theorem 3.46.

Note that each μ_i being a transition *probability*, any trailing $\prod_{i=n+1}^{m} S_i$ in the measurable rectangles keeps all the probability mass.

Chapter 4

Nondeterministic Labeled Markov Processes

In Chapter 2, we presented a few classes of discrete transition systems introducing external nondeterminism, internal nondeterminism and probabilities (see also Table 1.1). The discrete nature of these models simplifies all measurability issues. Apart from this simplification, there is no apparent reason why a model like PA should not be extended to continuous state space and distributions.

For continuous state space the most basic model is a Markov process [27], a mathematical tool devised to capture physical phenomena like Brownian motion. This model augmented with external nondeterminism through labels is called labeled Markov processes [20]. This extension is not minor (the same consideration applies for the difference between MC and PLTS), since labels introduce the ideas of behavioral equivalence, logical characterization through modal logics, controlled compositionallity, schedulers, etc.

In this chapter, we extend LMPs with internal nondeterminism. We call such extension *nondeterministic labeled Markov processes* (NLMPs). The extension is non-trivial since the nondeterminism has to be carefully embedded in the measure theoretic framework. In the following, we first recall LMPs and then introduce NLMPs. Afterwards we discuss two variants of NLMPs that will be recurrent along this thesis.

4.1 Labeled Markov Processes

Labeled Markov processes (LMPs) [20] are models that capture *probabilistic choice* and *external nondeterminism* in the setting of continuous state space. LMPs propose a model that deals explicitly with measurability problems.

The state space S is endowed with a σ-algebra Σ that defines measurable events. The transition relation is captured by an (arbitrary) set of labeled transition (sub)probability functions. Thus, for each label a we have the transition function τ_a such $\tau_a(s, Q)$ is the (sub)probability of reaching the event Q, from state s, by label a.

Definition 4.1 (LMP)**.** A *labeled Markov process* (LMP) [20, 21] is a triple $(S, \Sigma, \{\tau_a \mid a \in L\})$ where Σ is a σ-algebra on the set of states S, and for each label $a \in L$, $\tau_a : S \times \Sigma \to [0, 1]$ is a transition (sub)probability function.

By Lemma 3.41, we can say that $(S, \Sigma, \{\tau_a \mid a \in L\})$ is an LMP iff every $\tau_a : S \to \Delta(S)$ is measurable. The next example is system modeling the continuous motion of a fish with continuous stochastic behavior, where the probability distribution is position-dependent.

Example 4.2 (Stochastic Fish [5])**.** Suppose we model the nondeterministic and probabilistic behavior of a fish that lives in the (real) unidimensional aquarium $[0, 1]$ that is divided in two equal parts. From $[0, 1/2)$ it is filled with water, while in the other half $[1/2, 1]$ some kids put cola soda (Figure 4.1). The fish has two modes: *swim* and *jump*. In the swim mode, the fish moves errant in the *safe region* of the aquarium with transition probability defined by the rational-endpoint intervals,

$$\tau_{swim}(x, [p, q]) = \begin{cases} 2\lambda([p, q] \cap [0, 1/2)) & \text{if } x \in [0, 1/2) \\ 0 & \text{otherwise} \end{cases}$$

where λ is the Lebesgue measure (see Definition 3.28). The other mode implies a probabilistic daredevil jump to the right, falling into a possibly *unsafe region* of its living space. From position x there is $2(1/2 - x)$ chance of getting in the $[x, 1/2)$ safe part, and $2x$ of getting into poisoned water. This probability is clearly dependent on the fish position:

$$\tau_{jump}(x, [p, q]) = \begin{cases} 2\lambda([p, q] \cap [x, 1/2)) + 4x\lambda([p, q] \cap [1/2, 1]) & \text{if } x \in [0, 1/2) \\ 0 & \text{otherwise} \end{cases}$$

By Theorem 3.20 there are exactly two measures on $\mathcal{B}([0, 1])$ that coincides with τ_{swim} and τ_{jump} as defined above. Notice also that both τ_{swim} and τ_{jump} are transition probabilities, that is $\tau_{swim}(x, [0, 1]) = \tau_{jump}(x, [0, 1]) = 1$.

For the measurability of the transition function we use Lemma 3.39 since the events in $\mathcal{B}([0, 1])$ are generated by the π-system given by all the rational endpoint intervals $[p, q]$. We take the jump transition and rewrite the expression fixing the current position x and the target interval $[p, q]$:

$$\begin{aligned} \tau_{jump}(x, [p, q]) = \chi_{[0,1/2)}(x) \{&2(\max(\min(q, 1/2), x) - \min(\max(p, x), 1/2)) \\ &+ 4x(\max(\min(q, 1), 1/2) - \min(\max(p, 1/2), 1))\} \end{aligned}$$

Since the identity function $f(x) = x$ and $\chi_A(x)$ with $A \in \mathcal{B}([0,1])$ are measurable, by repeated uses of Theorem 3.26 and Lemma 3.24, the whole expression is measurable in the first argument. The transition τ_{swim} is similar.

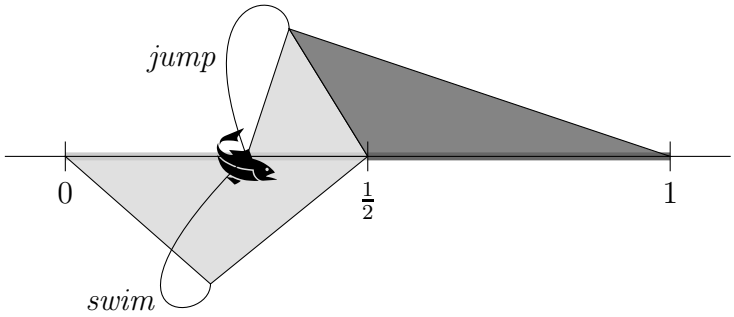


Figure 4.1: Stochastic fish modeled using labeled Markov process.

Notice that this LMP cannot be *lumped* to a discrete system, that is, there is no discrete probabilistic automata that is behaviorally equivalent to the stochastic fish.

4.2 Nondeterministic Labeled Markov Processes

There have been several attempts to define nondeterministic continuous probabilistic transition systems and all of them are straightforward extensions of (simpler) discrete versions. There are two fundamental differences in our new model. The first one is that the *nondeterministic transition function* T_a now maps states to *measurable sets of* probability measures rather than arbitrary sets as previous approaches. This is motivated by the fact that later on the nondeterminism has to be resolved using schedulers. If we allowed the target set of states to be an arbitrary subset (as some continuous ones [8, 12, 15]), the system as a whole could suffer from nonmeasurability issues and therefore it could not be quantified. (Rigorously speaking, labels should also be provided with a σ-algebra, but we omit it here since it is not needed.) The second difference is inspired by the definition of LMP and Lemma 3.41: we ask that, for each label $a \in L$, T_a is a measurable function. One of the reasons for this restriction is to have well defined modal operators of a probabilistic Hennessy-Milner logic, like in the LMP case.

Definition 4.3 (NLMP)**.** A *nondeteministic labeled Markov process* (NLMP) is a triple $(S, \Sigma, \{T_a \mid a \in L\})$ where Σ is a σ-algebra on the set of states S, and for each label $a \in L$,

$$T_a : S \to \Delta(\Sigma)$$

is a measurable function.

Notice that we changed the transition function symbol from τ_a in LMP to T_a in order to emphasize the difference. For the requirement that T_a is measurable, we need to endow $\Delta(\Sigma)$ with a σ-algebra. This is a key construction to forthcoming definitions and theorems.

Definition 4.4 (Hit σ-algebra)**.** Given measurable space (S, Σ), the *hit σ-algebra* over Σ is the one generated by the hit sets $H_A \doteq \{B \in \Sigma \mid B \cap A \neq \emptyset\}$ for $A \in \Sigma$. Therefore the hit σ-algebra is $H(\Sigma) \doteq \sigma(\{H_A \mid A \in \Sigma\})$.

This construction is similar to that of the Effros-Borel space [36] and resembles the so-called hit-and-miss topologies [48].

To prove measurability of T_a, is sufficient to check the generators of the hit σ-algebra $H(\Delta(\Sigma))$ (Proposition 3.11), that is, we have to show that for each $\xi \in \Delta(\Sigma)$, $T_a^{-1}(H_\xi) \in \Sigma$. Observe that

$$T_a^{-1}(H_\xi) = \{s \in S \mid T(s) \cap \xi \neq \emptyset\}$$

is the set of all states s such that, through label a, the transition function "hits" the set of measures in ξ. This forms the basis to *existentially quantify over the nondeterminism*, and it is fundamental for the definition of the bisimulation and the logic.

The next two examples (inspired by an example in [11]) show why T_a is required to map into measurable sets and to be measurable. For these examples we fix the state space and σ-algebra in the real unit interval with the standard Borel σ-algebra.

Example 4.5. Let $\mathcal{V} = \{\delta_q \mid q \in V\}$, where V is a non-measurable Vitali set in $[0, 1]$. The set $\mathcal{V}$ is not measurable in $\Delta(\Sigma)$. This can be verified by first noting that the function $(\lambda s : \delta_s) : S \to \Delta(S)$ is measurable: $(\lambda s : \delta_s)^{-1}(\Delta^{>q}(Q))$ is equal to Q if $q < 1$, otherwise it is $\emptyset$. Therefore if $\mathcal{V}$ were measurable then $(\lambda s : \delta_s)^{-1}(\mathcal{V}) = V$ would also be measurable, contradicting the nonmeasurability of V. We define the transition function (see left-hand side of Figure 4.2)

$$T_a(0) = \mathcal{V}$$

The resolution of the internal nondeterminism by means of schedulers would require to assign probabilities to all possible choices, and this amounts to

measure the nonmeasurable set $T_a(0)$. *This is why we require that T_a maps into measurable sets.*

Example 4.6. We define the following transition function as in the right-hand side of Figure 4.2,

$$T_a(0) = \{\mu\}$$
$$T_b(s) = \textbf{if } (s \in V) \textbf{ then } \{\delta_1\} \textbf{ else } \emptyset,$$

where $s \in [0,1]$, and μ is the uniform distribution in $[0,1]$. Notice that $T_a(s)$ and $T_b(s)$ are measurable sets for every s.

Assuming that there is a scheduler that chooses to first do a and then b starting at 0, the probability of such set of executions cannot be measured. It requires to apply μ to the set $T_b^{-1}(H_{\Delta(S)}) = V$ which is not measurable. *This is why we ask for measurable T_a with respect to the hit σ-algebra.* Besides, we will later need that sets $T_a^{-1}(H_\xi)$ are measurable so that the semantics of the logic maps into measurable sets.

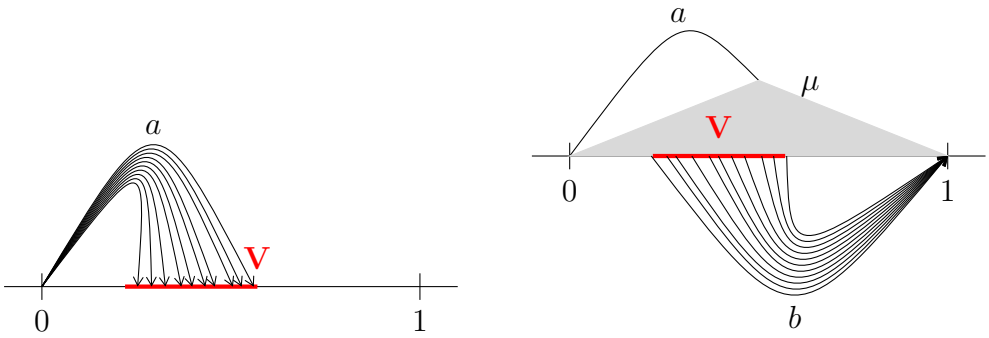


Figure 4.2: Two non-probabilistic continuous LTS showing measurability issues.

It is worth noting that the example do not make use of probabilistic choice in transitions, it is purely nondeterministic. This kind of continuous LTS endowed with a σ-algebra will continue to be used throughout this thesis.

We now show an example of a simple NLMP in terms of its definition, but it encodes a strong nondeterminism in the probabilistic choice that is dependent of the state.

Example 4.7. Let the structure $([0,1], \mathcal{B}([0,1]), \{T_a\})$ be such that

$$T_a(x) = \Delta^{\geq 1/2}([x,1]) \cap \Delta^{=1}([0,1]) \qquad \text{with } x \in [0,1]$$

It represents a position-dependent transition function of a continuous system. Notice that it encodes a strong nondeterminism, since given position x, any probability measure μ such that $1/2 \leq \mu([x,1])$ is included. It can be seen that the transition consists of intersection of $\Delta(\Sigma)$ generators, therefore $T_a(x) \in \Delta(\Sigma)$ for all $x \in [0,1]$.

In addition, for $\xi \in \Delta(\Sigma)$, the inverse image of a generator is $T_a^{-1}(H_\xi) = A_\xi = \{x \mid \Delta^{\geq 1/2}([x,1]) \cap \Delta^{=1}([0,1]) \cap \xi \neq \emptyset\}$. Suppose $x \in T_a^{-1}(H_\xi)$ and $0 \leq y \leq x$. By monotonicity of the measures, $\Delta^{\geq 1/2}([y,1]) \supseteq \Delta^{\geq 1/2}([x,1])$. Hence $T_a^{-1}(H_\xi) = [0, \sup(A_\xi)]$ if the supremum is included in A_ξ, or else $T_a^{-1}(H_\xi) = [0, \sup(A_\xi))$. Since both intervals are Borel measurable, we conclude that $([0,1], \mathcal{B}([0,1]), \{T_a\})$ is an NLMP.

NLMPs as a generalization of LMPs Notice that an LMP is an NLMP without internal nondeterminism. That is, an NLMP in which $T_a(s)$ is a *singleton* for all $a \in L$ and $s \in S$, is an LMP. In fact, an LMP can be encoded as an NLMP by taking $T_a(s) = \{\tau_a(s)\}$. (We will prove this formally in Proposition 4.2.) As a consequence it is necessary that singletons $\{\mu\}$ be measurable in $\Delta(\Sigma)$ for the NLMP to be well defined. The following lemma gives sufficient conditions on the carrier σ-algebra Σ to ensure that all singletons are measurable in $\Delta(\Sigma)$.

Lemma 4.1. *Let the σ-algebra Σ be countably generated. Then, for all σ-finite $\mu \in \Delta(S)$, $\{\mu\} \in \Delta(\Sigma)$.*

Proof. Using the remark below Proposition 3.3 we build $\mathcal{F}$, a denumerable algebra such that $\Sigma = \sigma(\mathcal{F})$. Notice that the set

$$\bigcap\{\Delta^{>q_i}(Q_i) \mid Q_i \in \mathcal{F}, q_i \in \mathbb{Q} \cap [0,1], q_i < \mu(Q_i)\} \cap \bigcap\{\Delta^{<q_i}(Q_i) \mid Q_i \in \mathcal{F}, q_i \in \mathbb{Q} \cap [0,1], \mu(Q_i) < q_i\} \tag{4.1}$$

is in $\Delta(\Sigma)$. Now it suffices to show that the set (4.1) is equal to $\{\mu\}$. By construction μ is in the non-empty intersection. Take μ' such that $\mu \neq \mu'$. By Theorem 3.21, there must be a $Q_i \in \mathcal{F}$ such that $\mu(Q_i) \neq \mu'(Q_i)$. If $\mu(Q_i) > \mu'(Q_i)$ then μ' does not belong to the first intersection; if $\mu(Q_i) < \mu'(Q_i)$, μ' does not belong to the second one. □

Note that Lemma 4.1 also gives sufficient conditions to define NLMPs with *finite and denumerable nondeterminism*. Once $\{\mu\}$ is measurable, we define $T_a(s) = \bigcup_i \{\mu_i\}$ with measurable image.

Notice also that asking for measurable singletons in $\Delta(\Sigma)$ does not trivialize Σ (in the sense that $\Sigma = 2^S$). A nontrivial example in which Lemma 4.1 holds is the standard Borel σ-algebra in $\mathbb{R}$. A less obvious example is the

σ-algebra **Q-coQ** from Example 3.9. Notice that **Q-coQ** cannot separate one irrational from another (let alone asking for all singletons being measurable). Nevertheless, as it is generated by a denumerable family, it is under the conditions of Lemma 4.1 and hence for every measure μ on it, $\{\mu\}$ is measurable on $\Delta(\mathbf{Q\text{-}coQ})$.

The formal connection between NLMPs and LMPs is an immediate consequence of the next proposition.

Proposition 4.2. *Let $T_a(s) = \{\tau_a(s)\}$ for all $s \in S$ and let Σ be a σ-algebra on S. Then $\tau_a : S \to \Delta(S)$ is measurable iff $T_a : S \to \Delta(\Sigma)$ is measurable.*

Proof. Let $\xi \in \Delta(\Sigma)$. Note that $T_a(s) \in H_\xi$ iff $\{\tau_a(s)\} \cap \xi \neq \emptyset$ iff $\tau_a(s) \in \xi$. Then $T_a^{-1}(H_\xi) = \tau_a^{-1}(\xi)$. Therefore τ_a is measurable whenever T_a is measurable. For the converse, we have that $T_a^{-1}(H_\xi)$ is measurable for all generators H_ξ. As a consequence T_a is measurable in general. □

4.3 Structure on the Labels

The definition of NLMPs does not impose any restriction on the set of labels L. This could lead to measurability issues in the external nondeterminism.

Example 4.8. We define an NLMP conforming with Definition 4.3 with two states $\{s, s'\}$, and a continuous set of labels $L = [0, 1]$. From initial state s the system loops if the label r is in a set V, otherwise it jumps to s', where V is a Vitali set (Figure 4.3). We write the transition function:

$$T_r(s) = \textbf{if } (r \in V) \textbf{ then } \{\delta_{s'}\} \textbf{ else } \{\delta_s\},$$

where $r \in [0, 1]$. Suppose there is an *external source of nondeterminism* choosing the label, and this external source is implemented by a probabilistic choice η. If we want to quantify the probability of jumping from s to s', this amounts to measure the nonmeasurable set

$$\{r \mid T_r(s) \cap \{\delta_{s'}\} \neq \emptyset\} = V$$

This example shows that in order to quantify the behavior of the environment, it is also needed to *impose structure on the labels*, even for continuous LTS without probabilistic choice and nondeterminism.

Definition 4.9 (NLMP with structure on the labels)**.** Let Σ_L be a σ-algebra on L such that the singletons are measurable. The structure (S, Σ, Σ_L, T) is

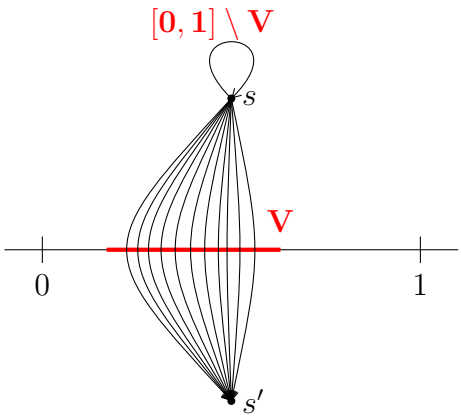


Figure 4.3: Continuous deterministic LTS showing a measurability issue on the labels.

an *NLMP with structure on the labels* if (S, Σ) is a measurable space and the transition function

$$T : S \to \Sigma_L \otimes \Delta(\Sigma)$$

is such that for all $a \in L$, $T(\cdot)_{|a} : S \to \Delta(\Sigma)$ is measurable.

Transition function $T(\cdot)_{|a}$ is well defined since sections of measurable sets in a product σ-algebra are also measurable sets (Proposition 3.17). We ask $\{a\} \in \Sigma_L$ for technical reasons. This is a reasonable restriction since we still want to provide the environment the possibility to *push individual buttons*.

Notice that the LTS in Example 4.8 does not conform to Definition 4.9. We can rewrite the transition function in the form of Definition 4.9 as $T(s) = (([0,1] \setminus V) \times \{\delta_s\}) \uplus (V \times \{\delta_{s'}\})$, however $T(s)_{|\delta_{s'}} = V \notin \Sigma_L$, where $\Sigma_L = \mathcal{B}([0,1])$, therefore $T(s) \notin \Sigma_L \otimes \Delta(\Sigma)$ by Proposition 3.17.

Also note that NLMPs with structure on the labels are particular cases of general NLMPs of Definition 4.3. Indeed, if (S, Σ, Σ_L, T) is an NLMP with structure on the labels, then $(S, \Sigma, \{T(\cdot)_{|a} \mid a \in L\})$ is an NLMP as in Definition 4.3.

Now we show that the transition function measurability condition of Definition 4.9 is strictly weaker than asking $T : S \to \Sigma_L \otimes \Delta(\Sigma)$ to be measurable with the hit structure on $\Sigma_L \otimes \Delta(\Sigma)$.

Proposition 4.3. *If $T : S \to \Sigma_L \otimes \Delta(\Sigma)$ is measurable, then for all $a \in L$, $T(\cdot)_{|a} : S \to \Delta(\Sigma)$ is measurable.*

Proof. Let $A \in \Sigma_L \otimes \Delta(\Sigma)$. Then $T^{-1}(H_A) \in \Sigma$. Therefore $T^{-1}(H_{\{a\}\times\xi}) = \{s \mid T(s) \cap (\{a\} \times \xi) \neq \emptyset\} \in \Sigma$, with $\xi \in \Delta(\Sigma)$. It is not hard to see this set is exactly $\{s \mid T(s)_{|a} \cap \xi \neq \emptyset\}$, concluding $T_{|a}^{-1}(H_\xi) \in \Sigma$. □

Because of Example 3.24 the reverse implication is not valid in general. The following example from [33] shows a concrete NLMP with structure on the labels where $T(\cdot)_{|a} : S \to \Delta(\Sigma)$ is measurable, but $T : S \to \Sigma_L \otimes \Delta(\Sigma)$ is not.

Example 4.10. Let $(\mathbb{R}, \mathcal{B}(\mathbb{R}), 2^{\mathbb{R}}, T)$ be an NLMP with structure on the labels where $L = \mathbb{R}$, $\Sigma_L = 2^{\mathbb{R}}$, and

$$T(s) = \{(s, \delta_1)\} \cup (\{s\}^{\mathsf{c}} \times \{\delta_0\})$$

Notice that the system moves deterministically to state 1 if the label agrees with the current value of the state, or else it jumps to state 0. It can be seen that $T(s) \in 2^{\mathbb{R}} \otimes \Delta(\mathcal{B}(\mathbb{R}))$ for all $s \in \mathbb{R}$. In addition, for $\xi \in \Delta(\mathcal{B}(\mathbb{R}))$, $T(\cdot)_{|s}^{-1}(H_\xi) \in \{\emptyset, \mathbb{R}, \{s\}, \{s\}^{\mathsf{c}}\}$, hence $T(\cdot)_{|s}$ measurable for all $s \in \mathbb{R}$. Now let $A = V \times \{\delta_1\} \in 2^{\mathbb{R}} \otimes \Delta(\mathcal{B}(\mathbb{R}))$, with V being a Vitali set. Note that $T^{-1}(H_A) = \{s \mid T(s) \cap (V \times \{\delta_1\}) \neq \emptyset\} = V \notin \mathcal{B}(\mathbb{R})$. Hence, even when $T(\cdot)_{|s}$ is measurable for all $s \in \mathbb{R}$, T is not.

Instead of the stronger condition of $T : S \to \Sigma_L \otimes \Delta(\Sigma)$ measurable, we stick to the measurability condition of Definition 4.9 that is strictly weaker in order to include more systems while keeping the structure that is needed.

4.4 Non-probabilistic NLMPs

Non-probabilistic NLMPs are a variation of the NLMPs where all transitions are Dirac measures. This is basically a continuous LTS with a σ-algebra attached. Let

$$\delta(S) \doteq \{\delta_s \mid s \in S\}$$

be the set of all Dirac probability measures over S. We define a restriction of standard NLMPs, using the relative σ-algebra $\Delta(\Sigma)|\delta(S)$.

Definition 4.11 (Non-probabilistic NLMP, I)**.** A *non-probabilistic NLMP* is a triple $(S, \Sigma, \{T_a \mid a \in L\})$ where Σ is a σ-algebra on the set of states S, and for each label $a \in L$,

$$T_a : S \to \Delta(\Sigma)|\delta(S)$$

is a measurable function. If we explicitly state the measurable spaces, $T_a : (S, \Sigma) \to (\Delta(\Sigma)|\delta(S), H(\Delta(\Sigma)|\delta(S)))$.

Although $\Delta(\Sigma)|\delta(S)$ is a σ-algebra. the set $\xi \cap \delta(S)$ is not necessarily in $\Delta(\Sigma)$ where $\xi \in \Delta(\Sigma)$. We show that given some reasonable conditions in the underlying σ-algebra, $\delta(S)$ is measurable; hence $\xi \cap \delta(S)$ is also measurable.

Proposition 4.4. *Given a countably generated σ-algebra Σ that separates points, then the set of Dirac measures is measurable, that is, $\delta(S) \in \Delta(\Sigma)$.*

Proof. Let $\mathcal{C}$ be the countable set of generators. Consider the set:

$$\Delta^{=1}(S) \cap \bigcap_{A_i \in \mathcal{C}} (\Delta^{=0}(A_i) \cup \Delta^{=0}(A_i^\complement)) \tag{4.2}$$

We show that this set in $\Delta(\Sigma)$ is equal to $\delta(S)$. For the right to left inclusion, let $\delta_s \in \delta(S)$ be a Dirac measure. Clearly $\delta_s(S) = 1$ so $\delta_s \in \Delta^{=1}(S)$. Using $\delta_s(S) = \delta_s(A_i) + \delta_s(A_i^\complement) = 1$ for all i, it holds that either $\delta_s(A_i) = 0$ or $\delta_s(A_i^\complement) = 0$, so $\delta_s \in \bigcap_{A_i \in \mathcal{C}} (\Delta^{=0}(A_i) \cup \Delta^{=0}(A_i^\complement))$.

Let μ be a measure in the left-hand side. Using the set construction and the property that Σ separates points, we will show there is $x \in S$ such that $\mu(\{x\}) = 1$.

First notice $\mu(S) = 1$. Moreover, for all $A_i \in \mathcal{C}$, $\mu(A_i) = 0$ or $\mu(A_i^\complement) = 0$. For each $A_i \in \mathcal{C}$, let $B_i = A_i$ if $\mu(A_i) = 1$ or $B_i = A_i^\complement$ if $\mu(A_i^\complement) = 1$. Notice that $\mu(B_i) = 1$ for all i. Using B_i we construct a decreasing sequence $C_0 = B_0$, $C_{n+1} = C_n \cap B_{n+1}$. The limit is the measurable set $C = \bigcap_i C_i$. Using Theorem 3.18, we conclude $\mu(C_i) \searrow \mu(C)$.

Next we show that the sequence $\mu(C_i)$ is constantly 1. This is valid for $\mu(C_0) = \mu(B_0) = 1$. Taking it is valid for n, $\mu(C_{n+1}) = \mu(C_n \cap B_{n+1}) = \mu(C_n) + \mu(B_{n+1}) - \mu(C_n \cup B_{n+1}) = 1 + 1 - 1 = 1$, being the second equality the inclusion-exclusion principle (Theorem 3.18). Therefore $\mu(C) = 1$.

It remains to show that C is a singleton. Notice first it is nonempty since $\mu(C) = 1$. Suppose towards a contradiction that there are $s, s' \in C$, $s \neq s'$. Since Σ is separative and generated by $\{A_i\}_i$ then, by Proposition 3.5, there is some A_j such that $s \notin A_j \ni s'$. Then, either $s \notin C = \bigcap_i B_i$ or $s' \notin C = \bigcap_i B_i$ depending on whether $B_j = A_j$ or $B_j = A_j^\complement$. □

The above result can be generalized to arbitrary measurable sets.

Corollary 4.5. *Let Σ be countably generated σ-algebra that separates points. Then set of Dirac measures over $Q \in \Sigma$ is measurable, that is $\delta(Q) \in \Delta(\Sigma)$.*

The previous result also implies that $Q \in \Sigma \Rightarrow \delta(Q) \in \Delta(\Sigma)|\delta(S)$. We now prove that the converse holds.

Proposition 4.6. *For all $\xi \in \Delta(\Sigma)|\delta(S)$, there is a $Q \in \Sigma$ such that $\xi = \delta(Q)$.*

Proof. Let the good sets $\mathcal{G} = \{\xi \in \Delta(\Sigma)|\delta(S) \mid \exists Q \in \Sigma, \xi = \delta(Q)\}$. The $\Delta(\Sigma)$ generators are included in $\mathcal{G}$ since $\Delta^{>q}(Q) \cap \delta(S)$ is either $\delta(Q)$ or $\emptyset$. Suppose $\xi = \delta(Q)$, then $\xi^{\mathsf{c}} \cap \delta(S) = \delta(Q)^{\mathsf{c}} \cap \delta(S) = \delta(Q^{\mathsf{c}})$, so $\mathcal{G}$ is closed under complements. Closure under denumerable union follows since $\bigcup_i \delta(Q_i) = \delta(\bigcup_i Q_i)$. Using the good sets principle, the property is valid for every measurable $\xi \in \Delta(\Sigma)|\delta(S)$. □

Corollary 4.5 and Proposition 4.6 shows there is a one-to-one correspondence between the sets in Σ and in $\Delta(\Sigma)|\delta(S)$. We can exploit this idea and give an alternative definition of non-probabilistic NLMPs. This definition does not involve the use of probabilities, since the hit σ-algebra can be taken over the σ-algebra of states Σ. We change the definition of NLMPs so that the target of the transition functions are measurable sets of states.

Definition 4.12 (Non-probabilistic NLMP, II)**.** A *non-probabilistic NLMP* is a triple $(S, \Sigma, \{\tilde{T}_a \mid a \in L\})$ where Σ is a σ-algebra on the set of states S, and for each label $a \in L$,

$$\tilde{T}_a : S \to \Sigma$$

is a measurable function. If we state the measurable spaces explicitly, $\tilde{T}_a : (S, \Sigma) \to (\Sigma, H(\Sigma))$.

The next proposition states that non-probabilistic NLMPs responding to Definition 4.11 are equally expressive to those responding to Definition 4.12.

Proposition 4.7. *Let Σ be a countably generated σ-algebra that separates points. Then*

i. if $(S, \Sigma, \{T_a \mid a \in L\})$ is a non-probabilistic NLMP in the sense of Definition 4.11, and $\tilde{T}_a(s) = \{t \mid \mu \in T_a(s) \wedge \mu(\{t\}) = 1\}$ for all $a \in L$ and $s \in S$, then $(S, \Sigma, \{\tilde{T}_a \mid a \in L\})$ is a non-probabilistic NLMP in the sense of Definition 4.12.

ii. if $(S, \Sigma, \{\tilde{T}_a \mid a \in L\})$ is a non-probabilistic NLMP in the sense of Definition 4.12, and $T_a(s) = \delta(\tilde{T}_a(s))$ for all $a \in L$ and $s \in S$, then $(S, \Sigma, \{T_a \mid a \in L\})$ is a non-probabilistic NLMP in the sense of Definition 4.11.

Proof. First notice that $(\lambda x : \delta_x)$ is an isomorphism from (S, Σ) to $(\delta(S), \Delta(\Sigma)|\delta(S))$ (this follows from Corollary 4.5 and Proposition 4.6).

For (i), we have that $T_a(s) = \delta(\tilde{T}_a(s))$ and $\tilde{T}_a(s)$ is measurable in Σ because of Proposition 4.6. Moreover, using the fact that $(\lambda x : \delta_x)$ is an isomorphism, we can calculate that $\tilde{T}_a^{-1}(H_Q) = T_a^{-1}(H_{\delta(Q)})$ for all $Q \in \Sigma$ and hence $\tilde{T}_a$ is measurable. Then this new NLMP is well defined.

For (ii) $T_a(s) = \delta(\tilde{T}_a(s))$ by definition and hence $T_a(s)$ is measurable by Corollary 4.5. Moreover, having Proposition 4.6 and the fact that $(\lambda x : \delta_x)$ is an isomorphism, we have that for all $\xi \in \Delta(\Sigma)|\delta(S)$, $T_a^{-1}(H_\xi) = \tilde{T}_a^{-1}(H_Q)$ for some $Q \in \Sigma$. Thus T_a is measurable and the new NLMP is well defined. □

4.5 Concluding Remarks

We generalized LMPs in order to include internal nondeterminism. We maintained the idea of transition function measurability to obtain measurable semantics for a (forthcoming) modal operator. In doing so, we have to define a custom σ-algebra, namely the hit σ-algebra. The image of the transition function was modified to allow for measurable nondeterminism. This measurability in the set of target probabilities will be essential to define schedulers. In fact we have been careful in giving the most general definition that rendered measurable systems. We also showed two transition systems (Examples 4.5, 4.6) to motivate why the transition function is required to map into measurable sets and to be measurable. The NLMP of Example 4.7 is new and it shows the modeling power of NLMPs. Throughout the chapter we gave NLMPs all the possible freedom to capture not only probability measures, but also subprobability, finite and σ-finite measures. For example Lemma 4.1 was generalized with respect to [17]. Example 4.8 showed weaknesses of the initial proposed model for uncountable labels set. We tightened the definition and showed that the new definition is not stronger than needed (Proposition 4.3 and Example 4.10). The definition of NLMPs with structure on the labels is new with respect to our previous publications. We also showed two alternative but equivalent definitions for non-probabilistic NLMPs, where the definition of the hit σ-algebra was crucial to equate them.

Chapter 5

Uses and Comparisons

The definition of the transition function in NLMPs is based mainly in two ideas: using measurable sets in $\Delta(\Sigma)$ to represent underspecification, and capturing existential quantification by endowing $\Delta(\Sigma)$ with a hit σ-algebra. In the first part of this chapter we develop the idea that $\Delta(\Sigma)$, and more precisely their generators $\Delta^{>q}(Q)$, form an appropriate basis for specifying sets of measures with meaningful properties. We already gave some examples in the previous chapter. In this one we explore and elaborate on the use of the set of generators as a specification framework. Next, we use NLMPs as the underlying semantics of two non-trivial modeling formalisms that allow for the specification of probabilistic and nondeterministic continuous behavior. Both formalisms were defined a priori and independently of NLMPs, therefore they constitute a good testing scenario on the way NLMPs was defined. We also show how NLMPs captures the semantics of pGCL, a probabilistic and nondeterministic programming language. Finally, we compare NLMPs to other classes of labeled transition systems including both probabilities and nondeterminism. We show, whenever possible, how this framework translates to NLMPs.

5.1 $\Delta(\Sigma)$ for Probabilistic Nondeterminism

The codomain of the transition functions in NLMPs, namely $\Delta(\Sigma)$, is a powerful tool to model continuous probabilistic nondeterminism. When using probabilistic automata (PA) and other discrete systems alike, transitions are usually specified by enumerating all possible target distributions. That is, we would usually write $s \xrightarrow{a} \{\mu_1, \dots, \mu_n\}$. Underspecification of probabilistic choices have also been represented through some kind of symbolic machinery. Thus, a dense set of (sub)probability functions can be defined as the convex

closure of a finite set of probability functions [58], the up-closure of a finite set of subprobability functions [46], the set of all subprobability functions matching a super-additive function [22], or the set of solutions of a linear constraint [19]. In our case, the generators of $\Delta(\Sigma)$ seem to provide a natural tool to describe underspecified models. Take, for example, the transition T_a defined on $[0,1]$ by:

$$T_a(0) = \Delta^{>1/2}([1/4, 3/4]) \cap \Delta^{=1}([0,1])$$

and $T_a(x) = \emptyset$ for $x \neq 0$. This encodes all possible distributions such that the probability of reaching states in $[1/4, 3/4]$ is greater than $1/2$ when starting at the state 0. Notice that $T_a(0)$ defines a continuous nondeterminism that includes discrete distributions like $\mu = \{0 \mapsto 1/5, 1/2 \mapsto 3/5, 1 \mapsto 1/5\}$, or continuous distributions like the normal distribution $\mathcal{N}(0.5, 0.1)$. Besides, notice that T_a is well defined in the Borel σ-algebra $\mathcal{B}([0,1])$. Indeed, T_a maps into measurable sets (either $\Delta^{>1/2}([1/4, 3/4]) \cap \Delta^{=1}([0,1])$ or $\emptyset$), and it is a measurable function as its preimage is either $\emptyset$ or $\{0\}$.

Denumerable set operations (unions, intersection and complements) on generators are the constructors of a very expressive language for probabilistic nondeterminism. This language is basically the language that constructs the σ-algebra $\Delta(\Sigma)$. We have already seen examples of the usage of this language where with a modest nesting level, we can capture interesting sets of measures. We briefly recall a few of them:

$$\begin{array}{rl} \bigcup_i \Delta^{<i}(S) & \text{Finite measures (3.6)} \\ \Delta^{=1}(S) \cap \bigcap_{A_i \in \mathcal{C}} (\Delta^{=0}(A_i) \cup \Delta^{=0}(A_i^{\mathsf{c}})) & \text{Delta Dirac measures (4.2)} \end{array}$$

We also show how the convex combination given by a probabilistic scheduler on the PA of Figure 5.1 (originally in Figure 1.1), can be compactly encoded in $\Delta(\Sigma)$.

$$\begin{aligned} T_{\mathit{flip}}(s_0) &= \Delta^{\geq 1/4}(\{sh\}) \cap \Delta^{\geq 1/4}(\{st\}) \cap \Delta^{=1}(\{sh, st\}) \\ T_{\mathit{heads}}(sh) &= \{\delta_{s_1}\} \\ T_{\mathit{tails}}(st) &= \{\delta_{s_2}\} \end{aligned}$$

That is: tossing the coin should outcome heads with probability at least $1/4$, likewise for tails, and it can only produce heads or tails with probability 1. If instead, we want to encode subprobability distributions in the same example, we can change the *flip* transition function to

$$T_{\mathit{flip}}(s_0) = \Delta^{\geq 1/4}(\{sh\}) \cap \Delta^{\geq 1/4}(\{st\}) \cap \Delta^{=0}(S \setminus \{sh, st\}) \cap \Delta^{=1}(S)$$

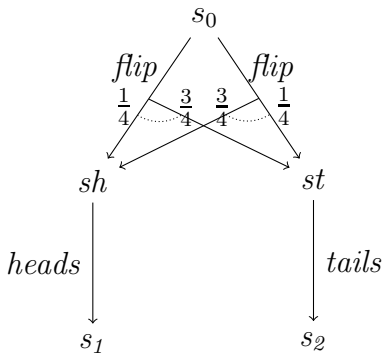


Figure 5.1: A dense set of probability functions is captured by convex combinations on PA with discrete nondeterminism.

Every discrete (sub)probability distribution, typically found in discrete probabilistic systems, can also be encoded in $\Delta(\Sigma)$ as follows:

$$\{s_i \mapsto p_i\}_{i=1}^n = \textstyle\bigcap_{i=1}^n \Delta^{=p_i}(\{s_i\}) \quad \text{with } \textstyle\sum_{i=1}^n p_i = 1$$

Up-closure of a discrete subprobability measure [46] $\{s_i \mapsto p'_i\}_{i=1}^n$ such that $\sum_{i=1}^n p'_i \leq 1$ can also be given in terms of $\Delta(\Sigma)$. This continuous set of discrete probability distribution is encoded as:

$$\begin{aligned}\left\{\{s_i \mapsto p_i\}_{i=1}^n \mid \textstyle\sum_{i=1}^n p_i = 1, \forall i \in [1..n], p'_i \leq p_i\right\} \\ = \Delta^{=1}(\{s_i\}_{i=1}^n) \cap \textstyle\bigcap_{i=1}^n \Delta^{\geq p'_i}(\{s_i\})\end{aligned}$$

The geometric distribution over the discrete measurable space $(\mathbb{N}, 2^{\mathbb{N}})$, $\mu(\{k\}) = (1-p)^k p$, is encoded as a denumerable intersection of generators in $\Delta(2^{\mathbb{N}})$:

$$\mu = \textstyle\bigcap_k \Delta^{=(1-p)^k p}(\{k\})$$

We have already shown that the set of all Dirac measures, namely $\delta(S)$, is measurable in $\Delta(\Sigma)$ provided Σ is countably generated and separates points (see Proposition 4.4). More precisely

$$\delta(S) = \Delta^{=1}(S) \cap \textstyle\bigcap_{A_i \in \mathcal{C}} (\Delta^{=0}(A_i^{\mathsf{c}}) \cup \Delta^{=0}(A_i))$$

where $\mathcal{C}$ is a countable set generating Σ. By slightly varying the proof of Proposition 4.4, we can see that the set

$$\Phi_{=1} \doteq \{c\delta_s \mid 0 < c, s \in S\} \tag{5.1}$$

of all one-point measures is also measurable. In fact

$$\Phi_{=1} = \Delta^{>0}(S) \cap \bigcap_{A_i \in \mathcal{C}} (\Delta^{=0}(A_i^{\mathsf{c}}) \cup \Delta^{=0}(A_i))$$

Moreover if $\overline{\mathbf{0}}$ is the null measure, i.e. $\overline{\mathbf{0}}(S) = 0$, then

$$\Phi_{\leq 1} \doteq \overline{\mathbf{0}} \cup \{c\delta_s \mid 0 < c, s \in S\} \tag{5.2}$$

Therefore, this set is also measurable since

$$\Phi_{\leq 1} = \{\overline{\mathbf{0}}\} \cup \Phi_{=1} = \bigcap_{A_i \in \mathcal{C}} (\Delta^{=0}(A_i^{\mathsf{c}}) \cup \Delta^{=0}(A_i)) \tag{5.3}$$

In Figure 5.2 it can be seen how $\Phi_{\leq 1}$ discards any two-point measure in $\Delta(\mathcal{B}(\mathbb{R}^+))$. Let $\mu_2 = \{s_1 \mapsto c_1, s_2 \mapsto c_2\}$, with $s_1, s_2 \in \mathbb{R}$, $s_1 < s_2$ and $0 < c_1, c_2$. There is a generator $[q_1, q_1']$ (closed interval with rational endpoints) such that $s_1 \in [q_1, q_1'] \not\ni s_2$. Therefore $\mu_2 \notin \Delta^{=0}([q_1, q_1']^{\mathsf{c}})$ and $\mu_2 \notin \Delta^{=0}([q_1, q_1'])$, concluding μ_2 is not in the intersection.

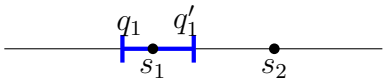


Figure 5.2: A two-point measure is not in the measurable description of $\Phi_{\leq 1}$.

The generalization of (5.1) and (5.2) to n points are defined by:

$$\Phi_{=n} \doteq \left\{ \textstyle\sum_{i=1}^{n} c_i\delta_{s_i} \mid \{c_i\}_{i=1}^{n} \subseteq \mathbb{R}^+ \setminus \{0\}, \{s_i\}_{i=1}^{n} \subseteq S \right\} \tag{5.4}$$

$$\Phi_{\leq n} \doteq \{\textstyle\sum_{i=1}^{j} c_i\delta_{s_i} \mid 0 \leq j \leq n, \{c_i\}_{i=1}^{j} \subseteq \mathbb{R}^+ \setminus \{0\}, \{s_i\}_{i=1}^{j} \subseteq S\} \tag{5.5}$$
$$\text{with } \textstyle\sum_{i=1}^{0} c_i\delta_{s_i} = \overline{\mathbf{0}}.$$

Equation (5.3) shows that is easier to define $\Phi_{\leq n}$ in terms of $\Delta^{>q}(Q)$ than $\Phi_{=n}$. We therefore capture $\Phi_{\leq n}$ in $\Delta(\Sigma)$ first, and express $\Phi_{=n}$ in terms of the former. First we motivate the generalization of (5.3) giving a measurable set that captures $\Phi_{\leq 2}$ on the Real measurable space:

$$\bigcap_{\substack{q_1 < q_1' < q_2 < q_2' \\ q_1, q_1', q_2, q_2' \in \mathbb{Q}}} \Delta^{=0}(([q_1, q_1'] \cup [q_2, q_2'])^{\mathsf{c}}) \cup \Delta^{=0}([q_1, q_1']) \cup \Delta^{=0}([q_2, q_2'])$$

Zero to two-point measures belong to the intersection since one of the three generators contains an interval with null measure. However measures with three or more points with positive measure are excluded. For example, the

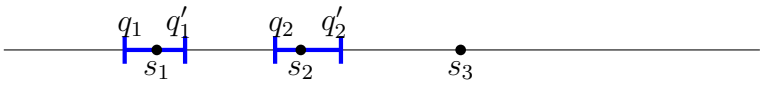


Figure 5.3: A three-point measure is not in the measurable description of $\Phi_{\leq 2}$.

discrete measure $\mu_3 = \{s_1 \mapsto c_1, s_2 \mapsto c_2, s_3 \mapsto c_3\}$ with $s_1 < s_2 < s_3$ and $0 < c_1, c_2, c_3$ is not in $\Phi_{\leq 2}$. Take $[q_1, q_1'], [q_2, q_2']$ around s_1, s_2 respectively as depicted in Figure 5.3. It can be observed that $\mu_3 \notin \Delta^{=0}([q_1, q_1'])$, $\mu_3 \notin \Delta^{=0}([q_2, q_2'])$, as well as $\mu_3 \notin \Delta^{=0}(([q_1, q_1'] \cup [q_2, q_2'])^{\mathsf{c}})$. Therefore the measure μ_3 is not in the intersection.

The next lemma is auxiliary to Proposition 5.2.

Lemma 5.1. *Given a measurable space (S, Σ) that is countably generated by an algebra $\mathcal{F}$ and separates points, for all $2 \leq n$, $\{s_i\}_{i=1}^n \subseteq S$, there is a partition $\{A_i\}_{i=1}^n \subseteq \mathcal{F}$ separating those n points, i.e. $\forall i, 1 \leq i \leq n, s_i \in A_i$.*

Proof. For the base case $n = 2$, we use Proposition 3.5 and without loss of generality there is $B_1 \in \mathcal{F}$ such that $s_1 \in B_1 \not\ni s_2$. The partition $A_1 = B_1$, $A_2 = B_1^{\mathsf{c}}$ separates $\{s_1, s_2\}$ and $\{A_1, A_2\} \subseteq \mathcal{F}$.

Suppose $\{s_i\}_{i=1}^n$ are separated by $\{A_i\}_{i=1}^n \subseteq \mathcal{F}$, and we add s_{n+1}. Again without loss of generality there are $\{B_i\}_{i=1}^n \subseteq \mathcal{F}$ such that for all $1 \leq i \leq n$, $s_{n+1} \in B_i \not\ni s_i$. The new partition is $A_i' = A_i \cap B_i^{\mathsf{c}}$ for $1 \leq i \leq n$, and $A_{n+1}' = \bigcup_{i=1}^n B_i$, is such that separates $\{s_i\}_{i=1}^{n+1}$ and $\{A_i'\}_{i=1}^{n+1} \subseteq \mathcal{F}$. □

Measurability of $\Phi_{\leq n}$ is given in the next result.

Proposition 5.2. *Given a measurable space (S, Σ) that is countably generated by $\mathcal{C}$ and separates points, then the set of all discrete measures up to n points is measurable, $\Phi_{\leq n} \in \Delta(\Sigma)$.*

Proof. Let $\mathcal{F}$ be the countable algebra generated by $\mathcal{C}$. We are going to prove that

$$\Phi_{\leq n} = \bigcap_{\substack{\{A_i\}_{i=0}^n \subseteq \mathcal{F} \\ Part(\{A_i\}_{i=0}^n)}} \left(\bigcup_{i=0}^n \Delta^{=0}(A_i) \right) \tag{5.6}$$

where $Part(\{A_i\}_{i=1}^n)$ is a predicate that is true iff $\{A_i\}_{i=1}^n$ is a partition of S. The generation of the algebra is needed to ensure the existence of disjoint sets covering S[1].

[1]For example $\mathcal{C} = \{(q, \infty) \mid q \in \mathbb{Q}\}$ is a denumerable set generating $\mathcal{B}(\mathbb{R})$ but every pair of sets in the family overlap.

For the left to right inclusion, let $\mu \in \Phi_{\leq n}$. There are $\{s_i\}_{i=1}^{j} \subseteq S$, with $0 \leq j \leq n$, having positive probability and $\mu(S \setminus \{s_i\}_{i=1}^{j}) = 0$. Given an arbitrary partition $\{A_i\}_{i=0}^{n} \subseteq \mathcal{F}$, it can be seen that $0 < |\{A_k \mid A_k \cap \{s_i\}_{i=1}^{j} = \emptyset, 0 \leq k \leq n\}|$, since the lower possible value is 1. This lower bound is attained if $j = n$ and $\forall i, 1 \leq i \leq n, s_i \in A_{f(i)}$, with an injection $f : [1..n] \to [0..n]$. Therefore $\mu \in \Delta^{=0}(A_i)$ for some i, hence $\mu \in \bigcup_{i=0}^{n} \Delta^{=0}(A_i)$, concluding that μ is in the denumerable intersection.

For the right to left inclusion, let μ belong to the intersection and $0 < \mu(S)$, since the null measure case is again handled trivially. Since $\mathcal{F}$ is countable, all the $n+1$ partitions can be enumerated by superscript j, $\{A_i^j\}_{i=0}^{n}$. Let $B_j = \biguplus\{A_i^j \mid 0 < \mu(A_i^j), 0 \leq i \leq n\}$ be the sets of the j-th partition with positive measure. Clearly $B_j^{\mathsf{c}} = \biguplus\{A_i^j \mid \mu(A_i^j) = 0, 0 \leq i \leq n\}$. It can be seen that for all j, $B_j \neq \emptyset$ and $B_j^{\mathsf{c}} \neq \emptyset$. The first follows by σ-additivity, since $0 < \mu(S) = \sum_{i=0}^{n} \mu(A_i^j)$ implies $0 < \mu(A_i^j)$ for some $0 \leq i \leq n$. The second follows since $\mu \in \bigcup_{i=0}^{n} \Delta^{=0}(A_i^j)$, so $\mu(A_i^j) = 0$ for some $0 \leq i \leq n$. We define the sequence $C_0 = B_0$, $C_{n+1} = C_n \cap B_{n+1}$. We prove inductively $\mu(C_i^{\mathsf{c}}) = 0$ for all i. The base case is $\mu(C_0^{\mathsf{c}}) = \mu(B_0^{\mathsf{c}}) = 0$ by definition of B_i. For the inductive step $\mu(C_{n+1}^{\mathsf{c}}) = \mu(C_n^{\mathsf{c}} \cup B_{n+1}^{\mathsf{c}}) \leq \mu(C_n^{\mathsf{c}}) + \mu(B_{n+1}^{\mathsf{c}}) = 0$. Therefore $\forall i, \mu(C_i^{\mathsf{c}}) = 0$, and given that $0 < \mu(S)$, it holds $\forall i, 0 < \mu(C_i)$. Given that $C_i^{\mathsf{c}} \nearrow C^{\mathsf{c}} = \bigcup_j B_j^{\mathsf{c}}$, and $\mu(C_i^{\mathsf{c}}) = 0$ for all i, it holds that $\mu(C^{\mathsf{c}}) = 0$, and $0 < \mu(C)$. The rest of the proof amounts to show that C has at most n elements. Suppose that C has at least $n+1$ elements. Then by Lemma 5.1, there is a partition j, $\{A_i^j\}_{i=0}^{n} \subseteq \mathcal{F}$ separating at least $n+1$ points. However it was shown that for all i, neither B_i nor B_i^{c} are empty, therefore B_j includes n elements at most, concluding that C cannot have $n+1$ elements. □

The measurability of $\Phi_{=n}$ is inherited from $\Phi_{\leq n}$. We can write $\Phi_{=n} = \Phi_{\leq n} \cap (\Phi_{\leq n-1})^{\mathsf{c}}$.

Observe that the set where PA probabilities lives, that is the set of all *finite discrete probabilities*, $Probs(X)$ in Segala's terms [57–59], is also $\Delta(\Sigma)$-measurable, $Probs(X) = \Delta^{=1}(S) \cap \bigcup_i \Phi_{\leq i}$.

Previous results and definitions can be useful for the Hybrid Systems community, where most of the systems live in the space $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$. This space is countably generated and separates points, therefore it satisfies the hypothesis of Proposition 5.2, and the sets $\Phi_{\leq n}$, $\Phi_{=n}$ are measurable in $\Delta(\mathcal{B}(\mathbb{R}^k))$.

5.2 Semantic Models Using NLMPs

In this section we give nontrivial examples on how NLMPs can be used to provide low-level semantics to probabilistic systems involving nondeterminism

in continuous state spaces. We give semantics to stochastic automata [15], stochastic hybrid automata [28] and the pGCL programming language [46]. Before we present some general results on measurable functions that ease the proof of the measurability of the transition functions for LMPs and NLMPs.

A deterministic measurable transition function $S \to S$ induces a measurable function $S \to \Delta(\Sigma)$.

Proposition 5.3. *Given σ-algebras Σ and $\Delta(\Sigma)$ with measurable singletons, a measurable function $\hat{T} : S \to S$ induces a measurable function $T : S \to \Delta(\Sigma)$, defined by $T(s) = \{\delta_{\hat{T}(s)}\}$.*

Proof. Let $\tilde{T}(s) = \{\hat{T}(s)\}$ be a non-probabilistic transition function similar to Definition 4.12. It is measurable since $\tilde{T}^{-1}(H_A) = \{s \mid \hat{T}(s) \in A\} = \hat{T}^{-1}(A) \in \Sigma$. By Proposition 4.7, the result follows. □

The convex combination of two measurable functions $S \to \Delta(S)$ is again a measurable function $S \to \Delta(S)$.

Proposition 5.4. *Let $\tau^1, \tau^2 : S \to \Delta^{=1}(S)$ be two measurable functions with image in the set of probability measures. The function $\tau = p\tau^1 + (1-p)\tau^2$ with $p \in [0,1]$ is also measurable, and $\tau(s)$ is a probability measure.*

Proof. It is easy to see that $\tau(s)$ is a probability measure, therefore in $\Delta(S) \cap \Delta^{=1}(S)$. Being $\tau(s)(Q)$ a linear combination of measurable functions, by Theorem 3.26 $\tau(\cdot, Q)$ is measurable, therefore by Lemma 3.41 $\tau : S \to \Delta(S)$ is also measurable. □

Observe that previous result is also valid for countable convex combinations. A countable set of measurable functions $S \to \Delta(S)$ forms a measurable function $S \to \Delta(\Sigma)$, provided that singletons are measurable in $\Delta(\Sigma)$.

Proposition 5.5. *Let $\{\tau^i\}_i$ be a denumerable set of measurable functions $S \to \Delta(S)$. If $\Delta(\Sigma)$ has measurable singletons, then the function $T(s) = \{\tau^i(s)\}_i$ from S to $\Delta(\Sigma)$ is measurable.*

Proof. Since $\Delta(\Sigma)$ has measurable singletons, $T(s) = \bigcup_i \{\tau^i(s)\} \in \Delta(\Sigma)$. Measurability of T is inherited from $\{\tau^i\}_i$ since $T^{-1}(H_\xi) = \bigcup_i (\tau^i)^{-1}(\xi)$. □

Notice that the converse is not valid. Take the measurable space $([0,1], \mathcal{B}([0,1]))$ and the measurable function $T(s) = \{\tau^1(s), \tau^2(s)\}$ from S to $\Delta(\Sigma)$. We define

$$\tau^1(s) = \textbf{if } s \in V \textbf{ then } \delta_1 \textbf{ else } \overline{\mathbf{0}}$$
$$\tau^2(s) = \textbf{if } s \notin V \textbf{ then } \delta_1 \textbf{ else } \overline{\mathbf{0}}$$

where V is a Vitali set in $[0,1]$. None of the functions are measurable since $(\tau^1)^{-1}(\Delta^{>0}(\{1\})) = V$ and $(\tau^1)^{-1}(\Delta^{>0}(\{1\})) = V^{\mathsf{c}}$, however for all $\xi \in \Delta(\Sigma)$, $T^{-1}(H_\xi) \in \{\emptyset, S\}$, therefore it is measurable.

Finally, a conditional choice of two measurable functions $S \to \Delta(\Sigma)$ with a measurable guard, is also measurable.

Proposition 5.6. *Let T^1, T^2 be two measurable functions $S \to \Delta(\Sigma)$ and $A \in \Sigma$ be a measurable set denoting a* condition*, then $T(s) =$ **if** $s \in A$ **then** $T^1(s)$ **else** $T^2(s)$ is a measurable function $S \to \Delta(\Sigma)$.*

Proof. The target of $T(s)$ is in $\Delta(\Sigma)$ since their choices $T^1(s)$ and $T^2(s)$ belong to $\Delta(\Sigma)$. Measurability of T is direct from $T^{-1}(H_\xi) = ((T^1)^{-1}(H_\xi) \cap A) \cup ((T^2)^{-1}(H_\xi) \cap A^{\mathsf{c}})$. □

The previous results provide some simple tools to compose LMPs and NLMPs.

We define deterministic LTS with a σ-algebra attached as $(S, \Sigma, \{\hat{T}_a \mid a \in L\})$, where $\hat{T}_a : S \to S$ is measurable. Proposition 5.3 embeds a measurable deterministic LTS into the NLMP $(S, \Sigma, \{T_a \mid a \in L\})$, where $T_a(s) = \{\delta_{\hat{T}_a(s)}\}$. Proposition 5.4 says that given two LMPs $(S, \Sigma, \{\tau^1_a \mid a \in L\})$, $(S, \Sigma, \{\tau^2_a \mid a \in L\})$, its convex combination $(S, \Sigma, \{p\tau^1_a + (1-p)\tau^2_a \mid a \in L\})$ is also an LMP. This can be extended to a denumerable set of LMPs and denumerable convex combinations. Proposition 5.5 is a partial extension of the LMP to NLMP embedding (Proposition 4.2). Given a denumerable set of LMPs $\{(S, \Sigma, \{\tau^i_a \mid a \in L\})\}_i$, we can define the denumerably branching NLMP $(S, \Sigma, \{T_a \mid a \in L\})$, where $T_a(s) = \{\tau^i_a(s)\}_i$. Finally Proposition 5.6 says that the conditional choice of NLMPs $(S, \Sigma, \{T^1_a \mid a \in L\})$, $(S, \Sigma, \{T^2_a \mid a \in L\})$, over a measurable condition or guard $A \in \Sigma$, i.e. the triple $(S, \Sigma, \{T_a \mid a \in L\})$, where $T(s) =$ **if** $s \in A$ **then** $T^1(s)$ **else** $T^2(s)$, is an NLMP.

Stochastic Automata Stochastic Automata [8,15,16] provide a symbolic framework to model soft real-timed systems. It can be seen as a nondeterministic extension of generalized semi-Markov processes that is amenable to composition. Stochastic Automata is usually interpreted in terms of continuous probabilistic transition systems. Here we show that such interpretation is indeed an NLMP, improving on [15] where measurability issues were not taken into account. The definition below is a simplification of that appearing in [8].

Definition 5.1. A *stochastic automaton* (SA) is a tuple $(St, Ck, Dst, Act, \rightarrow, s_0)$ where:

- St is a finite set of *control states* with $s_0 \in St$ being the *initial control state*,
- Ck is the finite set of *clock names*,
- $Dst : Ck \rightarrow \Delta(\mathbb{R}^+)$ assigns a probability measure to each clock,
- Act is a finite set of *actions*, $Act \cap \mathbb{R}^+ = \emptyset$, and
- $\rightarrow \subseteq St \times 2^{Ck} \times Act \times \Delta(2^{Ck} \times St)$ is the finite *control transition*.

We write $s \xrightarrow{C,a} \rho$ if $(s, C, a, \rho) \in \rightarrow$. A clock has two values associated: its current time (which increases synchronously with the current time of all other clocks) and a termination value (which is set randomly according to $Dst(c)$). We say that a clock has *terminated* if its current value exceeds the termination value, otherwise it is *active*. The meaning of a control transition $s \xrightarrow{C,a} \rho$ is the following. To trigger the transition, all clocks in the enabling set C must terminate, that is, the transition cannot be executed as long as a clock in C is active. Transitions are labeled with actions that can interact with the environment. We say that a clock c has *started*, when its current value is set to 0 and its termination value is set randomly according to $Dst(c)$. When a transition is executed, a probabilistic choice will take place according to ρ, where $\rho(C', s')$ is the probability that all clocks in C' are started and the system reaches the control state s'. Time is allowed to elapse as long as no control transition becomes enabled, i.e., as long as all control transitions still have some active clock.

Formally, an actual state records not only the control state $s \in St$ but also the current value and the termination value of each clock through the valuations $v \in (\mathbb{R}^+)^{Ck}$ and $e \in (\mathbb{R}^+)^{Ck}$, respectively. Hence, the set of states is $S = St \times (\mathbb{R}^+)^{Ck} \times (\mathbb{R}^+)^{Ck}$, and it is equipped with the standard product σ-algebra $\Sigma = 2^{St} \otimes (\mathcal{B}(\mathbb{R}^+))^{Ck} \otimes (\mathcal{B}(\mathbb{R}^+))^{Ck}$. The labels of an NLMP represent both the occurrence of actions in Act and the passage of time through positive reals. Then $L = Act \uplus \mathbb{R}^+$.

In the following we give some definitions to ease notation. Given $C \subseteq Ck$, by $e(C) \leq v(C)$ we denote the point-wise order relation $\forall c \in C, e(c) \leq v(c)$. For valuation v and $t \in \mathbb{R}$ we define valuation $v + t$ by $(v + t)(c) = v(c) + t$.

Before defining the nondeterministic transitions T_a, we construct probabilistic transitions following [8], which are given in the left of Table 5.1. First we define $\mu^{v,e}_{C',s'}$ a probability measure in (S, Σ) that *randomly activate clocks*

$$\frac{s \xrightarrow{C,a} \rho \quad e(C) \leq v(C)}{(s,v,e) \xrightarrow{a} \nu_\rho^{v,e}} \qquad \frac{0 < t \leq \min\{\max_{c \in C}(e(c) - v(c)) \mid \exists a, \rho : s \xrightarrow{C,a} \rho\}}{(s,v,e) \xrightarrow{t} \delta_{(s,v+t,e)}}$$

Table 5.1: Probabilistic and timed transition rules for Stochastic Automata.

in C' while moving to s' and leaving all other valuations of v, e unchanged:

$$\begin{aligned}\mu^{v,e}_{C',s'}(A \times \mathcal{I} \times \mathcal{I}') \;\doteq\; & \delta_{s'}(A) \cdot \textstyle\prod_{c \in C'} \delta_0(\mathcal{I}(c)) \cdot \prod_{c \in Ck - C'} \delta_{v(c)}(\mathcal{I}(c)) \cdot \\ & \cdot \textstyle\prod_{c \in C'} Dst(c)(\mathcal{I}'(c)) \cdot \prod_{c \in Ck - C'} \delta_{e(c)}(\mathcal{I}'(c))\end{aligned}$$

where $A \subseteq St$ and $\mathcal{I}$ and $\mathcal{I}'$ are measurable rectangles in $(\mathcal{B}(\mathbb{R}^+))^{Ck}$, i.e. $\forall c \in Ck : \mathcal{I}(c), \mathcal{I}'(c) \in \mathcal{B}(\mathbb{R}^+)$. Using Theorem 3.20, the probability measure $\mu^{v,e}_{C',s'}$ on Σ is uniquely defined by its values on the rectangles.

The *target distribution* is then defined by a finite convex combination, according to distribution ρ, of continuous probability measures $\mu^{v,e}_{C',s'} \in \Delta(S)$.

$$\nu^{v,e}_\rho \doteq \textstyle\sum_{s' \in St, C' \subseteq Ck} \rho(C', s') \mu^{v,e}_{C',s'}$$

The timed transition is defined in the right-hand side of Table 5.1. From rules in Table 5.1 we obtain that:

$$\begin{aligned} T_a(s,v,e) &= \{\nu^{v,e}_\rho \mid s \xrightarrow{C,a} \rho \wedge e(C) \leq v(C)\} \qquad \text{for } a \in Act \\ T_t(s,v,e) &= \{\delta_{(s,v+t,e)}\} \qquad \text{if } 0 < t \leq \min\{\max_{c \in C}(e(c) - v(c)) \mid \exists a, \rho : s \xrightarrow{C,a} \rho\} \end{aligned}$$

Notice that T_a, $a \in Act$, defines a finite nondeterministic choice of continuous probabilities; while T_t, where $t \in \mathbb{R}^+$, defines a deterministic change in the state. To claim that $(S, \Sigma, \{T_a \mid a \in L\})$ is indeed an NLMP we have to show that for all $a \in Act$ and $t \in \mathbb{R}^+$, $T_a(s,v,e)$ and $T_t(s,v,e)$ are in $\Delta(\Sigma)$, and T_a and T_t are measurable.

Since $S = St \times (\mathbb{R}^+)^{Ck} \times (\mathbb{R}^+)^{Ck}$ is countably generated, by Lemma 4.1 singletons in $\Delta(S)$ are measurable. Given that $T_t(s,v,e)$, $t \in \mathbb{R}^+$ and $T_a(s,v,e)$, $a \in Act$ have finite image, they are measurable. This concludes the proof of the first part.

It remains to prove that T_t and T_a are measurable. We first show the case in which $a \in Act$. For $\xi \in \Delta(\Sigma)$ we write:

$$\begin{aligned} T_a^{-1}(H_\xi) &= \{(s,v,e) \mid \{\nu^{v,e}_\rho \mid s \xrightarrow{C,a} \rho \wedge e(C) \leq v(C)\} \cap \xi \neq \emptyset\} \\ &= \textstyle\bigcup_{s \xrightarrow{C,a} \rho} \left((\nu^{v,e}_\rho)^{-1}(\xi) \cap \{(s,v,e) \mid e(C) \leq v(C)\}\right) \end{aligned}$$

Notice that the union is bounded by control transition, therefore it is finite. The set $\{(s,v,e) \mid e(C) \leq v(C)\} = S \times \{(v,e) \mid e(C) \leq v(C)\}$, where the right factor is a closed set in $(\mathbb{R}^+)^{Ck} \times (\mathbb{R}^+)^{Ck}$, therefore the product is measurable. The only proof obligation is that $\nu_\rho^{v,e}$ is measurable. Being $\nu_\rho^{v,e}$ is a finite convex sum of $\mu_{C',s'}^{v,e}$, Proposition 5.4 and the next result completes the proof.

Proposition 5.7. *Let $f_{C',s'} : S \to \Delta(S)$ be the map defined by $f_{C',s'}(s,v,e) = \mu_{C',s'}^{v,e}$. Then $f_{C',s'}$ is measurable.*

Proof. First notice that the set of rectangles generating S form a π-system. By Lemma 3.39, it suffices to prove that $f_{C',s'}^{-1}(\Delta^{>q}(Q))$ is measurable for every rectangle Q and $q \in \mathbb{Q} \cap [0,1]$. Since $Q = A \times \mathcal{I} \times \mathcal{I}'$ is a rectangle, then, for arbitrary v and e,

$$\begin{aligned} q < \mu_{C',s'}^{v,e}(Q) \text{ iff} \quad & s' \in A \\ & \wedge\ (\forall c \in C' : 0 \in \mathcal{I}(c)) \ \wedge\ (\forall c \in Ck - C' : v(c) \in \mathcal{I}(c)) \\ & \wedge\ (q < \textstyle\prod_{c \in C'} Dst(c)(\mathcal{I}'(c))) \ \wedge\ (\forall c \in Ck - C' : e(c) \in \mathcal{I}'(c)) \end{aligned}$$

As a consequence, if $s' \in A$, $(\forall c \in C' : 0 \in \mathcal{I}(c))$, and $q < \prod_{c \in C'} Dst(c)(\mathcal{I}'(c))$, $f_{C',s'}^{-1}(\Delta^{>q}(Q)) = \{(s,v,e) \mid s \in St \wedge \forall c \in Ck - C' : v(c) \in \mathcal{I}(c) \wedge e(c) \in \mathcal{I}'(c)\}$, and $f_{C',s'}^{-1}(\Delta^{>q}(Q)) = \emptyset$, otherwise. So $f_{C',s'}^{-1}(\Delta^{>q}(Q))$ is a rectangle and hence measurable. □

It remains to be shown that T_t is measurable for all $t \in \mathbb{R}^+$. We fix the time label t, and define the measurable set $B_t = \{(s,v,e) \mid 0 < t \leq \min\{\max_{c \in C}(e(c) - v(c)) \mid \exists a, \rho : s \xrightarrow{C,a} \rho\}\}$. We develop the inverse image of a generator:

$$T_t^{-1}(H_\xi) = \{(s,v,e) \mid \{\delta_{(s,v+t,e)}\} \cap \xi \neq \emptyset\} \cap B_t$$

Let $\hat{T}_t(s,v,e) = (s, v+t, e)$ be a function $S \to S$. It is measurable since $(\hat{T}_t)^{-1}(A \times \mathcal{I} \times \mathcal{I}') = A \times (\mathcal{I} - t) \times \mathcal{I}'$ and this set is measurable (see Example 3.15). By Proposition 5.3 the result follows.

In Chapter 4 we have already discussed the need for structure on the labels. In particular Example 4.10 shows that a strong correlation between continuous labels and states could give rise to measurability problems. In the following, we show that the same definition of Table 5.1 yields an NLMP with structure on the labels with $\Sigma_L = 2^{Act} \oplus \mathcal{B}(\mathbb{R}^+)$ being the σ-algebra on the set $Act \oplus \mathbb{R}^+$ of labels. (Notice that Σ_L contains all singletons.) Now,

function T is defined:

$$T(s,v,e)=\{(a,\nu_\rho^{v,e}) \mid a\in Act, s\xrightarrow{C,a}\rho, e(C)\leq v(C)\} \quad (5.7)$$
$$\uplus\,\{(t,\delta_{(s,v+t,e)}) \mid t\in\mathbb{R}^+, 0<t\leq\min\{\max_{c\in C}(e(c)-v(c)) \mid \exists a,\rho : s\xrightarrow{C,a}\rho\}\}$$

Notice that $T(s,v,e) = \big(\bigcup_{a\in Act}\{a\}\times T_a(s,v,e)\big)\uplus\big(\bigcup_{t\in\mathbb{R}^+}\{t\}\times T_t(s,v,e)\big)$. Hence $T(\cdot)_{|l} = T_l$ for all $l\in Act\oplus\mathbb{R}^+$. As a consequence $T(\cdot)_{|l}$ is measurable. It remains to show that $T(s,v,e)\in\Sigma_L\otimes\Delta(\Sigma)$. We separately show that the two sets in the union of (5.7) are measurable. The first set is finite and hence measurable since singletons sets are measurable. For the second set, we have that $\Delta(\Sigma)$, that is countably generated and separates points, is isomorphic to some Borel σ-algebra [36, Proposition 12.1]. Moreover, fixing s, v and e, the mapping $t\mapsto\delta_{(s,v+t,e)}$ is measurable. Then, by [36, Proposition 12.4], the graph defined by such measurable mapping is measurable in $\Sigma_L\otimes\Delta(\Sigma)$. Therefore, the second set of the union in (5.7) is also measurable.

The semantics of SA discussed above is what is called *residual lifetime semantics* in [8]. Although calculations were not straightforward, it should not be surprising that this semantics is an NLMP since the ingredients (a mix of finite sets and standard Borel spaces on the reals) have fine properties in measure theory. For this same reason we expect that the so called *spent lifetime semantics* [8] also can be modeled as NLMPs.

Stochastic Hybrid Automata This type of models deals with the interplay of continuous time and randomness. In the following, we describe *stochastic hybrid automata* [28], in which we allow continuous probability distributions and uncountable non-determinism in discrete assignments, yet not over continuous distributions. We show that the underlying semantics is an NLMP.

Let m denote a variable ranging over a finite set of *modes* $\mathtt{M} = \{m_1,\ldots,m_n\}$, and let $\boldsymbol{x} = (x_1,\ldots,x_k)$ be a vector of variables ranging over real numbers $\mathbb{R}$, representing the *position* of some object. The derivative of $\boldsymbol{x}$ representing the *speed* is denoted with $\dot{\boldsymbol{x}} = (\dot{x}_1,\ldots,\dot{x}_k)$, and it ranges over $\mathbb{R}^k$. With m' and $\boldsymbol{x}' = (x'_1,\ldots,x'_k)$ we denote primed versions of m and $\boldsymbol{x}$ respectively. This is subsequently used to specify values resulting from discrete jumps of a hybrid automaton.

Later on, $S = \mathtt{M}\times\mathbb{R}^k$ will denote the state-space of the semantics of the hybrid automaton. We let $\Sigma = 2^{\mathtt{M}}\otimes\mathcal{B}(\mathbb{R}^k)$ denote the product σ-algebra on the state-space. A *state-space constraint* is a set $\mathtt{s}\subseteq\mathtt{M}\times\mathbb{R}^k$ over modes and variables. A *flow constraint* is a set $\mathtt{f}\subseteq\mathtt{M}\times\mathbb{R}^k\times\mathbb{R}^k$ over the variables m, $\boldsymbol{x}$, $\dot{\boldsymbol{x}}$, that is, over mode, position and speed.

A *probabilistic guarded command* $\mathtt{c}$ is defined as

$$\mathit{condition} \to p_1 : \mathit{update}_1 + \ldots + p_n : \mathit{update}_n$$

where $1 \leq n$ denotes the cardinality of the probabilistic branching of $\mathtt{c}$ with $0 < p_i$ for $i \in [1..n]$ and $\sum_{i=1}^{n} p_i = 1$. We demand that $\mathit{condition} \in \Sigma$ is a measurable constraint over $(m, \boldsymbol{x})$, and that $\mathit{update}_i : (S, \Sigma) \to (\Sigma, H(\Sigma))$ is a non-probabilistic measurable function (like in Definition 4.12) denoting a *reset mapping* for m and $\boldsymbol{x}$ for all $i \in [1..n]$. Observe that for $i \neq j$, it could be the case that $\mathit{update}_i(m, \boldsymbol{x}) \cap \mathit{update}_j(m, \boldsymbol{x}) \neq \emptyset$. Variables not appearing primed in an update, remain unchanged.

Example 5.2.

$$\begin{aligned} m = m_1 \;\; \to \;\; & 0.2 : m' = m_2 \wedge x_1' \leq x_2 - 0.84 \\ + \;\; & 0.2 : m' = m_2 \wedge x_2 - 0.85 \leq x_1' \leq x_2 - 0.25 \\ + \;\; & 0.2 : m' = m_2 \wedge x_2 - 0.26 \leq x_1' \leq x_2 + 0.26 \\ + \;\; & 0.2 : m' = m_2 \wedge x_2 + 0.25 \leq x_1' \leq x_2 + 0.85 \\ + \;\; & 0.2 : m' = m_2 \wedge x_1' \geq x_2 + 0.84 \end{aligned}$$

is a probabilistic guarded command. It can be executed when the system is in mode m_1. With probability 1, we move to mode m_2. With probability 0.2, a particular interval is chosen, and the variable x_1 is non-deterministically set to some value within this interval. The endpoints of the intervals depend on the value of x_2. Other variables remain unchanged.

While in a previous work [68] the model was restricted to commands where each update_i maps a state to a unique successor, we here allow updates to be predicates over successor states. This leads to a possibly uncountable non-determinism, as in Example 5.2.

To model continuous measures, we introduce an additional form of guarded commands. Let $\tau : (S, \Sigma) \to (\Delta(S), \Delta(\Sigma))$ be a measurable function mapping states to probability measures, i.e. a Markov process. A *stochastic guarded command* is of the form:

$$\mathit{condition} \to \tau$$

Example 5.3. We specify $\tau(m_1, x_1, x_2, \ldots, x_n)$ as

$$\begin{aligned} &\tau(m_1, x_1, x_2, \ldots, x_n)\,(M \times \textstyle\prod_{i=1}^{n} [a_i, b_i]) \\ &\qquad = \delta_{m_2}(M) \cdot \left(\frac{1}{\sqrt{2\pi}} \int_{a_1}^{b_1} e^{-\frac{1}{2}(x - x_2)^2} dx \right) \cdot \textstyle\prod_{i=2}^{n} \delta_{x_i}([a_i, b_i]) \end{aligned}$$

and there is exactly one measure on the whole σ-algebra coinciding with this function on the rectangles (Theorem 3.20). Then $m = m_1 \to \tau$ is a stochastic guarded command, which we denote by $\mathtt{c}$. It can execute when the system is in state m_1 and, once executed, it changes to state m_2. Variable x_1 is set according to the normal distribution $\mathcal{N}(x_2, 1)$ with expected value x_2 and standard deviation 1. The other variables do not change. In practice, the normal distribution models perturbations arising from inexact measurements, deviations of production parameters in a production line, etc.

Now, we can define stochastic hybrid automata as follows.

Definition 5.4 (Stochastic Hybrid Automata)**.** A *stochastic hybrid automaton* is a tuple $\mathcal{H} = (\mathtt{M}, \boldsymbol{x}, \mathit{Init}, \mathit{Flow}, \mathtt{C}, \mathit{UnSafe})$ where

- $\mathtt{M}$ is a finite set of modes and $\boldsymbol{x}$ is a set of k variables,
- $\mathit{Init} \subseteq \mathtt{M} \times \mathbb{R}^k$ is a constraint on the initial states,
- $\mathit{UnSafe} \subseteq \mathtt{M} \times \mathbb{R}^k$ is a constrain describing the unsafe states,
- $\mathit{Flow} \subseteq \mathtt{M} \times \mathbb{R}^k \times \mathbb{R}^k$ is a flow constraint and
- $\mathtt{C}$ is a finite set of guarded commands. We denote the subset of probabilistic guarded commands as $\mathtt{C}_f$ and the subset of stochastic guarded commands as $\mathtt{C}_c$.

We require *Flow* to be measurable in the following sense: for each $m \in \mathtt{M}$, the *pre-post-relation* or *transfer relation* is:

$$T_m = \left\{ (\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^k \times \mathbb{R}^k \;\middle|\; \begin{array}{l} \exists e \geq 0, f : [0, e] \to \mathbb{R}^k \text{ differentiable} : \\ \left(\begin{array}{r} f(0) = \boldsymbol{x} \\ \wedge f(e) = \boldsymbol{y} \\ \wedge \forall t \in [0, e] : (m, f(t), \dot{f}(t)) \in \mathit{Flow} \end{array} \right) \end{array} \right\}$$

given that the continuous flow T_m is a measurable set in $\mathcal{B}(\mathbb{R}^k \times \mathbb{R}^k)$. Moreover, we require $\mathit{post}^m(\boldsymbol{x}) \doteq T_{m_{|\boldsymbol{x}}}$ to be measurable, i.e. $\mathit{post}^m : (\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k)) \to (\mathcal{B}(\mathbb{R}^k), H(\mathcal{B}(\mathbb{R}^k)))$. Furthermore, we require *Init* and *UnSafe* to be measurable sets in $2^{\mathtt{M}} \otimes \mathcal{B}(\mathbb{R}^k)$.

Notice that the definition explicitly requires that the transfer relation T given by the nondeterministic differential equations implied by *Flow* is measurable forward and backwards.

The *semantics* of a stochastic hybrid automaton $\mathcal{H} = (\mathtt{M}, \boldsymbol{x}, \mathit{Init}, \mathit{Flow}, \mathtt{C}, \mathit{UnSafe})$ is the tuple $[\![\mathcal{H}]\!] \doteq (S, \Sigma, \mathit{Init}, \mathit{Steps}, \mathit{UnSafe})$ where $S \doteq \mathtt{M} \times \mathbb{R}^k$, $\Sigma \doteq 2^{\mathtt{M}} \otimes \mathcal{B}(\mathbb{R}^k)$ and we define *Steps* as the union of two transition relations

$Steps_{\mathcal{T}}, Steps_{\mathcal{J}} : S \to \Delta(\Sigma)$. Since SHA do not consider labels, neither does our transition relation. The semantics of timed steps is defined as:

$$Steps_{\mathcal{T}}((m, \boldsymbol{x})) \doteq \{\delta_{(m,\boldsymbol{x}')} \mid \boldsymbol{x}' \in post^m(\boldsymbol{x})\}$$

Now we define the semantics of guarded commands. To start with, we define the semantics of a probabilistic guarded command $\mathtt{C} = cond \to p_1 : u_1 + \ldots + p_n : u_n$ by: $Steps_{\mathtt{c}}(s) \doteq \emptyset$ if $s \notin cond$, and otherwise:

$$Steps_{\mathtt{c}}(s) \doteq \{\textstyle\sum_{i=1}^{n} p_i \delta_{s_i} \,|\, (s_1, \ldots, s_n) \in \prod_{i=1}^{n} u_i(s)\}$$

Inside the previous formula, we have weighted sums of Dirac probability measures. A step induced by a probabilistic guarded command has as successors all discrete measures, such that with probability p_i chooses some state of the ith update. In case a state is the successor of two different updates, their probabilities are added up.

Next, for a stochastic guarded command $\mathtt{c} = cond \to \tau$ we define $Steps_{\mathtt{c}}(s) \doteq \emptyset$ if $s \notin cond$, and otherwise:

$$Steps_{\mathtt{c}}(s) \doteq \{\tau(s)\}$$

Then for $s \in S$, we let:

$$Steps_{\mathcal{J}}(s) \doteq \textstyle\bigcup_{\mathtt{c} \in \mathtt{C}} Steps_{\mathtt{c}}(s)$$

and

$$Steps(s) \doteq \begin{cases} Steps_{\mathcal{T}}(s) \cup Steps_{\mathcal{J}}(s) & \text{if } Steps_{\mathcal{T}}(s) \cup Steps_{\mathcal{J}}(s) \neq \emptyset \\ \{\delta_s\} & \text{otherwise} \end{cases}$$

The possible steps in the semantics are thus all possible transitions induced by jumps or timed transitions. Self-loops introduced using Dirac distributions are necessary to guarantee that each state has at least one successor measure.

It remains to show that the semantics is well-defined. That is, we have to show that for all $s \in S$, $Steps(s)$ is a measurable set in $\Delta(\Sigma)$ and that $Steps$ is a measurable function.

Lemma 5.8. *Let $[\![\mathcal{H}]\!] = (S, \Sigma, Init, Steps, UnSafe)$ be a tuple that is the semantics of a stochastic hybrid automaton $\mathcal{H}$. Then, $Steps : (S, \Sigma) \to (\Delta(\Sigma), H(\Delta(\Sigma)))$ is a measurable function mapping states to elements of $\Delta(\Sigma)$, that is, $[\![\mathcal{H}]\!]$ is an NLMP augmented with initial states Init and unsafe states UnSafe. In case $\mathcal{H}$ is purely probabilistic, $[\![\mathcal{H}]\!]$ is a PA.*

Proof. We handle the two parts of *Steps* separately. Namely $Steps_{\mathcal{T}}$ given by flows, and $Steps_{\mathcal{J}}$ defined by probabilistic guarded commands and stochastic guarded commands, are proven to be measurable functions $S \to \Delta(\Sigma)$. Moreover, we only consider semantics with *true* guards, since conditions different than *true* are handled by Proposition 5.6.

The proof stands on two results. First that $f : S \to \Delta(S)$ is measurable iff its uncurried version $f : S \times \Sigma \to [0, 1]$ is measurable in its first coordinate (Lemma 3.41). Second, that $\delta(s) = \delta_s$ is a measurable function $S \to \Delta(S)$.

For the case $Steps_{\mathcal{T}}$ we write $Steps_{\mathcal{T}}(s) = \{\delta_s \mid s \in post(s)\} = \delta(post(s))$. Since the underlying measurable space $(\mathtt{M} \times \mathbb{R}^k, 2^{\mathtt{M}} \otimes \mathcal{B}(\mathbb{R}^k))$ is countably generated and separates points, and for all $s \in S$, $post(s) \in \Sigma$, by Corollary 4.5, $\delta(post(s)) \in \Delta(\Sigma)$.

To show that $Steps_{\mathcal{T}}$ is measurable notice that

$$\begin{aligned} Steps_{\mathcal{T}}^{-1}(H_\xi) &= \{s \mid Steps_{\mathcal{T}}(s) \cap \xi \neq \emptyset\} \\ &= \{s \mid \delta(post(s)) \cap \xi \neq \emptyset\} \\ &= post^{-1}(H_{\delta^{-1}(\xi)}) \end{aligned}$$

Since *post* and δ are measurable, so is $Steps_{\mathcal{T}}$.

For $Steps_{\mathcal{J}}$ there are two components: probabilistic guarded commands and stochastic guarded commands. We analyze each case separately. Let $\mathsf{c} = true \to p_1 : u_1 + \ldots + p_n : u_n$ a probabilistic guarded command with n alternatives. We have to show that $Steps_{\mathsf{c}}(s) = \{\sum_{i=1}^n p_i \delta_{s_i} \mid (s_1, \ldots, s_n) \in \prod_{i=1}^n u_i(s)\}$ is a measurable set for each s and that the function $Steps_{\mathsf{c}}(s)$ is measurable $(S, \Sigma) \to (\Delta(S), H(\Delta(\Sigma)))$. It can be shown that for *disjoint* u_i, the semantics is the set of point measures in $\Phi_{\leq n}$ (5.5), such that any of those discrete measures when applied to event $u_i(s)$ is equal to p_i. We write:

$$Steps_{\mathsf{c}}(s) = \Phi_{\leq n} \cap \textstyle\bigcap_{i=1}^n \Delta^{=p_i}(u_i(s))$$

Therefore $Steps_{\mathsf{c}}(s)$ is measurable in $\Delta(\Sigma)$ by Proposition 5.2. Adding the mode component of the state space S does not raise any measurability issue, since it is a finite set having a finite number of subsets as measurable events.

If the sets $u_i(s)$ are *not disjoint*, the expression is still measurable. We show this by taking into account the intersections where two different update functions can choose the point measure, where the probabilities have to be added. There are as many conditions of intersection combinations as partitions of the set $[1..n]$. We denote the *family of partitions* of a discrete set A as $SetPart(B) \doteq \{B \mid Part(A)\}$. For example $SetPart([1..3])$ is:

$$\Big\{\big\{\{1,2,3\}\big\}, \big\{\{1,2\},\{3\}\big\}, \big\{\{1\},\{2,3\}\big\}, \big\{\{1,3\},\{2\}\big\}, \big\{\{1\},\{2\},\{3\}\big\}\Big\}$$

Here we use the measurable set $\Phi_{=n}$ of n-points measures to express the semantics of overlapping updates

$$Steps_{\mathsf{c}}(s) = \bigcup_{P \in SetPart([1..n])} \Phi_{=|P|} \cap \bigcap_{i=1}^{|P|} \Delta^{=\sum_{j \in P_i} p_j}(\bigcap_{j \in P_i} u_j(s))$$

Let $f(s_1, \ldots, s_n) = p_1\delta_{s_1} + \cdots + p_n\delta_{s_n}$ be a function in $S^n \to \Delta(S)$ generating discrete measures, and let $u(s) = (u_1(s), \ldots, u_n(s))$ be a function in $S \to S^n$ building the cross product of the nondeterministic update functions. Taking the set-wise extension of f we have that $Steps_{\mathsf{c}}(s) = f(u(s))$, for $\mathsf{c} = true \to p_1 : u_1 + \ldots + p_n : u_n$.

To show that $Steps_{\mathsf{c}}$ is measurable notice that:

$$\begin{array}{ll} & Steps_{\mathsf{c}}^{-1}(H_\xi) \\ = & \{s \mid f(u(s)) \cap \xi \neq \emptyset\} \\ = & \{s \mid \exists (s_1, \ldots, s_n) \in u(s), f(s_1, \ldots, s_n) \in \xi\} \\ = & u^{-1}(H_{f^{-1}(\xi)}) \end{array}$$

The uncurried version of f, namely $f(s_1, \ldots, s_n, Q)$, is a convex combination of measurable functions $(\lambda s : \delta_s)$, hence by Proposition 5.4, $f : S \to \Delta(S)$ is measurable. The function u is measurable since its components are measurable (Proposition 3.25). Therefore we conclude that $Steps_{\mathsf{c}}^{-1}(H_\xi)$ is a measurable set.

For the stochastic guarded command component of $Steps_{\mathcal{J}}$, notice that $Steps_{\mathcal{J}}(s) = \{\tau(s)\}$. By Proposition 4.2 $Steps_{\mathcal{J}}(s)$ is measurable since $\tau(s)$ is measurable. Notice that, by Lemma 4.1, singletons $\{\mu\}$ are measurable in $\Delta(\Sigma)$ since the reals are generated by a denumerable π-system. □

Probabilistic Guarded Command Language Although Probabilistic Guarded Command Language (pGCL) [46] semantics is not given in terms of NLMPs, their semantics show interesting concepts. Therefore it is worth checking if the semantics of pGCL can be captured by an NLMP. pGCL is a probabilistic extension of the Dijkstra's Guarded Command Language (GCL) [24]. It features, besides nondeterministic choice $P_1 \sqcap P_2$, a discrete probabilistic choice in the form of $P_1 \oplus_p P_2$, where program P_1 is executed with probability p and P_2 with probability $1 - p$. The semantics of pGCL is given as a state transformer function $S \to 2^{\Delta(S)}$. Such a state transformer represents a nondeterministic choice of probabilistic behaviors, on a *discrete*

state space S. However not every set of distributions is appropriate. [46] considers only non-empty, up-closed, convex-closed, and Cauchy-closed sets. We revisit *up-closed* and *convex-closed* properties since they are similar to other probabilistic and nondeterministic models.

The semantics of pGCL also considers subprobabilities. Given a subprobability measure μ, the "missing" part of the probability, i.e. $1 - \mu(S)$, is interpreted as the probability that the program aborts. That is, if the measure given by a probabilistic and nondeterministic program over the whole state space is less than one, the remaining part of the behavior is unpredictable. This is consistent with the non-probabilistic case, where the semantics of the aborting program is every possible final state. In the probabilistic semantics, every probability measure that is event-wise greater than the sub-probability measure is possible, so the semantics of a pGCL program should be *up-closed*.

Definition 5.5. A set of distributions $\Omega \subseteq \Delta(S)$ is *up-closed* if $\mu \in \Omega$, and $\mu \leq \mu'$, then $\mu' \in \Omega$, where $\leq$ is lifted to measures by the point-wise ordering, that is, $\mu \leq \mu'$ *iff* $\forall A \in \Sigma, \mu(A) \leq \mu'(A)$. In other words, a set of subprobability distributions is up-closed if it is $\leq$ -closed (2.3).

Discrete nondeterminism in pGCL is semantically treated as all the convex combinations of nondeterministic choices, so that probabilistic choice is a refinement of nondeterministic choice. For example if the semantic interpretation of a program is given by two probability distributions $\{\mu_1, \mu_2\}$, the semantics should also include every convex combination $p\mu_1 + (1-p)\mu_2$.

Definition 5.6. A set of distributions $\Omega \subseteq \Delta(S)$ is *convex-closed* if for every $\mu_1, \mu_2 \in \Omega$, the measure $p\mu_1 + (1-p)\mu_2$ is also in Ω, for every $p \in [0, 1]$.

The semantics $T_P : S \to 2^{\Delta(S)}$ of relevant program constructors P of pGCL is given in Table 5.2.

$$
\begin{aligned}
T_{P_1;P_2}(s) &\doteq \{\textstyle\int T_{P_2}(s')d\mu(s') \mid \mu \in T_{P_1}(s)\} \\
T_{P_1 \oplus_p P_2}(s) &\doteq \{p\mu_1 + (1-p)\mu_2 \mid \mu_1 \in T_{P_1}(s), \mu_2 \in T_{P_2}(s)\} \\
T_{P_1 \sqcap P_2}(s) &\doteq \textstyle\bigcup_{p\in[0,1]} T_{P_1 \oplus_p P_2}(s)
\end{aligned}
$$

Table 5.2: Semantics of sequential, probabilistic choice and nondeterministic choice constructors of pGCL [46].

We are going to show that the semantic function T_P is in $S \to \Delta(\Sigma)$. Instead of proving it by induction on the program structure, we will use that

the semantics of pGCL is Cauchy-closed together with the following result. It states that there is a bijection between certain measurable sets in $[0,1]^n$ and the measurable sets of discrete subprobability measures over state space of n elements.

Lemma 5.9. *Let $S = \{s_i\}_{i=1}^n$ be a finite state space, and let $P = \{(x_1,\ldots,x_n) \mid \sum_{i=1}^n x_i \leq 1\} \in \mathcal{B}([0,1]^n)$ be a measurable subspace of $[0,1]^n$. We define the bijection $f : P \to \Delta(S) \cap \Delta^{\leq 1}(S)$, that sends a point of a subspace of $[0,1]^n$ to a discrete subprobability distribution, i.e. $f((x_1,\ldots,x_n)) = \{s_i \mapsto x_i\}_{i=1}^n$. If f is lifted to sets, then for all $B \in \mathcal{B}([0,1]^n)|P$, $f(B) \in \Delta(2^S)|\Delta^{\leq 1}(S)$. Moreover, if $\xi \in \Delta(2^S)|\Delta^{\leq 1}(S)$, then $f^{-1}(\xi) \in \mathcal{B}([0,1]^n)|P$.*

Proof. First observe that $f(P) = \Delta^{\leq 1}(S)$ and $f^{-1}(\Delta^{\leq 1}(S)) = P$. The generators in $\mathcal{B}([0,1]^n)|P$ are $(\prod_{i=1}^n A_i) \cap P$ with $A_j = (q_j, \infty)$, $1 \leq j \leq n$, and $\forall i, 1 \leq i \leq n, i \neq j, A_i = [0,1]$, that corresponds exactly to the generators $\Delta^{>q_j}(\{s_j\}) \cap \Delta^{\leq 1}(S)$. If f is lifted to sets, then it preserves set operations. As f^{-1} also preserve set operations by definition, we conclude that the lifting of f to sets establishes a bijection between measurable sets in $\mathcal{B}([0,1]^n)|P$ and measurable sets in $\Delta(2^S) \cap \Delta^{\leq 1}(S)$. □

In pGCL Relational Semantics [46, Section 5.4], it is shown that the semantics given in Table 5.2, $T_P : S \to 2^{\Delta(S)}$, is Cauchy-closed, and this means that for all $s \in S$ and pGCL program P, $f^{-1}(T_P(s))$ is a closed set in $\mathcal{B}([0,1]^n)$. By Lemma 5.9 this means that $T_P(s) \in \Delta(2^S)$. Measurability of the transition function is direct since S is finite.

5.3 Similar Models

Here we revisit similar transition systems models that capture continuous nondeterminism of probabilistic choices. For the sake of completeness, we first show that PAs are captured by NLMPs. The recent abstract probabilistic automata [19] also fit in the NLMP model. Finally we consider the infimum labeled Markov processes (infLMPs) [22].

Probabilistic Automata The encoding of PA (Definition 2.8) to NLMPs is direct, the image of a transition is the set of all related measures, Given a PA $(S, L, \to)$ like in Definition 2.8, we define the triple $(S, 2^S, \{T_a \mid a \in L\})$ where

$$T_a(s) = \{\mu \mid s \xrightarrow{a} \mu\}$$

Since the state space S and the transition relation $\to$ are countable sets, the powerset σ-algebra 2^S is countably generated and by Lemma 4.1 $\{\mu\} \in \Delta(2^S)$. Therefore $T_a(s) = \bigcup\{\mu \mid s \xrightarrow{a} \mu\} \in \Delta(2^S)$. Measurability of T_a is direct since in 2^S every set is measurable. We conclude that the triple $(S, 2^S, \{T_a \mid a \in L\})$ is an NLMP.

There are two important aspects of the definition of PA that impact the translation. One is the cardinality of the state space, while the other is the cardinality of the transition relation.

It is worth noting that in some definitions of PA the state space and the transition relation are not restricted, and there the measurability problems previously shown in Example 4.5 and Example 4.6 can arise.

It is interesting to remark that even if we restrict to finite state space S, and labels L, but the *transition relation is unrestricted* there could be measurability problems. The next example shows this.

Example 5.7. Let the PA $(\{s_0, s_1\}, \{a\}, \to)$ where the transition relation is defined

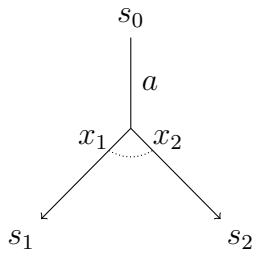
$$s_0 \xrightarrow{a} \{p\delta_{s_0} + (1-p)\delta_{s_1} \mid p \in V\} \tag{5.8}$$

and V a Vitali set in $[0, 1]$. We will show that the translation $T_a(s_0) = \{\mu \mid s \xrightarrow{a} \mu\} \notin \Delta(\Sigma)$, where $\Sigma = 2^{\{s_0, s_1\}}$.

Let $S = \{s_0, s_1\}$ and notice that by Lemma 5.9, there is a measurable bijection f between $\Delta(S) \cap \Delta^{=1}(S)$ and the antidiagonal $\{(x, 1-x) \mid x \in [0, 1]\}$. The function $ad(x) = (x, 1-x)$ is measurable, therefore for all $\xi \in \Delta(\Sigma)$, $ad^{-1}(f^{-1}(\xi)) \in \mathcal{B}([0, 1])$. However $ad^{-1}(f^{-1}(T_a(s))) = V$, hence $T_a(s_0) \notin \Delta(\Sigma)$, and the above PA cannot be encoded as an NLMP.

Abstract Probabilistic Automata Abstract Probabilistic Automata (APA) [19] presents a symbolic framework that allows for the partial representation of PA. An APA is a structure $(S, L, \to)$ with a finite set of states $S = \{s_i\}_{i=1}^n$, a finite set of labels L, and a finite transition relation $\to \subseteq S \times L \times C(S)$, where $C(S)$ is the set of finite Boolean expressions with atoms being linear constraints with variables in $\{x_i\}_{i=1}^n$. Each constraint in $C(S)$ should contain the equality $\sum_{i=1}^n x_i = 1$, thus defining that each possible solution is also a probability distribution on S. (Notice that we do not consider may and must modalities.) In Figure 5.4 a simple APA is shown.

A given PA is an *implementation* of an APA if for all abstract transitions $s \xrightarrow{a} \varphi$ in the APA, there is a concrete transition $s \xrightarrow{a} \mu$ in the PA, such that $\mu = \{s_i \mapsto x_i\}_{i=1}^n$ and $(x_1, \ldots, x_n) \models \varphi$.



$$0.9 \leq x_1 \wedge x_0 + x_1 = 1$$

Figure 5.4: A simple abstract probabilistic automata (APA).

We now show some examples on how $\Delta(2^S)$ is able to capture the *constraint specification language* for discrete probability measures. The simple example from Figure 5.4 defines the set of measures $\{\{s_1 \mapsto x_1, s_2 \mapsto x_2\} \mid 0.9 \leq x_2 \wedge x_1 + x_2 = 1\}$. This set can be expressed with $\Delta(2^S)$ generators as $\Delta^{\geq 0.9}(\{s_2\}) \cap \Delta^{=1}(\{s_1, s_2\})$. Table 5.3 shows two more translations from constraint specification language for discrete probabilities found in [19] to measurable sets in $\Delta(2^S)$.

APA Constraint
$0.7 \leq x_2 + x_3 \wedge 0.2 \leq x_4 + x_5 \wedge x_2 + x_3 + x_4 + x_5 = 1$
$\Delta(2^S)$ **expression**
$\Delta^{\geq 0.7}(\{s_2, s_3\}) \cap \Delta^{\geq 0.2}(\{s_4, s_5\}) \cap \Delta^{=1}(\{s_2, s_3, s_4, s_5\})$
APA Constraint
$0 \leq x_2 \leq 0.5 \wedge 0.2 \leq x_3 \leq 0.7 \wedge 0 \leq x_4 \leq 0.5 \wedge 0.4 \leq x_2 + x_3 \leq 0.8$ $\wedge x_2 + x_3 + x_4 = 1$
$\Delta(2^S)$ **expression**
$\Delta^{\leq 0.5}(\{s_2\}) \cap \Delta^{\geq 0.2, \leq 0.7}(\{s_3\}) \cap \Delta^{\leq 0.5}(\{s_4\}) \cap \Delta^{\geq 0.4, \leq 0.8}(\{s_2, s_3\})$ $\cap \Delta^{=1}(\{s_2, s_3, s_4\})$

Table 5.3: Translation of APA constraints of discrete measures to $\Delta(2^S)$ measurable sets.

We now formalize the syntax and semantics of the constraint language. The constraint language is a Boolean algebra with atoms in linear (in)equalities:

$$\varphi ::= \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \neg\varphi \mid \mathit{LinIneq}(\{x_i\}_{i=1}^n)$$
$$\mathit{LinIneq}(\{x_i\}_{i=1}^n) ::= c_0 \bowtie \sum_{i=1}^n c_i x_i$$

where $\{c_i\}_{i=0}^n \subseteq \mathbb{R}$, and $\bowtie \in \{>, <, \geq, \leq, =\}$. The satisfaction relation is given by $[\![\varphi]\!] = \{(x_1, \ldots, x_n) \mid (x_1, \ldots, x_n) \models \varphi\} \subseteq \mathbb{R}^n$ is defined below:

$$[\![\varphi_1 \wedge \varphi_2]\!] = [\![\varphi_1]\!] \cap [\![\varphi_2]\!] \qquad [\![\varphi_1 \vee \varphi_2]\!] = [\![\varphi_1]\!] \cup [\![\varphi_2]\!] \qquad [\![\neg\varphi]\!] = [\![\varphi]\!]^{\mathsf{c}}$$
$$[\![c_0 \bowtie \textstyle\sum_{i=1}^n c_i x_i]\!] = \{(x_1, \ldots, x_n) \mid c_0 \bowtie \textstyle\sum_{i=1}^n c_i x_i\}$$

Given a APA $(S, L, \to)$, we define the triple $(S, 2^S, \{T_a \mid a \in L\})$, where the transition function is defined:

$$T_a(s) = \{\mu \mid s \xrightarrow{a} \varphi \wedge \mu \models \varphi\}$$

Given that $[\![\varphi]\!]$ is a measurable set in $\mathcal{B}([0,1]^n)$ (unions, intersection and complements of open and closed sets), by Lemma 5.9, $T_a(s) \in \Delta(2^S)$. Since S is finite, measurability of T_a is direct.

Infimum Labeled Markov Process This model presented in [22] departs from the PA line giving a completely different approach for subspecification of continuous probabilistic systems. Its definition is a minor variation of the LMP [20]. It only relaxes the σ-additivity of the transition subprobability to *super-additivity*, and here is where the nondeterminism of probabilistic choices is captured.

Definition 5.8 (Super-additivity)**.** A set function f is *super-additive* if $f(A) + f(B) \leq f(A \uplus B)$.

The characterization of this subspecification in probabilistic choice is the following.

Definition 5.9. Given super-additive function f, the set of all *(sub)probabilistic realizations* in a measurable space (S, Σ) are the point-to-point greater (sub)probability measures, i.e.

$$\Theta_f \doteq \{\mu \in \Delta(S) \mid \forall Q \in \Sigma, f(Q) \leq \mu(Q)\} \tag{5.9}$$

In a nutshell the gap between $f(A) + f(B)$ and $f(A \uplus B)$ is not known and every σ-additive probability measure point-to-point greater is possible.

Example 5.10. Let f be defined on 2^S with $S = \{s_1, s_2, s_3, s_4\}$ as follows.

$$f(\{s_1\}) = f(\{s_2\}) = f(\{s_3\}) = 1/8$$
$$f(\{s_1, s_2\}) = 2/3 \qquad f(\{s_1, s_3\}) = f(\{s_2, s_3\}) = 1/4$$
$$f(\{s_1, s_2, s_3\}) = 1$$

Clearly f is super-additive. Let μ be a discrete probability measure defined by $\mu(\{s_1\}) = \mu(\{s_2\}) = \mu(\{s_3\}) = 1/3$. Then $\mu \in \Theta_f$. Moreover notice that

$$\Theta_f = \{\mu \mid 1/8 \leq \mu(\{s_i\}), i \in [1..3], \text{ and } 1/3 \leq \mu(\{s_1, s_2\})\}$$

is the (dense) set of probabilistic realizations of f.

The formal definition of infLMPs is as follows.

Definition 5.11 (Infimum labeled Markov process)**.** An *infimum Labeled Markov Process* (infLMP) [22] is a triple $(S, \Sigma, \{\tau_a \mid a \in L\})$ where Σ is a σ-algebra on the set of states S, $\tau_a : S \times \Sigma \to [0, 1]$ is a function such that $\tau_a(s, \cdot)$ is a super-additive and $\tau_a(\cdot, Q)$ is measurable, for each label $a \in L$, state $s \in S$ and measurable set $Q \in \Sigma$.

Not surprisingly most of the definitions, theorems and proofs behind LMPs remain the same for infLMPs, simply because σ-additivity is not required on those places. Super-additivity allows subspecification of LMPs as the following examples show.

Example 5.12. Let the measurable space (S, Σ) and super-additive function f as in Example 5.10. We define the infLMP $(S, 2^S, \{\tau_a\})$ partially represented in Figure 5.5, where $\tau_a(s_0) = f$. Observe not only it encodes lower bounds on probabilities, but also it is encoding *upper bounds* on probabilities. For example the singleton $\{s_3\}$ is explicitly lower bounded but also implicitly upper bounded by $\tau_a(s_0, \{s_1, s_2\}) = 2/3$ and $\tau_a(s_0, S) = 1$, so we have $1/8 \leq \tau_a(s_0, \{s_3\}) \leq 1/3$. An LMP realization of $(S, 2^S, \{\tau_a\})$ is, for example, $(S, 2^S, \{\tau'_a\})$, where $\tau'_a(s_0, \{s_i\}) = 1/3$ for $i = 1, 2, 3$.

Example 5.13. Consider an infLMP $(S, \Sigma, \{\tau_a\})$ with transition function τ_a such that $\tau_a(s, A) = 0$, $\tau_a(s, B) = 0$ and $\tau_a(s, A \uplus B) = 1$. The transition function $\tau_a(s, \cdot)$ is super-additive. This means that from s and through label a, it is certain that either A or B will occur, but it is not known the probability of each individual event. It does not mean that event A or B are individually impossible, but rather that under a demonic view of nondeterminism their chances are null. Notice that this example also shows that infLMPs can encode pure nondeterminism.

Given Definition 5.9, the translation from infLMP $(S, \Sigma, \{\tau_a \mid a \in L\})$ to an NLMP seems direct. We define the triple $(S, \Sigma, \{T_a \mid a \in L\})$, where $T_a(s) = \Theta_{\tau_a(s)}$. The next proposition shows that if the underlying σ-algebra is countably generated, the set $T_a(s)$ is measurable.

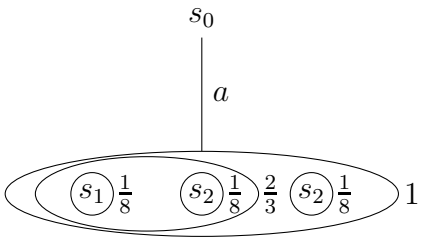


Figure 5.5: infLMP showing probabilistic subspecification of probabilistic choice.

Proposition 5.10. *Given measurable space (S, Σ) and super-additive set-wise function f on Σ, if Σ is countably generated by $\mathcal{C}$, then the set Θ_f a measurable set in $\Delta(\Sigma)$.*

Proof. Let $\mathcal{F}$ be the denumerable algebra generated by $\mathcal{C}$. We will show that

$$\Theta_f = \bigcap_{Q \in \mathcal{F}} \Delta^{\geq f(Q)}(Q)$$

Left to right inclusion is direct since the inequality holds for every measurable set, in particular for the sets in $\mathcal{F}$.

For the other inclusion suppose $\mu \notin \Theta_f$. Therefore, there exists $Q \in \Sigma$ such that $\mu(Q) < f(Q)$. Using the approximation result of Corollary 3.22 there is $Q' \in \mathcal{F}$ with $Q \subseteq Q'$ such that $\mu(Q) \leq \mu(Q') < f(Q) \leq f(Q')$. Therefore Q' witnesses that μ is not in the right-hand side. □

The remaining proof obligation, namely that T_a is a measurable function, cannot be deduced from the measurability of the infLMP transition function τ_a. Therefore it is still not clear if infLMPs can be embedded into NLMPs.

For the other inclusion, an NLMP with a simple discrete nondeterminism cannot be embedded in an infLMP. Consider the NLMP $(\{0,1\}, 2^{\{0,1\}}, \{T_a\})$, where $T_a = (\Delta^{=1/3}(\{0\}) \cap \Delta^{=2/3}(\{1\})) \cup (\Delta^{=2/3}(\{0\}) \cap \Delta^{=1/3}(\{1\}))$ (observe this is basically a PA). The best attempt to capture that discrete nondeterminism over discrete state space with a super-additive function f would be: $f(\{0\}) = f(\{1\}) = 1/3$ and $f(\{0,1\}) = 1$. However this super-additive function includes realizations like $\mu(\{0\}) = \mu(\{1\}) = 1/2$, that are not present in the former NLMP. We can say that NLMPs can express subspecifications of probabilistic choices in a finer way.

Observe that in [22] a translation from PA to infLMPs that is *simulation preserving* is shown. The authors do not claim that PA can be embedded in infLMPs.

5.4 Concluding Remarks

An important portion of this chapter is unpublished material. Section 5.1 is new, and for Section 5.2 although the SHA semantics main result was in [28, Lemma 3], its proof was not part of the final published version, therefore we present it here in full detail. Section 5.3 is also completely new.

The exercise of modeling underspecified systems using $\Delta(\Sigma)$ was fruitful in many aspects. It gives us confidence in the model and it also augments the well-known sets of measures that live in $\Delta(\Sigma)$. Discrete measures were formally shown to be in $\Delta(\Sigma)$. For more inspiring examples of measurable sets that are not trivial to define, please refer to [36, Examples, p.69–70] where rather complex sets of functions are shown to be measurable. The use of $\Delta^{\bowtie q}(Q)$ and set operators also triggered some ideas on symbolic model checking for NLMPs. We will discuss them in the conclusions.

The semantics of soft real-time systems modeled through SA was developed using NLMPs. We proved it using simple and easily reusable tools. Since the set of labels is dense to represent time passage, we also showed that the semantics was conforming the requirement on an NLMP with structure on the labels. It must be remarked that Lemma 3.39 was fundamental in the core of the proofs of measurability.

The semantics of a hybrid system featuring continuous nondeterminism was also captured by NLMPs. This proof was important since the definition of the hit σ-algebra was extensively used to show measurability of the construction. In doing this we gained more confidence in the way NLMPs was defined since it allowed us to capture a rather complex model involving continuous nondeterminism and continuous probabilities. It would be desirable that a measurable *Flow* restriction relation for the nondeterministic differential equations renders a measurable transfer function $post^m : (\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k)) \to (\mathcal{B}(\mathbb{R}^k), H(\mathcal{B}(\mathbb{R}^k)))$. A result like this could not be worked out, and the Hybrid Systems literature does not seem to deal with it in full generality. The classical approach is dealing with a vector field and ordinary differential equations (ODEs), that is deterministic equations. There, *Lipschitz continuity*, is a sufficient condition for existence and uniqueness of a solution [6]. In the presence of nondeterminism in the vector field (*differential inclusion* as stated in [6]), the existence, uniqueness or even the analysis of the transfer relation are not taken into account, see for example [55] that uses a similar *Flow* relation. Others like [35] derive measurability of the transfer function in terms of a measurable flow restriction that is inherently nondeterministic. However the flow restrictions must be rectangular regions and it is only valid for unidimensional trajectories. The basic idea is using

the mean value theorem so that the pair $(\boldsymbol{x}, \boldsymbol{y})$ is in the flow relation if and only if $(\boldsymbol{x}-\boldsymbol{y})/e \in \mathit{Flow}$. An approach yet to be explored is to export results from the field of Differential Topology [53]. There each point of the space is associated with a *cone*, a subspace that is closed under vector sums. Cones are functions of the position and they restrict the future trajectories, but they are intrinsically nondeterministic. In this setting, one particular problem is defining the region of a hyperplane the particle *trips* could hit. The similarities are striking, but the barriers of the uses and terminology between the two fields have to be broken.

We also showed that pGCL [46] semantics can be captured by NLMPs. Although the state space is discrete, its semantics deals with continuous nondeterminism in the form of up-closed and convex-closed sets of subprobability measures. We use Lemma 5.9 relating the space $[0,1]^n$ with subprobability distributions in a finite space in order to prove that the transfer function has measurable image. This lemma is a valuable tool to attack general problems of continuous nondeterminism in discrete probability spaces.

NLMPs can also capture the semantics of structures like Markov set-chains [34] and timed automata augmented with discrete probabilistic choice [39, 40]. They are not included here since they are not essentially different from what has been already shown.

A comparison with similar models was given in Section 5.3. First we discuss how to capture PA. Although for finite state space, labels and transition relation the embedding is direct, we show in Example 5.7 that just allowing arbitrary transition relations, measurability problems arise. This is, to the author's knowledge, the first time that PA shows a definability problem, and this is not a minor issue since no scheduler can quantify over this non-measurable nondeterminism. We also encode APA in NLMPs, and show how to use $\Delta^{\bowtie q}(Q)$ to capture some particular examples. Here again Lemma 5.9 proved useful to show that the embedding is an NLMP. infLMPs present a promising alternative for the subspecification of continuous probabilistic systems. It is really interesting in many aspects. The gaps given by super-additivity are the source of nondeterminism, and from pure nondeterministic choice to pure probabilistic choice, it seems that every intermediate possibility is captured in a grain that is as fine or coarse as needed. Some more work on infLMP model is needed in order to establish definitions and results concerning the resolution of this continuous nondeterminism. Schedulers and path probability definitions are missing in the original work [22]. Some more comparisons on the relative merits of the two forms of probabilistic subspecification (super-additivity vs. $\Delta^{\bowtie}(Q)$ generators) are still needed. A first impression is that infLMPs is not well suited to (sub)specify systems,

however it has a really simple definition that could render proofs that are shorter than in the NLMP counterpart. We expect that further work on infLMPs can answer these questions and show how infLMPs behaves as a modeling tool for probabilistic and nondeterministic continuous systems.

Even though stochastic transition systems (STS) [11,12] are presented as the continuous counterpart of PA, we decided not to include them in this chapter. STS departs from NLMPs and related models in one fundamental aspect, instead of restricting the transition relation, they limit the power of schedulers to take into account only those systems rendering a sound definition of path measure. They attack measurability problems in the semantic part of the model. The referred work does not show any concrete example of a system being an STS. Besides the main characterization result [12, Proposition 1], does not provide any insight on what conditions render a so-called *measurable scheduler* in the system itself. In that work, the comparison with established work like LMPs, also falls short. It is nowhere defined why the conditions on the definition or in the scheduler implies that an STS encoding of an LMP sends back measurable sets to measurable sets, a fundamental property for the semantics of any modal operator.

Chapter 6

Bisimulations and Logics

In the previous two chapters we defined the NLMP model and gave enough evidence of its applicability in the realm of real-time and hybrid systems, as well as showing through examples that the generators of $\Delta(\Sigma)$ provide adequate building blocks for probabilistic underspecification. In this chapter we extend the notions of bisimulation for LMPs [21] and PA [58] to our model. We give different generalizations of bisimulation to continuous nondeterminism and compare them exhaustively.

The work is based on the ideas of traditional bisimulation for discrete probabilistic systems [41] and LMPs [21]. We also took the concept of event bisimulation [14] for LMPs. These ideas give rise to the three notions we present here, two of which are an extension of traditional and event bisimulation, while the other is new. Traditional bisimulation is an extension of [21, 41], and it is strongly point-wise. Event bisimulation is an extension of [14], and its definition is done completely in terms of measurable sets and functions. The new notion is called *state bisimulation*. This definition is in the middle of traditional and event bisimulation, introducing a novel mix of point-wise and event-wise definitions.

We also extend the simple modal logic that characterizes bisimulation for LMPs, to finite and continuous nondeterminism, showing that this logic characterizes event bisimulation, but also state and traditional bisimulation for some restricted spaces.

In order to give a more amenable organization to the chapter, we first provide a detailed explanation of the bisimulations and the logic on LMPs (Section 6.1). Occasionally, we provide revisited proofs of the results. Once we explain the theory on this simpler framework, we introduce our results generalizing the theory of bisimulations and logic to NLMPs (Section 6.2).

6.1 Bisimulations and Logics in LMPs

In this section we give a summary of bisimulation and its logical characterization for LMPs. This work is based on [13, 14, 17, 18, 20, 21]. First we introduce bisimulation in two flavors, traditional point-wise bisimulation [21] and the relatively new event-wise bisimulation [14]. Then, we give the logical characterization of these bisimulations in terms of a Hennessy-Milner-like logic.

Relations, Measures, and σ-algebras. Given that bisimulation is a relation and LMPs are tied to σ-algebras, here we deal with some definitions and results that relate σ-algebras and relations.

Definition 6.1 (R-closed)**.** Given a relation $R \subseteq S \times S$, the predicate *R-closed*(Q) is used to denote $R(Q) \subseteq Q$, where $R(Q) \doteq \{t \mid s \in Q, s\ R\ t\}$. Notice that if R is symmetric, R-closed(Q) iff $\forall s, t, s\ R\ t, s \in Q \Leftrightarrow t \in Q$.

Given a symmetric relation R and a σ-algebra Σ, we can define the sub-σ-algebra of R-closed sets.

Definition 6.2. Let (S, Σ) be a measurable space and let R be a symmetric relation. We define $\Sigma(R) \doteq \{Q \in \Sigma \mid R\text{-closed}(Q)\}$.

Observe that $\Sigma(R)$ is the sub-σ-algebra of Σ containing all R-closed Σ-measurable sets.

It is important to remark that requiring that R is symmetric is sufficient for $\Sigma(R)$ to be complement-closed[1], that is $\forall Q \in \Sigma, R(Q) \subseteq Q \Rightarrow R(Q^c) \subseteq Q^c$ holds if $R = R^{-1}$. The converse is not valid, for example $\Sigma = \{\emptyset, \{1, 2\}\}$ and $R = \{(1, 2)\}$, showing there are weaker notions to fulfill the requirement that $\Sigma(R)$ is a σ-algebra. We stick to $R = R^{-1}$.

The next proposition states that the inclusion order between two relations transfers inversely to the σ-algebras induced by them and to the σ-algebra of measures applied to these σ-algebras.

Proposition 6.1. *Let R and R' be symmetric relations such that $R \subseteq R'$. Then*

i. $\Sigma(R) \supseteq \Sigma(R')$ *and*

ii. $\Delta(\Sigma(R)) \supseteq \Delta(\Sigma(R'))$.

[1]This point is missing in some publications about LMPs [14, 20, 21].

Proof. (i) follows from the fact that any measurable set that is R'-closed is also R-closed whenever $R \subseteq R'$. For (ii), recall that $\Delta(\Sigma(R'))$ is generated by $\mathcal{A}' = \{\Delta^{>q}(Q) \mid q \in \mathbb{Q}^+, Q \in \Sigma(R')\}$. Since $\Sigma(R') \subseteq \Sigma(R)$ (by (i)), then $\mathcal{A}' \subseteq \Delta(\Sigma(R))$ from which the proposition follows. □

We can lift R to an equivalence relation in $\Delta(S)$ as follows:

$$\mu \; R \; \mu' \quad \text{iff} \quad \forall Q \in \Sigma(R), \mu(Q) = \mu'(Q)$$

Then, the notion of R-closed can be defined on subsets of $\Delta(S)$ just like before. The following proposition will be useful.

Proposition 6.2. *If R is a symmetric relation, every $\Delta(\Sigma(R))$-measurable set is R-closed.*

Proof. We show this using good sets principle with $\mathcal{G} = \{\xi \in \Delta(\Sigma(R)) \mid R\text{-closed}(\xi)\}$. For $\Delta^{>q}(Q)$ with $q \in \mathbb{Q}^+$ and $Q \in \Sigma(R)$ it holds that $\mu \; R \; \mu' \Rightarrow (\mu \in \Delta^{>q}(Q) \Leftrightarrow \mu' \in \Delta^{>q}(Q))$. Therefore $\Delta^{>q}(Q)$ is R-closed.

Moreover, for any symmetric R, the property of being R-closed is preserved by denumerable union and complement. Since the lifted R is symmetric, we can conclude that every measurable set in $\Delta(\Sigma(R))$ is R-closed. □

A σ-algebra induces a relation in the sense that two elements cannot be distinguished iff they cannot be separated by any measurable set.

Definition 6.3. A σ-algebra Σ defines an *equivalence relation* $\mathcal{R}(\Sigma)$ as follows:

$$s \; \mathcal{R}(\Sigma) \; t \quad \text{iff} \quad \forall Q \in \Sigma, s \in Q \Leftrightarrow t \in Q$$

If Σ is a σ-algebra that separates points then $\mathcal{R}(\Sigma)$ is the identity. It is also relevant to know that the inclusion of σ-algebras transfer inversely to the relation.

Proposition 6.3. *Let Λ and Λ' be two σ-algebras such that $\Lambda \subseteq \Lambda'$. Then $\mathcal{R}(\Lambda) \supseteq \mathcal{R}(\Lambda')$.*

The following properties (due to [14]) appear here for the sake of completeness; they relate σ-algebras and relations. In particular, (v) is a consequence of (i) and (ii).

Proposition 6.4. *Let (S, Σ) be a measurable space, R a symmetric relation on S, and $\Lambda \subseteq \Sigma$ a sub-σ-algebra of Σ. Then,*

i. $\Lambda \subseteq \Sigma(\mathcal{R}(\Lambda))$;

ii. $R \subseteq \mathcal{R}(\Sigma(R))$;

iii. if each R-equivalence class is in Σ, then $R = \mathcal{R}(\Sigma(R))$;

iv. $\mathcal{R}(\Lambda) = \mathcal{R}(\Sigma(\mathcal{R}(\Lambda)))$*; and*

v. $\Sigma(R) = \Sigma(\mathcal{R}(\Sigma(R)))$[2].

In [20,21], a notion of behavioral equivalence similar to probabilistic bisimulation [41] is introduced.

Definition 6.4 (State bisimulation on LMP)**.** $R \subseteq S \times S$ is a *state bisimulation on LMP* $(S, \Sigma, \{\tau_a \mid a \in L\})$ if it is symmetric[3] and for all $s, t \in S$, $a \in L$, $s\ R\ t$ implies that $\tau_a(s)\ R\ \tau_a(t)$, i.e., for all $Q \in \Sigma(R)$, $\tau_a(s, Q) = \tau_a(t, Q)$.

This definition is point-wise and not event-wise as one should expect in a measure-theoretic setting; besides R has no restriction about measurability. The largest state bisimulation, called *state bisimilarity* and denoted $\sim_s$, is the union of all bisimulation relations:

$$\sim_s \doteq \bigcup\{R \mid R \text{ is a state bisimulation}\}$$

It is customary to prove that bisimilarity is an equivalence relation [45]. The relation $\sim_s$ is reflexive because the identity is a state bisimulation and symmetric since it is a restriction to properly define $\Sigma(R)$. The proof of transitivity provided by [20, 21] is painfully complicated, and it only applies to a restricted class of state spaces (Polish or analytic). With the tools we have nowadays, we provide in Section 6.2 a very simple and elegant proof showing that $\sim_s$ is an equivalence.

In [14] a measure-theory aware notion of behavioral equivalence is introduced.

Definition 6.5 (Event bisimulation on LMP)**.** An *event bisimulation on LMP* $(S, \Sigma, \{\tau_a \mid a \in L\})$ is a sub-σ-algebra Λ of Σ such that τ_a is Λ-measurable function for all $a \in L$.

With this notion the largest event bisimulation relation (induced by the operator $\mathcal{R}$), the *event bisimilarity* is defined by:

$$\sim_e \doteq \bigcup\{\mathcal{R}(\Lambda) \mid \Lambda \text{ is an event bisimulation}\}$$

Showing that $\sim_e$ is an equivalence relation is straightforward: we only have to prove that $\sim_e$ is transitive, since the union of equivalence relations is reflexive and symmetric, but not necessarily transitive. Notice that, if $s\ \mathcal{R}(\Lambda)\ t$

[2]Proposition 6.4(v) appears in [14] but with the unnecessary condition that R is a state bisimulation.

[3]The requirement of symmetry is needed otherwise $\Sigma(R)$ may not be a σ-algebra.

and $t\ \mathcal{R}(\Lambda')\ u$ then $s\ \mathcal{R}(\Lambda \cap \Lambda')\ u$. Moreover, if Λ and Λ' are event bisimulations so is $\Lambda \cap \Lambda'$. As a consequence $\sim_e$ is transitive. Using Proposition 6.3, the largest event bisimulation relation is defined by the smallest event bisimulation, that is $\sim_e = \mathcal{R}(\Lambda)$ such that any other event bisimulation Λ' contains Λ. We write $\sim_e = \mathcal{R}(\bigcap\{\Lambda \mid \Lambda \text{ is an event bisimulation}\})$.

In [14] it is shown that R is a state bisimulation iff $\Sigma(R)$ is an event bisimulation. This is an important result that leads to prove that the largest state bisimulation is also an event bisimulation.

Lemma 6.5. *Given an LMP $(S, \Sigma, \{\tau_a \mid a \in L\})$ and a symmetric relation $R \subseteq S \times S$, R is state bisimulation iff $\Sigma(R)$ is an event bisimulation.*

Proof. Let $q \in \mathbb{Q} \cap [0,1]$ and $Q \in \Sigma(R)$ Since $\tau^{-1}(\Delta^{>q}(Q))$ is in Σ by the definition of LMP, we only have to prove that R is a state bisimulation iff $\tau_a^{-1}(\Delta^{>q}(Q))$ is R-closed. For this we calculate:

R is a state bisimulation
iff (Def.)
$s\ R\ t \Rightarrow \forall Q \in \Sigma(R), \tau_a(s,Q) = \tau_a(t,Q)$
iff (Equiv. in $\mathbb{R}$)
$s\ R\ t \Rightarrow \forall Q \in \Sigma(R), \forall q \in \mathbb{Q} \cap [0,1], q < \tau_a(s,Q) \Leftrightarrow q < \tau_a(t,Q)$
iff (Def. of inv. and currification)
$s\ R\ t \Rightarrow \forall Q \in \Sigma(R), \forall q \in \mathbb{Q} \cap [0,1], s \in \tau_a^{-1}(\Delta^{>q}(Q)) \Leftrightarrow t \in \tau_a^{-1}(\Delta^{>q}(Q))$
iff (Def. of R-closed)
$\forall Q \in \Sigma(R), \forall q \in \mathbb{Q} \cap [0,1], R\text{-closed}(\tau_a^{-1}(\Delta^{>q}(Q)))$

□

Hence every state bisimulation is also an event bisimulation (via the operator $\mathcal{R}$), therefore we have the following result.

Corollary 6.6. $\sim_s \subseteq \sim_e$.

Logical Characterization. A Hennessy-Milner probabilistic logic for LMP $\mathcal{L}_s$ is based on a modal-probabilistic operator indexed by $q \in \mathbb{Q} \cap [0,1]$, the binary conjunction, and the top element:

$$\phi ::= \top \mid \phi \wedge \phi \mid \langle a \rangle_q \phi \tag{6.1}$$

Observe that the set of terms produced by this syntactic definition is countable if and only if the set of labels is countable. The default assumption is that L is an arbitrary set. The semantics of ϕ, denoted by $[\![\phi]\!]$, is given by

the set of states that make a formula ϕ valid. It is defined recursively as follows:

$$[\![\top]\!] = S, \qquad [\![\phi_1 \wedge \phi_2]\!] = [\![\phi_1]\!] \cap [\![\phi_2]\!], \qquad [\![\langle a \rangle_q \phi]\!] = \tau_a^{-1}(\Delta^{>q}([\![\phi]\!]))$$

Let $[\![\mathcal{L}_s]\!] = \{[\![\phi]\!] \mid \phi \in \mathcal{L}_s\}$. Note that $[\![\mathcal{L}_s]\!]$ is a π-system and that the semantics of every logical expression is a measurable set. The last claim is a consequence of τ_a being measurable.

Proposition 6.7. *For every $\phi \in \mathcal{L}_s$, we have $[\![\phi]\!] \in \Sigma$.*

Moreover $[\![\phi]\!]$ is closed for every state bisimulation relation.

Proposition 6.8. *Let R be a state bisimulation. Then, $\forall \phi \in \mathcal{L}_s$, R-closed$([\![\phi]\!])$.*

Proof. We proceed by induction on the structure of ϕ. For the base case notice that $[\![\top]\!] = S$ is trivially R-closed. For case $\phi_1 \wedge \phi_2$, by induction hypothesis $s \in [\![\phi_i]\!] \Leftrightarrow t \in [\![\phi_i]\!]$ with $i \in \{1, 2\}$; therefore $s \in [\![\phi_1]\!] \cap [\![\phi_2]\!] \Leftrightarrow t \in [\![\phi_1]\!] \cap [\![\phi_2]\!]$. For $\phi = \langle a \rangle_q \phi'$ suppose by induction hypothesis that R-closed$([\![\phi']\!])$. Take $s \; R \; t$. Since R is a state bisimulation, $\forall Q \in \Sigma(R), \tau_a(s, Q) = \tau_a(t, Q)$. Therefore $q < \tau_a(s, [\![\phi']\!]) \Leftrightarrow q < \tau_a(t, [\![\phi']\!])$ and hence $s \in [\![\phi]\!] \Leftrightarrow t \in [\![\phi]\!]$. So R-closed$([\![\phi]\!])$. □

By Proposition 6.7 and Proposition 6.8, we have that $[\![\mathcal{L}_s]\!] \subseteq \Sigma(R)$ for every state bisimulation R.

Even though this logic is very terse, it allows to capture the differences in small-step behavior for LMPs. Notice that the cardinality of the state space does not play any role in the complexity of the logic. What is mandatory for its distinguishing capabilities is that the modal operator captures a subset of the $\Delta(\Sigma)$ generators. Therefore, the modal operator has enough separation power to distinguish two different probability measures.

An important result relates the state bisimilarity and the semantics of the logic. We say that the logic $\mathcal{L}_s$ characterizes state bisimulation if

$$s \sim_s t \Leftrightarrow s \; \mathcal{R}(\mathcal{L}_s) \; t$$

where $s \; \mathcal{R}(\mathcal{L}_s) \; t$ is defined as $\forall \phi \in \mathcal{L}_s, s \in [\![\phi]\!] \Leftrightarrow t \in [\![\phi]\!]$. However this result is not valid in general. We need to restrict the state space and the cardinality of the label set.

Definition 6.6 (Polish and analytic spaces)**.** A topological space is *Polish* if it is separable (i.e. it contains a countable dense subset) and completely metrizable. A topological space is *analytic* if it is the continuous image of a Polish space. A measurable space is *analytic (standard) Borel* if it is isomorphic to $(X, \sigma(\mathcal{T}))$ where $\mathcal{T}$ is an analytic (Polish) topology on X.

Every standard Borel space is analytic, but the converse is false. The real line with the usual Borel σ-algebra, and more generally, $A^{\mathbb{N}}$ with A a countable discrete space, are standard Borel and therefore, analytic.

The next theorem from [23] essentially shows that in analytic Borel spaces, R-closed measurable sets are well behaved when the relation R is defined in terms of a sequence of measurable sets.

Theorem 6.9. *Let (S, Σ) be an analytic Borel space. Let $\mathcal{C} \subseteq \Sigma$ be countable and assume $S \in \mathcal{C}$. Then $\Sigma(\mathcal{R}(\mathcal{C})) = \sigma(\mathcal{C})$.*

The next lemma appears in [20,21]. We provide here an alternative proof.

Lemma 6.10. *$\mathcal{L}_s$ characterizes state bisimulation on LMPs with analytic state spaces and countable set of labels.*

Proof. The left to right implication follows immediately from Proposition 6.8.

For the right to left implication, we show that $\mathcal{R}(\mathcal{L}_s)$ is a state bisimulation, that is, we have to show that

$$s\ \mathcal{R}(\mathcal{L}_s)\ t \Rightarrow \forall Q \in \Sigma(\mathcal{R}(\mathcal{L}_s)), \tau_a(s, Q) = \tau_a(t, Q)$$

Since $\mathcal{L}_s$ is countable by the denumerability of L, by Theorem 6.9 we have that the above expression is equivalent to

$$s\ \mathcal{R}(\mathcal{L}_s)\ t \Rightarrow \forall Q \in \sigma([\![\mathcal{L}_s]\!]), \tau_a(s, Q) = \tau_a(t, Q)$$

However, given an arbitrary $\phi \in \mathcal{L}_s$, $s\ \mathcal{R}(\mathcal{L}_s)\ t$ implies that $\forall q \in \mathbb{Q} \cap [0,1], s \in [\![\langle a \rangle_q(\phi)]\!] \Leftrightarrow t \in [\![\langle a \rangle_q(\phi)]\!]$. This is equivalent to $\forall q \in \mathbb{Q} \cap [0,1], q < \tau_a(s, [\![\phi]\!]) \Leftrightarrow q < \tau_a(t, [\![\phi]\!])$, while in turn implies that $\tau_a(s, [\![\phi]\!]) = \tau_a(t, [\![\phi]\!])$. It remains to be shown that this equality extends to $\sigma([\![\mathcal{L}_s]\!])$, but this is valid by Theorem 3.20 since $[\![\mathcal{L}_s]\!]$ is a π-system. □

For event bisimulations there is no restriction on the state space, and this gives more evidence on the appropriateness of an event-wise behavioral equivalence. As previously stated, $\sim_e$ is given by $\mathcal{R}(\Lambda)$, where Λ is the smallest event bisimulation. We show that $\sigma([\![\mathcal{L}_s]\!])$ is such smallest σ-algebra. We remark that an event bisimulation is a σ-algebra that is *stable*, where stability is defined as follows.

Definition 6.7 (Stability)**.** The family $\mathcal{A} \subseteq \Sigma$ is *stable* for the LMP $(S, \Sigma, \{\tau_a \mid a \in A\})$ if for all $a \in L$, $q \in \mathbb{Q} \cap [0,1]$, $Q \in \mathcal{A}$, $\tau_a^{-1}(\Delta^{>q}(Q)) \in \mathcal{A}$.

We can alternatively express stability of the family $\mathcal{A}$ by $\{\tau_a^{-1}(\Delta^{>q}(Q)) \mid a \in L, q \in \mathbb{Q} \cap [0,1], Q \in \mathcal{A}\} \subseteq \mathcal{A}$.

Lemma 6.11. $[\![\mathcal{L}_s]\!]$ *is the smallest stable* π*-system containing* S *for* $(S, \Sigma, \{\tau_a \mid a \in L\})$.

Proof. We already know that $[\![\mathcal{L}_s]\!]$ is a π-system It is also stable since the syntax of the modal operator is interpreted as τ_a^{-1}. Finally suppose $\mathcal{P}$ is another stable π-system containing S. By induction on the formula structure, $[\![\top]\!] = S \in \mathcal{P}$, since $\mathcal{P}$ contains S. If $[\![\phi_1]\!], [\![\phi_2]\!] \in \mathcal{P}$, then, given that $\mathcal{P}$ is a π-system, $[\![\phi_1 \wedge \phi_2]\!] \in \mathcal{P}$. For $[\![\phi]\!] \in \mathcal{P}$, since $\mathcal{P}$ is stable, $\tau_a^{-1}(\Delta^{>q}([\![\phi]\!])) \in \mathcal{P}$, and hence $[\![\langle a \rangle_q \phi]\!] \in \mathcal{P}$. Therefore $[\![\mathcal{L}_s]\!] \subseteq \mathcal{P}$. □

Lemma 6.12. *If* $\mathcal{P}$ *is a stable* π*-system containing* S *for the LMP* $(S, \Sigma, \{\tau_a \mid a \in L\})$ *where* $\forall s \in S, a \in L, \tau_a(s) \in \Delta^{=1}(S)$*, then* $\sigma(\mathcal{P})$ *is also stable.*

Proof. We prove it using Dynkin's π-λ Lemma (Lemma 3.9). Let

$$\mathcal{G} = \{A \in \Sigma \mid \forall a \in L, \forall q \in \mathbb{Q} \cap [0,1], \tau_a^{-1}(\Delta^{>q}(A)) \in \sigma(\mathcal{P})\}$$

be the family of good sets and notice that $\mathcal{P} \subseteq \mathcal{G}$ because $\mathcal{P}$ is stable. If we show that $\mathcal{G}$ forms a λ-system, then $\sigma(\mathcal{P}) \subseteq \mathcal{G}$, concluding that $\sigma(\mathcal{P})$ is stable.

The family $\mathcal{G}$ is nonempty since $S \in \mathcal{P}$. For complement and disjoint union closure, we first show that the expressions $\Delta^{>q}(A^{\mathsf{c}})$ and $\Delta^{>q}(\biguplus_i A_i)$ can be conveniently rewritten in terms of denumerable set operations on $\Delta^{>q'}(A)$ and $\Delta^{>q'_i}(A_i)$ respectively:

$$\Delta^{>p}(A^{\mathsf{c}}) = \bigcap\nolimits_{n > \frac{1}{1-p}} (\Delta^{>1-p-\frac{1}{n}}(A))^{\mathsf{c}} \tag{6.2}$$

$$\Delta^{>p}(A_1 \uplus A_2) = \bigcap\nolimits_n \bigcup\nolimits_{I_n} (\Delta^{>\frac{l}{n}}(A_1) \cap \Delta^{>\frac{m}{n}}(A_2)) \tag{6.3}$$

$$\Delta^{>p}(\biguplus\nolimits_i A_i) = \bigcap\nolimits_{m \geq \frac{1}{p}} \bigcup\nolimits_n \Delta^{>p-\frac{1}{m}}(\biguplus\nolimits_{i=1}^{n} A_i) \tag{6.4}$$

where $I_n = \{(l,m) \mid l, m \in \mathbb{N} \ \wedge \ l, m \leq n \ \wedge \ l + m \geq \lfloor np \rfloor - 2\}$ for a given p. These expressions are a minor adaptation of similar expressions appearing in the proof of [65, Lemma 3.6][4].

Let $A \in \mathcal{G}$, then for arbitrary $q \in \mathbb{Q} \cap [0,1]$, $\tau_a^{-1}(\Delta^{>q}(A)) \in \sigma(\mathcal{P})$. Using expression (6.2) and observing that τ_a^{-1} commutes with set operations, we conclude that given an arbitrary $q \in \mathbb{Q} \cap [0,1]$, $\tau_a^{-1}(\Delta^{>q}(A^{\mathsf{c}})) \in \sigma(\mathcal{P})$, therefore $A^{\mathsf{c}} \in \mathcal{G}$. For disjoint union we proceed similarly. □

Observe that in the previous proof, the equation (6.2) uses the hypothesis $\mu(S) = 1$, therefore the logical characterization of event bisimulation is only valid for probability measures. It can be easily generalized to finite measures including subprobabilities.

[4] In [65], the notation $\beta^p(A)$ is the generator $\Delta^{\geq p}(A)$. For $\Delta^{>p}(A)$ expression (6.2) had to be changed, while the other two remained unchanged. (6.3) is valid since $p < 1$.

Proposition 6.13. *$\sigma([\![\mathcal{L}_s]\!])$ is the smallest stable σ-algebra included in Σ.*

Proof. $[\![\mathcal{L}_s]\!]$ is a stable π-system by Lemma 6.11, therefore by Lemma 6.12, $\sigma([\![\mathcal{L}_s]\!])$ is a stable σ-algebra included in Σ (Proposition 6.7). Let Σ_m be a stable σ-algebra included in Σ. Since $[\![\mathcal{L}_s]\!]$ is the smallest stable π-system, $[\![\mathcal{L}_s]\!] \subseteq \Sigma_m$, since Σ_m is in particular a π-system. Therefore $\sigma([\![\mathcal{L}_s]\!]) \subseteq \Sigma_m$, concluding the proof. □

Lemma 6.14. *$\mathcal{L}_s$ characterizes event bisimulations on LMPs.*

Proof. Since $\sigma([\![\mathcal{L}_s]\!])$ is stable, it is also an event bisimulation. Since it also is the smallest event bisimulation, it follows that $\mathcal{R}(\sigma([\![\mathcal{L}_s]\!])) = \sim_e$. □

Now, the obvious question is whether $\sim_s = \sim_e$ or there is an event bisimulation on an LMP that is not an state bisimulation. Recently in [56], it was shown that state and event bisimulation for LMPs differ outside analytic spaces. This raises the question of whether we should abandon state bisimulation in favor of event bisimulation. For NLMPs the situation will be different.

6.2 Bisimulations and Logics in NLMPs

Event bisimulation in NLMPs is defined exactly in the same way as for LMPs: an event bisimulation is a sub-σ-algebra for which the transition function is measurable.

Definition 6.8 (Event bisimulation on NLMP)**.** An *event bisimulation on an NLMP* $(S, \Sigma, \{T_a \mid a \in L\})$ is a sub-σ-algebra Λ of Σ such that $T_a : (S, \Lambda) \to (\Delta(\Sigma), H(\Delta(\Lambda)))$ is measurable for each $a \in L$.

Notice that the transition function T_a does not change its domain and codomain base sets. On the other hand, the σ-algebras attached do change. For Λ to be an event bisimulation, T_a should be measurable from Λ to $H(\Delta(\Lambda))$. Here, $H(\Delta(\Lambda))$ is the sub-σ-algebra of $H(\Delta(\Sigma))$ generated by $\{H_\xi \mid \xi \in \Delta(\Lambda)\}$. If we take $T_a^{-1}(H_\xi) = \{s \mid T_a(s) \cap \xi \neq \emptyset\}$ with $\xi \in \Delta(\Lambda)$, the inner intersection is between two slightly different type of sets. The elements of the left-hand side are in $\Sigma \to [0, 1]$, while the elements of the right-hand side are in $\Lambda \to [0, 1]$. The solution is restrict Σ to Λ, since a measure in $\Sigma \to [0, 1]$ is, by restriction, a measure in $\Lambda \to [0, 1]$. A similar digression is applicable to LMPs event bisimulations.

The definition of state bisimulation is less standard. Following the original definition of [44] (which was lifted to discrete probabilistic models by [41]), a

traditional definition of bisimulation (see Definition 6.10) verifies that, whenever $s\ R\ t$, every measure on $T_a(s)$ has a corresponding one (modulo R) in $T_a(t)$. Rather than looking point-wise at probability measures, our definition follows the idea of Definition 4.4 and verifies that both $T_a(s)$ and $T_a(t)$ *hit* the same measurable sets of measures.

Definition 6.9 (State bisimulation on an NLMP)**.** A relation $R \subseteq S \times S$ is a *state bisimulation on an NLMP* $(S, \Sigma, \{T_a \mid a \in L\})$ if it is symmetric and for all $a \in L$, $s\ R\ t$ implies $\forall \xi \in \Delta(\Sigma(R)), T_a(s) \cap \xi \neq \emptyset \Leftrightarrow T_a(t) \cap \xi \neq \emptyset$.

The following property, which also holds in LMPs, states the fundamental relation between state bisimulation and event bisimulation.

Lemma 6.15. *Given an NLMP $(S, \Sigma, \{T_a \mid a \in L\})$ and a symmetric $R \subseteq S \times S$, R is state bisimulation iff $\Sigma(R)$ is an event bisimulation.*

Proof. By Definition 6.8, $\Sigma(R)$ is an event bisimulation iff T_a is $\Sigma(R)$-measurable. Since T_a is Σ-measurable, it suffices to prove that $T_a^{-1}(H_\xi)$ is R-closed for all labels $a \in L$ and generators H_ξ, $\xi \in \Delta(\Sigma(R))$. We calculate:

$$
\begin{aligned}
&R\text{-closed}(T_a^{-1}(H_\xi)) \\
&\text{iff} && (R \text{ is symmetric}) \\
&s\ R\ t \Rightarrow \left(s \in T_a^{-1}(H_\xi) \Leftrightarrow t \in T_a^{-1}(H_\xi)\right) \\
&\text{iff} && (\text{Def. inverse function}) \\
&s\ R\ t \Rightarrow (T_a(s) \in H_\xi \Leftrightarrow T_a(t) \in H_\xi) \\
&\text{iff} && (\text{Def. of } H_\xi) \\
&s\ R\ t \Rightarrow (T_a(s) \cap \xi \neq \emptyset \Leftrightarrow T_a(t) \cap \xi \neq \emptyset)\,.
\end{aligned}
$$

This completes the proof as the last statement is the definition of state bisimulation. □

The following results are consequences of Proposition 6.4 and, for the case of Lemma 6.16 (iii), Lemma 6.15 and the fact that $\mathcal{R}(\Lambda)$ is an equivalence relation. The proofs are the same as the proofs of similar results for LMPs in [14].

Lemma 6.16. *Let R be a state bisimulation. Then:*

i. R is an event bisimulation iff $R = \mathcal{R}(\Sigma(R))$.

ii. If the equivalence classes of R are in Σ, R is an event bisimulation.

iii. $\mathcal{R}(\Sigma(R))$ is both a state bisimulation and an event bisimulation.

Let $\sim_s = \bigcup\{R \mid R \text{ is a state bisimulation}\}$. In the following we show that $\sim_s$ is also a state bisimulation and hence the largest one. Moreover, we show that $\sim_s$ is also an event bisimulation and, as a consequence, an equivalence relation.

Theorem 6.17. *$\sim_s$ is:*

i. the largest state bisimulation;

ii. an event bisimulation (and hence $\sim_s \subseteq \sim_e$); and

iii. an equivalence relation.

Proof. (i) Take $s, t \in S$ such that $s \sim_s t$. Then there is a state bisimulation R with $s\ R\ t$. Take a measurable set $\xi \in \Delta(\Sigma(\sim_s))$. Since $R \subseteq \sim_s$, by Proposition 6.1, $\Delta(\Sigma(R)) \supseteq \Delta(\Sigma(\sim_s))$. Hence $\xi \in \Delta(\Sigma(R))$ and by Definition 6.9, $T_a(s) \cap \xi \neq \emptyset \Leftrightarrow T_a(t) \cap \xi \neq \emptyset$ which proves that $\sim_s$ is a state bisimulation. By definition, it is the largest one. (ii) Since $\sim_s$ is a state bisimulation, $\mathcal{R}(\Sigma(\sim_s))$ is a state bisimulation and an event bisimulation (Lemma 6.16 (iii)). Since $\sim_s$ is the largest bisimulation then $\sim_s = \mathcal{R}(\Sigma(\sim_s))$ and hence it is an event bisimulation. (iii) By definition, every event bisimulation is an equivalence relation. □

We have already stated that our definition of state bisimulation differs from a more traditional view such as those in [7,8,15,16,62]. These definitions closely resemble the definition given in [41]. The only difference is that two measures are considered equivalent if they agree in every *measurable union of* equivalence classes induced by the relation. We will now give a variant using the twice lifted relation R as in the PA bisimulation (see Definition 2.9).

Definition 6.10 (Traditional bisimulation on an NLMP)**.** A relation R is a *traditional bisimulation on an NLMP* $(S, \Sigma, \{T_a \mid a \in L\})$ if it is symmetric and for all $a \in L$, $s\ R\ t$ implies $T_a(s)\ R\ T_a(t)$. We say that $s, t \in S$ are *traditionally bisimilar*, denoted by $s \sim_t t$, if there is a traditional bisimulation R such that $s\ R\ t$.

The proof of the next proposition follows the standard strategy of the classic bisimulation as in [44]. Other than in the probabilistic treatment, it only differs in that the composition $R \circ R'$ is granted to be a traditional bisimulation if R and R' are *reflexive* traditional bisimulations. (If one of R or R' is not reflexive, $R \circ R'$ may not be a traditional bisimulation.)

Proposition 6.18. *$\sim_t$ is a traditional bisimulation and an equivalence relation.*

In the following we discuss the relation between state bisimulations and traditional bisimulations. Lemma 6.19 states that every traditional bisimulation is a state bisimulation. Theorems 6.20 and 6.21 give sufficient conditions to strengthen Lemma 6.19 so that the converse also holds.

Lemma 6.19. *If R is a traditional bisimulation, then R is a state bisimulation.*

Proof. Let $s\ R\ t$ and $\xi \in \Delta(\Sigma(R))$. If $T_a(s) \cap \xi \neq \emptyset$, then there is $\mu \in T_a(s)$ such that $\mu \in \xi$. Since R is a traditional bisimulation, $T_a(s)\ R\ T_a(t)$, i.e., there is $\mu' \in T_a(t)$ such that $\mu\ R\ \mu'$. By Proposition 6.2 R-closed(ξ), so $\mu' \in \xi$, and hence $T_a(t) \cap \xi \neq \emptyset$ as required. The other implication follows by symmetry. □

In the following we give two sufficient conditions that ensure that a state bisimulation is also a traditional bisimulation. The first condition focuses on the NLMP, it requires the NLMP to be *image denumerable*. (By a minor extension of Proposition 5.5, any countable set of LMPs form a denumerably branching NLMP.)

Definition 6.11. An NLMP $(S, \Sigma, \{T_a \mid a \in L\})$ is *image denumerable* iff for all $a \in L, s \in S$, the image of the transition $T_a(s)$ is denumerable.

Theorem 6.20. *Let $(S, \Sigma, \{T_a \mid a \in L\})$ be an image denumerable NLMP. Then R is a traditional bisimulation iff it is a state bisimulation.*

Proof. The left to right implication is Lemma 6.19. For the other implication we proceed as follows.

Let $s\ R\ t$ and for all $\xi \in \Delta(\Sigma(R))$, $T_a(s) \cap \xi \neq \emptyset \Leftrightarrow T_a(t) \cap \xi \neq \emptyset$. Suppose towards a contradiction that $T_a(s)\ \not R\ T_a(t)$, i.e. $\exists \mu \in T_a(s), \forall \mu'_i \in T_a(t) : \exists Q_i \in \Sigma(R) : \mu(Q_i) \bowtie_i \mu'_i(Q_i)$, where $\{\bowtie_i\}_i \subseteq \{>, <\}$ (the NLMP is image denumerable). By density of the rationals, there are $\{q_i\}_i \subseteq \mathbb{Q}^+$ such that $\mu(Q_i) \bowtie_i q_i \bowtie_i \mu'_i(Q_i)$. Then $\mu \in \Delta^{\bowtie_i q_i}(Q_i) \not\ni \mu'_i$. Let $\xi \doteq \bigcap_i \Delta^{\bowtie_i q_i}(Q_i)$. This set is measurable, moreover, since every $Q_i \in \Sigma(R)$, so $\xi \in \Delta(\Sigma(R))$. Then $\mu \in T_a(s) \cap \xi$, but $T_a(t) \cap \xi = \emptyset$, hence contradicting the assumption. □

After reading the proof, it should be clear that we can relax the sufficient condition so that we only require that the partition $T_a(s)/R$ is denumerable for each state s and label a instead of requiring image denumerability.

Observe that a state bisimulation on an LMP is a traditional bisimulation on the encoding NLMP and vice versa since $\{\tau_a(s)\} = T_a(s)\ R\ T_a(t) = \{\tau_a(t)\}$ iff $\tau_a(s)\ R\ \tau_a(t)$. As a consequence of Lemma 6.19 and Theorem 6.20 (a deterministic NLMP is image denumerable), we conclude that a state

bisimulation on an LMP is a state bisimulation on the encoding NLMP and vice versa.

The second sufficient condition looks at the σ-algebra $\Sigma(R)$ induced by the state bisimulation R. It turns out that if $\Sigma(R)$ is countably generated, then R is also a traditional bisimulation.

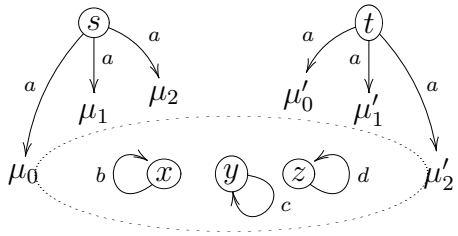
Theorem 6.21. *Let $(S, \Sigma, \{T_a \mid a \in L\})$ be an NLMP and R a symmetric relation such that $\Sigma(R)$ is generated by a denumerable family $\mathcal{C}$. Then R is a traditional bisimulation iff it is a state bisimulation.*

Proof. As before, the left to right implication is Lemma 6.19. For the other implication we proceed as follows. Suppose towards a contradiction that $s \ R \ t$ and $T_a(s) \not R \ T_a(t)$, that is $\exists \mu \in T_a(s), \forall \mu' \in T_a(t) : \mu \not R \ \mu'$. First we generate the denumerable algebra $\mathcal{F}$ out of $\mathcal{C}$ (see below Proposition 3.3). By Theorem 3.21, this implies that there exists $Q_i \in \mathcal{F}$ such that $\mu(Q_i) \neq \mu'(Q_i)$. The rest of the proof is as in Theorem 6.20. □

Logical Characterization. For systems (i.e. LMPs) with no internal nondeterminism, the logic that characterizes bisimulation is very simple: it suffices with a binary conjunction and a modal operator. This logic could have been equivalently written in a two-level syntax, a first level of formulas valid on states: $\phi ::= \top \mid \phi_1 \wedge \phi_2 \mid \langle a \rangle \psi$, and a second level describing sets of probabilities over states $\psi ::= [\phi]_q$. The semantics of the new operators are given by $[\![\langle a \rangle \psi]\!] = \tau_a^{-1}([\![\psi]\!])$, and $[\![[\phi]_q]\!] = \Delta^{>q}([\![\phi]\!])$. We decoupled the modal operator of (6.1) in order to stress how the logical characterization works. The first level of the logic forms a π-system on states, and the second level are the generators on $\Delta(\Sigma)$. The following example due to [13], shows that for a finitely branching NLMP, the logic that only captures generators in the second level of the logic is not sufficient. The semantics of this two-level logic is interpreted over NLMPs as follows

$$[\![\langle a \rangle \psi]\!] = T_a^{-1}(H_{[\![\psi]\!]}) \qquad [\![[\phi]_q]\!] = \Delta^{>q}([\![\phi]\!])$$

Example 6.12. Take the discrete NLMP depicted in Figure 6.1. States s and t are not bisimilar since given a $\mu \in T_a(s)$, there is no $\mu' \in T_a(t)$ such that $\mu(Q) = \mu'(Q)$ for all $Q \in \{\{x\}, \{y\}, \{z\}\}$ (which are the only relevant possible R-closed sets). A logic having modalities that can only describe one behavior after a label will not be able to distinguish between s and t. For example, $[\![\langle a \rangle [\phi]_q]\!] = \{s \mid T_a(s) \cap \Delta^{>q}([\![\phi]\!]) \neq \emptyset\}$ will always have s and t together. Observe that negation, denumerable conjunction or disjunction, do not add any distinguishing power (on an image finite setting).



	μ_0	μ_1	μ_2	μ'_0	μ'_1	μ'_2
$\{x\}$	$\frac{1}{3}$	$\frac{2}{3}$	0	$\frac{2}{3}$	$\frac{1}{3}$	0
$\{y\}$	0	$\frac{1}{3}$	$\frac{2}{3}$	0	$\frac{2}{3}$	$\frac{1}{3}$
$\{z\}$	$\frac{2}{3}$	0	$\frac{1}{3}$	$\frac{1}{3}$	0	$\frac{2}{3}$

Figure 6.1: s and t are not bisimilar

The essential need for a more expressive second level of the logic also shows that our σ-algebra $H(\Delta(\Sigma))$ in Definition 4.4 cannot be simplified to $\sigma(\{H_{\Delta^{>q}(Q)} : q \in \mathbb{Q} \cap [0,1], Q \in \Sigma\})$. States s and t in the example above should be observationally distinguished from each other. Formally, this amounts to saying that there must be some label a and some measurable $\Theta \in H(\Delta(\Sigma))$ such that $T_a^{-1}(\Theta)$ separates $\{s\}$ from $\{t\}$. Therefore, the same must be true for some generator Θ, but this does not hold for the family $\{H_{\Delta^{>q}(Q)} : q \in \mathbb{Q} \cap [0,1], Q \in \Sigma\}$. The problem is that hit sets do not transfer intersections nor complements.

The hypothesis is that the logic should accompany the branching of the nondeterminism. Following that idea, [13, 18] define a two-level logic $\mathcal{L}_f$ where the probabilistic level is able to distinguish between two *finite* sets of measures. We will quickly review it and present with more detail the logic that captures unbounded nondeterminism. The syntax is as follows:

$$\begin{array}{lcl} \phi & ::= & \top \mid \phi_1 \wedge \phi_2 \mid \langle a \rangle \psi \\ \psi & ::= & [_{\bowtie_i q_i} \phi_i]_{i=1}^n \end{array}$$

where $a \in L$, $\bowtie_i \in \{>, <\}$, $q \in \mathbb{Q} \cap [0,1]$. The semantics of the modal and probabilistic level is:

$$\begin{array}{c} [\![\langle a \rangle \psi]\!] = T_a^{-1}(H_{[\![\psi]\!]}) \\ [\![[_{\bowtie_i q_i} \phi_i]_{i=1}^n]\!] = \bigcap_{i=1}^n \Delta^{\bowtie_i q_i}([\![\phi_i]\!]) \end{array}$$

Notice that $\mathcal{L}_s \subseteq \mathcal{L}_f$, an that $\mathcal{L}_f$ is countable if and only if L is countable. The fact that $\mathcal{L}_f$ is countable is important for the proof of logical characterization of traditional bisimulation on analytic spaces.

We need the possibility to quantify with both inequalities, $>$ and $<$. Without them, the LTSs shown in Figure 6.2 cannot be distinguished, since bisimulation on nondeterministic models requires some form of negation. The only formula distinguishing them in Hennessy-Milner logic is $\langle a \rangle \neg \langle b \rangle \top$, which is translated to $\langle a \rangle [_{>0} \langle b \rangle [_{<0} \top]]$ in our logic.

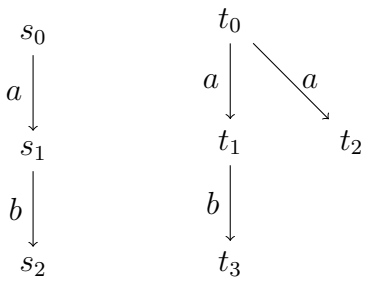


Figure 6.2: Two discrete space LTSs showing that $[\,_{>q_i}\phi_i]_{i=1}^n$ is not sufficient in presence of finitary nondeterminism.

In the following, we show that for image finite NLMPs, the logic $\mathcal{L}_f$ is enough to single out non-bisimilar states, that is, the so called *transfer property* can be encoded using the modality.

Lemma 6.22. *Let $(S, \Sigma, \{T_a \mid a \in L\})$ be an image finite NLMP. Then for every pair of states such that $s\ \mathcal{R}(\mathcal{L}_f)\ t$ and $\mu \in T_a(s)$, there is a $\mu' \in T_a(t)$ such that $\forall \phi \in \mathcal{L}_f, \mu([\![\phi]\!]) = \mu'([\![\phi]\!])$.*

Proof. Suppose towards a contradiction that there are s, t with $s\ \mathcal{R}(\mathcal{L}_f)\ t$ and there is a $\mu \in T_a(s)$, such that for all $\mu'_i \in T_a(t)$ there is a formula $\phi_i \in \mathcal{L}_f$ with $\mu([\![\phi_i]\!]) \neq \mu'_i([\![\phi_i]\!])$. Since $T_a(t)$ is finite, there are at most n different μ'_i. We can choose $\bowtie_i \in \{>, <\}, q_i \in \mathbb{Q} \cap [0,1]$ accordingly to make $\mu([\![\phi_i]\!]) \bowtie_i q_i \bowtie_i \mu'_i([\![\phi_i]\!])$. Take $\phi = \langle a \rangle [\,_{\bowtie_i q_i}\phi_i]_{i=1}^n$. Then $s \in [\![\phi]\!]$ but $t \notin [\![\phi]\!]$ contradicting $s\ \mathcal{R}(\mathcal{L}_f)\ t$. □

The main result from [13] (presented also in [18]) says that traditional bisimilarity is characterized by $\mathcal{L}_f$ for finitely branching NLMPs on analytic state spaces. Its proof is similar to its LMP counterpart (Lemma 6.10).

Lemma 6.23. *Let $(S, \Sigma, \{T_a \mid a \in L\})$ be an image finite NLMP, with (S, Σ) analytic and denumerable L, then $\sim_t = \mathcal{R}(\mathcal{L}_f)$.*

Proof. Left to right implication follows from a minor extension of Proposition 6.8 to include the modal operator $\langle a \rangle [\,_{\bowtie_i q_i}\phi_i]_{i=1}^n$.

For the right to left implication we show that $\mathcal{R}(\mathcal{L}_f)$ is a traditional bisimulation, that is,

$$s\ \mathcal{R}(\mathcal{L}_f)\ t \Rightarrow \big(\forall \mu \in T_a(s), \exists \mu' \in T_a(t), \forall Q \in \Sigma(\mathcal{R}(\mathcal{L}_f)), \mu(Q) = \mu'(Q)\big)$$

Since $\mathcal{L}_f$ is countable, by Theorem 6.9 we have that the previous expression is equivalent to

$$s\ \mathcal{R}(\mathcal{L}_f)\ t \Rightarrow \big(\forall \mu \in T_a(s), \exists \mu' \in T_a(t), \forall Q \in \sigma([\![\mathcal{L}_f]\!]), \mu(Q) = \mu'(Q)\big)$$

The above expression is basically Lemma 6.22, where we extend the equality from values $Q \in [\![\mathcal{L}_f]\!]$ to values $Q \in \sigma([\![\mathcal{L}_f]\!])$ using Theorem 3.20 since $[\![\mathcal{L}_f]\!]$ is a π-system. □

We present in the following a more general logic that completely characterizes event bisimulation on arbitrary NLMPs.

$$\begin{array}{rcl} \phi & ::= & \top \mid \phi_1 \wedge \phi_2 \mid \langle a \rangle \psi \\ \psi & ::= & \bigvee_i \psi_i \mid \neg\psi \mid [\phi]_q \end{array}$$

where $a \in L$, the disjunction is denumerable, and $q \in \mathbb{Q} \cap [0,1]$. We denote by $\mathcal{L}$ the set of all formulas generated by the first production and by $\mathcal{L}_\Delta$ the set of all formulas generated by the second production. Notice that the countability of the logic is lost, even if we restrict L to be countable.

The semantics is defined with respect to an NLMP $(S, \Sigma, \{T_a \mid a \in L\})$. Formulas in $\mathcal{L}$ are interpreted as sets of states in which they are true, and formulas in $\mathcal{L}_\Delta$ are interpreted as sets of measures on the state space as follows:

$$\begin{array}{ll} [\![\top]\!] = S & [\![\bigvee_i \psi_i]\!] = \bigcup_i [\![\psi_i]\!] \\ [\![\phi_1 \wedge \phi_2]\!] = [\![\phi_1]\!] \cap [\![\phi_2]\!] & [\![\neg\psi]\!] = [\![\psi]\!]^{\mathsf{c}} \\ [\![\langle a \rangle \psi]\!] = T_a^{-1}(H_{[\![\psi]\!]}) & [\![[\phi]_q]\!] = \Delta^{>q}([\![\phi]\!]) \end{array}$$

In particular, notice that $\langle a \rangle \psi$ is valid in a state s whenever there is some measure $\mu \in T_a(s)$ that makes ψ valid (existential quantification over nondeterminism), and that $[\phi]_q$ is valid in a measure μ whenever $q < \mu([\![\phi]\!])$. As a consequence, we need the sets $[\![\phi]\!]$ and $[\![\psi]\!]$ to be measurable in Σ and $\Delta(\Sigma)$, respectively. Indeed, this follows in the same way we proved it for LMPs, by induction on the construction of the formula, after observing that all operations involved in the definition of the semantics preserve measurability (in particular T_a is a measurable function). For the rest of this section, we fix $[\![\mathcal{L}]\!] = \{[\![\phi]\!] \mid \phi \in \mathcal{L}\}$ and $[\![\mathcal{L}_\Delta]\!] = \{[\![\psi]\!] \mid \psi \in \mathcal{L}_\Delta\}$.

We show that $\mathcal{L}$ characterizes event bisimulation. This is again an immediate consequence of the fact that $\sigma([\![\mathcal{L}]\!])$, the σ-algebra generated by the logic $\mathcal{L}$, is the smallest event bisimulation, which is what we aim to prove in this part of the section. The proof strategy resembles that of Section 6.1 but it is properly tailored to our two level logic. Moreover, such a separation and the definition of the hit σ-algebra allowed us to find an alternative to Dynkin's Lemma (used in Section 6.1 and in the original proof [14]).

The concept of *stable* family of measurable sets is crucial to the proof of Theorem 6.28.

Definition 6.13 (NLMP stable)**.** The family $\mathcal{A} \subseteq \Sigma$ is *stable* for NLMP $(S, \Sigma, \{T_a \mid a \in A\})$ if for all $a \in L$, $\xi \in \Delta(\mathcal{A})$, we have $T_a^{-1}(H_\xi) \in \mathcal{A}$.

Notice that $\Delta(\mathcal{A})$ is the σ-algebra generated by $\Delta^{>q}(Q)$ where $Q \in \mathcal{A}$, therefore $\mathcal{A}$ is an event bisimulation iff it is a stable σ-algebra.

The key point of the proof is to show that $[\![\mathcal{L}]\!]$ is the smallest stable π-system, which is stated in Lemma 6.25. The next lemma is auxiliary to Lemma 6.25.

Lemma 6.24. $[\![\mathcal{L}_\Delta]\!] = \Delta([\![\mathcal{L}]\!])$.

Proof. $[\![\mathcal{L}_\Delta]\!]$ is a σ-algebra since:

i. $\emptyset = [\![[\top]_1]\!] \in [\![\mathcal{L}_\Delta]\!]$, therefore it is nonempty;

ii. for $\xi_i \in [\![\mathcal{L}_\Delta]\!]$ there are $\psi_i \in \mathcal{L}_\Delta$ such that $\xi_i = [\![\psi_i]\!]$, and hence $\bigcup_i \xi_i = \bigcup_i [\![\psi_i]\!] = [\![\bigvee_i \psi_i]\!] \in [\![\mathcal{L}_\Delta]\!]$; and

iii. for $\xi \in [\![\mathcal{L}_\Delta]\!]$ there is $\psi \in \mathcal{L}_\Delta$ such that $\xi = [\![\psi]\!]$, and hence $\xi^{\mathsf{c}} = [\![\psi]\!]^{\mathsf{c}} = [\![\neg\psi]\!] \in [\![\mathcal{L}_\Delta]\!]$.

Moreover, since $[\![[\phi]_q]\!] = \Delta^{>q}([\![\phi]\!])$, every generator set of $\Delta([\![\mathcal{L}]\!])$ is in $[\![\mathcal{L}_\Delta]\!]$ and hence $\Delta([\![\mathcal{L}]\!]) \subseteq [\![\mathcal{L}_\Delta]\!]$.

Finally, it can be proven by induction on the depth of the formula that $[\![\mathcal{L}_\Delta]\!] \subseteq \mathcal{C}$ for any σ-algebra $\mathcal{C}$ containing all the sets $[\![[\phi]_q]\!] = \Delta^{>q}([\![\phi]\!])$ for $q \in \mathbb{Q} \cap [0,1]$ and $\phi \in \mathcal{L}$. Then $[\![\mathcal{L}_\Delta]\!]$ is the smallest σ-algebra containing all generator sets of $\Delta([\![\mathcal{L}]\!])$. Therefore $[\![\mathcal{L}_\Delta]\!] = \Delta([\![\mathcal{L}]\!])$. □

Lemma 6.25. $[\![\mathcal{L}]\!]$ *is the smallest stable* π*-system containing* S *for a given NLMP* $(S, \Sigma, \{T_a \mid a \in L\})$.

Proof. $[\![\mathcal{L}]\!]$ contains S since $[\![\top]\!] = S$ and it is a π-system since for $Q_1, Q_2 \in [\![\mathcal{L}]\!]$ there are $\phi_1, \phi_2 \in \mathcal{L}$ such that $Q_1 = [\![\phi_1]\!]$ and $Q_2 = [\![\phi_2]\!]$, and hence $Q_1 \cap Q_2 = [\![\phi_1]\!] \cap [\![\phi_2]\!] = [\![\phi_1 \wedge \phi_2]\!] \in [\![\mathcal{L}]\!]$.

For stability, let $\xi \in \Delta([\![\mathcal{L}]\!])$. By Lemma 6.24, there is $\psi \in \mathcal{L}_\Delta$ such that $[\![\psi]\!] = \xi$. Then $T_a^{-1}(H_\xi) = T_a^{-1}(H_{[\![\psi]\!]}) = [\![\langle a \rangle \psi]\!] \in [\![\mathcal{L}]\!]$.

Let $\mathcal{P}$ be another stable π-system for the NLMP $(S, \Sigma, \{T_a \mid a \in L\})$ containing S. By induction on the depth of the formula we show simultaneously that $[\![\mathcal{L}]\!] \subseteq \mathcal{P}$ and $\Delta([\![\mathcal{L}]\!]) \subseteq \Delta(\mathcal{P})$. First note that $[\![\top]\!] = S \in \mathcal{P}$ since $\mathcal{P}$ contains S. Now suppose as induction hypothesis that $[\![\phi]\!], [\![\phi_1]\!], [\![\phi_2]\!] \in \mathcal{P}$ and $[\![\psi]\!], [\![\psi_i]\!] \in \Delta(\mathcal{P})$. Then:

i. $[\![\phi_1 \wedge \phi_2]\!] = [\![\phi_1]\!] \cap [\![\phi_2]\!] \in \mathcal{P}$, because $\mathcal{P}$ is a π-system;

ii. $[\![\langle a \rangle \psi]\!] = T_a^{-1}(H_{[\![\psi]\!]}) \in \mathcal{P}$, because $\mathcal{P}$ is stable;

iii. $[\![\bigvee_i \psi_i]\!] = \bigcup_i [\![\psi_i]\!] \in \Delta(\mathcal{P})$ and

iv. $[\![\neg\psi]\!] = [\![\psi]\!]^{\mathsf{c}} \in \Delta(\mathcal{P})$ because $\Delta(\mathcal{P})$ is a σ-algebra; and finally,

v. $[\![[\phi]_p]\!] = \Delta^{>p}([\![\phi]\!]) \in \Delta(\mathcal{P})$ by definition of generator set of $\Delta(\mathcal{P})$.

□

Lemma 6.26 is auxiliary to Lemma 6.27. It is also significantly simpler than its related lemma in [14, Lemma 5.4] and its alternative version given in Lemma 6.12. This is due to our definition of stability and the powerful technical result from [65].

Lemma 6.26. *If $\mathcal{P}$ is a stable π-system for the NLMP $(S, \Sigma, \{T_a \mid a \in L\})$, then $\sigma(\mathcal{P})$ is also stable.*

Proof. First notice that $\mathcal{P}$ is stable iff $\{T_a^{-1}(H_\xi) \mid a \in L, \xi \in \Delta(\mathcal{P})\} \subseteq \mathcal{P}$. By Lemma 3.39, $\Delta(\mathcal{P}) = \Delta(\sigma(\mathcal{P}))$. Then $\{T_a^{-1}(H_\xi) \mid a \in L, \xi \in \Delta(\sigma(\mathcal{P}))\} \subseteq \mathcal{P} \subseteq \sigma(\mathcal{P})$, which proves that $\sigma(\mathcal{P})$ is stable. □

The next lemma is central to the proof that $\mathcal{L}$ characterizes event bisimulation, which is then presented in Theorem 6.28.

Lemma 6.27. *$\sigma(\mathcal{L})$ is the smallest stable σ-algebra included in Σ.*

Proof. Let $\mathcal{A}$ be the smallest stable σ-algebra included in Σ. By Lemma 6.25, $[\![\mathcal{L}]\!] \subseteq \mathcal{A}$, since $\mathcal{A}$ is a stable π-system. Therefore $\sigma(\mathcal{L}) \subseteq \mathcal{A}$ since $\mathcal{A}$ is also a σ-algebra. For the other inclusion, we observe that $[\![\mathcal{L}]\!]$ is a stable π-system because of Lemma 6.25. Then by Lemma 6.26, $\sigma(\mathcal{L})$ is stable, and thus it contains $\mathcal{A}$. □

Theorem 6.28. *The logic $\mathcal{L}$ completely characterizes event bisimulations for NLMPs. In other words, $\mathcal{R}(\mathcal{L}) = \sim_e$*

Proof. Lemma 6.27 establishes that $\sigma(\mathcal{L})$ is stable, that is, it is an event bisimulation. Being the smallest, it implies that any other event bisimulation preserves $\mathcal{L}$ formulas. □

A consequence of this theorem together with Theorem 6.17 and Lemma 6.19 is that both traditional and state bisimulation are sound for $\mathcal{L}$, that is, they preserve the validity of formulas.

Theorem 6.29. $\sim_t \subseteq \sim_s \subseteq \sim_e = \mathcal{R}(\mathcal{L})$.

An immediate corollary of this result, together with Lemma 6.23, is that, for image finite NLMP on analytic spaces, the four notions of behavioral equivalence coincide.

Corollary 6.30. *For a finitely branching NLMP $(S, \Sigma, \{T_a \mid a \in L\})$ over an analytic space S, it holds $\sim_t = \sim_s = \sim_e = \mathcal{R}(\mathcal{L}_f)$.*

Proof. Since $[\![\mathcal{L}_f]\!] \subseteq [\![\mathcal{L}]\!]$, by Proposition 6.3 we have $\mathcal{R}(\mathcal{L}) \subseteq \mathcal{R}(\mathcal{L}_f)$. By Lemma 6.23 $\sim_t = \mathcal{R}(\mathcal{L}_f)$, transforming the two inclusions of Theorem 6.29 into equalities. □

Examples separating NLMP bisimulations. In the following, we provide examples showing that $\sim_t \subsetneq \sim_s \subsetneq \sim_e$, therefore strengthening Theorem 6.29. We construct the examples using non-probabilistic NLMPs (Definition 4.12) with transitions of the form $\tilde{T}_a : S \to \Sigma$.

These LTSs with a σ-algebra attached have the following definitions for event, state and traditional bisimulation. Using the results of Section 4.4, they could be shown to be equivalent to the already introduced notions of bisimulation for general NLMPs.

Definition 6.14. Given a non-probabilistic NLMP $(S, \Sigma, \{\tilde{T}_a \mid a \in L\})$, where $\tilde{T}_a : (S, \Sigma) \to (\Sigma, H(\Sigma))$ we have

i. A sub-σ-algebra Λ of Σ is an *event bisimulation* if $\forall Q \in \Lambda, \tilde{T}_a^{-1}(H_Q) \in \Lambda$.

ii. A symmetric $R \subseteq S \times S$ is a *state bisimulation* if $s \ R \ t \Rightarrow \forall Q \in \Sigma(R), \big(s \in \tilde{T}_a^{-1}(H_Q) \Leftrightarrow t \in \tilde{T}_a^{-1}(H_Q)\big)$.

iii. A symmetric $R \subseteq S \times S$ is a *traditional bisimulation* if $s \ R \ t \Rightarrow \tilde{T}_a(s) \ \mathcal{R}(\Sigma(R)) \ \tilde{T}_a(t)$.

Notice that the measurable notion of traditional bisimulation is weaker than the (traditional) LTS bisimulation from [44] expressed in Equation (2.2). The result given in Lemma 6.15 is (obviously) valid for non-probabilistic NLMP:

Lemma 6.31. *Given a non-probabilistic NLMP $(S, \Sigma, \{\tilde{T}_a \mid a \in L\})$ and a symmetric relation $R \subseteq S \times S$, R is a state bisimulation iff $\Sigma(R)$ is an event bisimulation.*

Next, we present the example that separates traditional bisimulation from the other ones. The proof is reported in Theorem 6.33.

Example 6.15. Consider the standard Borel space $(S_1, \Sigma_1) = ([0,1] \uplus [2,3] \uplus \{s,t,x\}, \mathcal{B}([0,1] \uplus [2,3] \uplus \{s,t,x\}))$, where $\{s,t,x\} \subseteq \mathbb{R} \setminus [0,3]$. Let V be a non-Borel subset of $[2.5, 3]$. Take the injection $[0,1] \to [2,3] \setminus V$, that contracts linearly $[0,1]$ to $[2, 2.5]$, and the injection $[2,3] \setminus V \to [0,1]$ that moves the

points to the left by 2. By the Cantor-Bernstein-Schröeder Theorem [1], there is a bijection f between $[0,1]$ and $[2,3] \setminus V$ (they are equinumerous). The continuous label set is $L_1 = \{a\} \uplus [0,1]$, and the transition function $\tilde{T}_x$ is defined as follows:

$$\begin{aligned}
&\tilde{T}_a(s) = [0,1] \\
&\tilde{T}_a(t) = [2,3] \\
&\tilde{T}_r(r) = \tilde{T}_r(f(r)) = \{x\} && \text{if } r \in [0,1] \\
&\tilde{T}_c(y) = \emptyset && \text{otherwise.}
\end{aligned}$$

It can be seen that $\mathbf{S_1} = (S_1, \Sigma_1, \{\tilde{T}_x \mid x \in L_1\})$ is a non-probabilistic NLMP conforming Definition 4.12. This non-probabilistic NLMP is schematically depicted in Figure 6.3.

We also define the family $\mathcal{A} = \big\{\{s,t\}, \{r, f(r)\}_{r\in[0,1]}, \{x\}\big\}$ and the relation $R = \mathcal{R}(\sigma(\mathcal{A}))$.

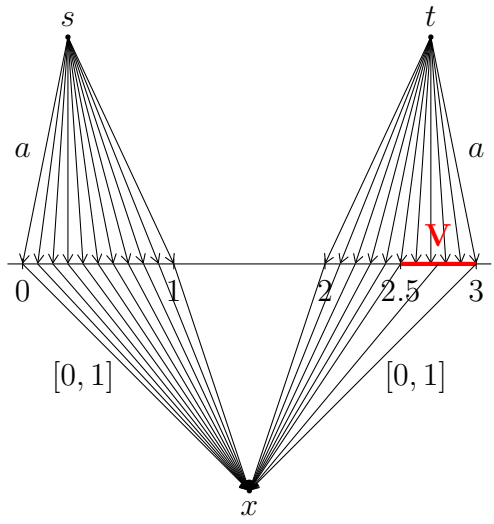


Figure 6.3: Non-probabilistic NLMP $\mathbf{S_1}$ showing that state and traditional bisimilarity differ.

Lemma 6.32. *$\sigma(\mathcal{A})$ is an event bisimulation and R is a state bisimulation,*

Proof. We first check that $\sigma(\mathcal{A})$ is an event bisimulation. Observe that for any $r \in [0,1]$, $\{r, f(r)\}$ is not separable in $\sigma(\mathcal{A})$. Thus for the label a and for all $Q \in \sigma(\mathcal{A})$, $\tilde{T}_a^{-1}(H_Q)$ is either empty or equal to $\{s,t\}$ and hence it

belongs to $\sigma(\mathcal{A})$. For the continuous labels $r \in [0,1]$, $\tilde{T}_r^{-1}(H_Q)$ is nonempty iff $x \in Q$, and in that case $\tilde{T}_r^{-1}(H_Q) = \{r, f(r)\} \in \sigma(\mathcal{A})$.

Then, to show that R is a state bisimulation, by Lemma 6.31 it suffices to show that $\tilde{T}_x^{-1}(H_Q) \in \Sigma_1(R)$ if $Q \in \Sigma_1(R)$. We divide again by labels. For $r \in [0,1]$, $\tilde{T}_r^{-1}(H_Q)$ is nonempty only if $x \in Q$, and in this case $\tilde{T}_r^{-1}(H_Q) = \{r, f(r)\}$. Therefore $\tilde{T}_r^{-1}(H_Q)$ is R-closed and Σ_1-measurable. For label a, $\tilde{T}_a^{-1}(H_Q) \in 2^{\{s,t\}}$, so, the only possibility not to obtain an R-closed set is separating s and t. If we want to obtain $\{s\}$ we have to take a measurable set $Q \subseteq [0,1]$. But we also need that Q is R-closed, so Q needs to have at least one point in $[2,3]\setminus V$, therefore t is also in the pre-image. The remaining choice is looking for measurable subsets of V since $\tilde{T}_a^{-1}(H_V) = \{t\}$, however there is no proper subset of V that is R-closed, and V itself is not Borel measurable. □

Theorem 6.33. *State bisimilarity (respectively, event bisimilarity) and traditional bisimilarity differ in* $\mathbf{S_1}$.

Proof. It suffices to show that s and t are not traditionally bisimilar. That is, if s and t are not traditionally bisimilar we are done because, by Lemma 6.32, $s \sim_s t$ and $s \sim_e t$ in $\mathbf{S_1}$.

First notice that for all $0 \leq r \leq 1$ we have $r \not\sim_t y$ if $y \notin \{r, f(r)\}$ since $\tilde{T}_r(y)$ is nonempty iff $y \in \{r, f(r)\}$. Therefore $\{r, f(r)\}$ is $\sim_t$ -closed for every $0 \leq r \leq 1$ and hence $\{r, f(r)\} \in \Sigma_1(\sim_t)$.

In order to show a contradiction, we now assume $s \sim_t t$ and take $y \in V \subseteq [2.5, 3]$. Since $y \in \tilde{T}_a(t)$, by definition of traditional bisimulation, there must be $0 \leq r \leq 1$ such that $y \, \mathcal{R}(\Sigma_1(\sim_t)) \, r$. However $1 < y$ and, by construction, y is not in the image of f. Therefore $\{r, f(r)\} \in \Sigma_1(\sim_t)$ separates y from r. So, for every $0 \leq r \leq 1$, it does not hold that $y \, \mathcal{R}(\Sigma_1(\sim_t)) \, r$, which contradicts the fact that $s \sim_t t$.

Since $\sim_s \subseteq \sim_e$, traditional bisimilarity and event bisimilarity also differ in $\mathbf{S_1}$. □

In the final part, we prove that the largest event bisimulation $\sim_e$ is not contained in $\sim_s$. We do this by slightly modifying $\mathbf{S_1}$. Instead of a nonmeasurable set $V \subseteq [2.5, 3]$ we pick the interval $I = (2.5, 3]$, and add self-loop transitions in the state x through the new label b[5].

Example 6.16. Take the standard Borel space $(S_2, \Sigma_2) = ([0,1] \uplus [2,3] \uplus \{s,t,x\}, \mathcal{B}([0,1] \uplus [2,3] \uplus \{s,t,x\}))$, where $\{s,t,x\} \subseteq \mathbb{R} \setminus [0,3]$. Let I be the measurable interval $(2.5, 3]$. We again establish a bijection f between $[0,1]$

[5]Carlos Budde pointed out a mistake in our original proof in [17, Theorem 6.9]. The proof we present here has also appeared in [10].

and $[2,3] \setminus I = [2, 2.5]$. The continuous label set is $L_2 = \{a, b\} \uplus [0,1]$, and the transition function $\tilde{T}_x$ is similar to that of Example 6.15 adding the self-loop in x through label b:

$$\begin{aligned}
&\tilde{T}_a(s) = [0,1] \\
&\tilde{T}_a(t) = [2,3] \\
&\tilde{T}_r(r) = \tilde{T}_r(f(r)) = \{x\} && \text{if } r \in [0,1] \\
&\tilde{T}_b(x) = \{x\} \\
&\tilde{T}_c(y) = \emptyset && \text{otherwise.}
\end{aligned}$$

$\mathbf{S_2} = (S_2, \Sigma_2, \{\tilde{T}_x \mid x \in L_2\})$ is a non-probabilistic NLMP. It is schematically represented in Figure 6.4.

We again define the family $\mathcal{A} = \big\{\{s,t\}, \{r, f(r)\}_{r\in[0,1]}, \{x\}\big\}$ and the relation $R = \mathcal{R}(\sigma(\mathcal{A}))$.

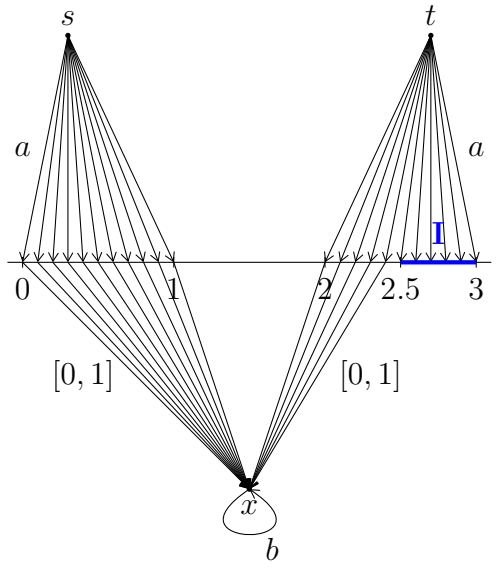


Figure 6.4: Non-probabilistic NLMP $\mathbf{S_2}$ showing that event and state bisimilarity differ.

Lemma 6.34. *$I \notin \sigma(\mathcal{A})$.*

Proof. It is clear that every member of $\sigma(\mathcal{A})$ is countable or has a countable complement, from which the lemma follows. □

In the following we prove that $R = \sim_e$ using Proposition 6.3 and showing that $\sigma(\mathcal{A})$ is the smallest event bisimulation for $\mathbf{S_2}$.

Lemma 6.35. *For non-probabilistic NLMP* $\mathbf{S_2} = (S_2, \Sigma_2, \{\tilde{T}_x : x \in L_2\})$, $\sigma(\mathcal{A})$ *is the smallest event bisimulation.*

Proof. The proof of Lemma 6.32 can be reused. We only have to add the minor observation that $\tilde{T}_b^{-1}(H_Q)$ is either $\emptyset$ or $\{x\}$, concluding that $\sigma(\mathcal{A})$ is an event bisimulation for $\mathbf{S_2}$.

Let $\Lambda \in \Sigma_2$ an arbitrary stable σ-algebra for NLMP $\mathbf{S_2}$ (Definition 6.13). In particular, $S_2 \in \Lambda$, and we calculate:

$$\begin{aligned} T_a^{-1}(H_{S_2}) &= \{s, t\} \\ T_r^{-1}(H_{S_2}) &= \{r, f(r)\}, \quad \forall r \in [0,1] \\ T_b^{-1}(H_{S_2}) &= \{x\} \end{aligned}$$

Since Λ is NLMP $\mathbf{S_2}$ stable, $\{\{s,t\}, \{r, f(r)\}_{r\in[0,1]}, \{x\}\} = \mathcal{A} \subseteq \Lambda$. By Proposition 3.1, we conclude $\sigma(\mathcal{A}) \subseteq \Lambda$, so it is the smallest stable σ-algebra for $\mathbf{S_2}$, that is, the smallest event bisimulation for $\mathbf{S_2}$. □

Theorem 6.36. *Event and state bisimilarity differ in* $\mathbf{S_2}$.

Proof. Since $(s,t) \in R = \sim_e$, we just have to show that $s \not\sim_s t$. Observe that $I \in \Sigma_2(R)$. If s and t were state-bisimilar, by Definition 6.14, it would be the case that $s \in \tilde{T}_a^{-1}(H_I)$ iff $t \in \tilde{T}_a^{-1}(H_I)$. But this is absurd since $\tilde{T}_a(s) \cap I = \emptyset$ and $3 \in \tilde{T}_a(t) \cap I$. □

6.3 Concluding Remarks

In this Chapter we reorganized the material found in [17], first revisiting the LMP notions of bisimulations and logical characterization to introduce the concepts and proof ideas in a simpler context. Some proofs for LMPs are new, most notably the proof corresponding to [14, Lemma 5.4]. That lemma has a small but solvable flaw in its proof. Besides, we used a slightly different hypothesis for the Dynkin Lemma (λ-systems instead of d-systems). Therefore we give an alternative proof in Lemma 6.12 based on the main idea underlying [65, Lemma 3.6]. The aforementioned lemma is used later in the proof of the logical characterization of event bisimulation for NLMPs.

In the proofs of Lemma 6.12 and Lemma 6.26 (required to show that event bisimilarity is completely characterized by $\mathcal{L}_s$ and $\mathcal{L}$, respectively), the use of $\mu(S) = 1$ is essential to capture complements. Therefore the logical characterization of event bisimulation is only valid for probability measures.

This can be easily generalized to finite measures including subprobabilities. However, for σ-finite measures it should be changed. One possibility, yet to be explored, is that the state logic captures an algebra instead of a π-system. Having this structure, the monotone family theorem (Theorem 3.10) can be applied to avoid any class of complementation in the proofs.

The logic $\mathcal{L}$ for uncountable branching NLMPs is similar (it was developed at the same time) to the logic given in [52]. The main difference is that we structured it better by considering two kinds of formulas: one that is interpreted over states, and the other that is interpreted over probability measures.

Counterexample 6.12 not only showed the need of a more expressive logic than $\mathcal{L}_s$, but it also evidenced another aspect where the definition of NLMPs is tight: the generators $H_{\Delta^{>q}(Q)}$ are not enough to generate a hit σ-algebra that captures finite nondeterminism, not to mention denumerable or continuous.

It is important to remark that there are weaker conditions than restricting NLMPs to finite nondeterminism to equate traditional and state bisimulation. One possibility is restricting to image denumerable NLMPs, other possibility is that the partitions induced in T_a by the bisimulation relation are also denumerable, and finally we could restrict to a measurable space such that $\Sigma(R)$ is countably generated.

Some additional observations on the counterexamples 6.15 and 6.16 are in order. First we restated the systems to conform non-probabilistic NLMPs of the form $S \to \Sigma$, instead of working in the restriction to Dirac measures $S \to \Delta(\Sigma)|\delta(S)$. In doing so, we rephrased the three notions of bisimilarity for these "measurable" LTSs. Second, counterexample $\mathbf{S_1}$ in Theorem 6.33 relies on the fact that a state bisimulation cannot distinguish a non-measurable set V while a traditional bisimulation can. From our point of view, such distinction should not be possible since V is not measurable. Third, counterexample $\mathbf{S_2}$ in Theorem 6.36 makes a distinction on the measurable set I that the event bisimulations cannot distinguish. In our opinion, such distinction should be observed since a possible *scheduler* may lead to such set of states with certain probability. Notice that in this example, states in I do not allow the system to reach state x from t, while x can always be reached from s. In this sense, we argue that state bisimulation is the most appropriate definition.

This is rather disappointing since the logic $\mathcal{L}$ has a natural definition but, as it completely characterizes event bisimulation, it will not be able to test the presence of states like those in I in $\mathbf{S_2}$. This is due to the fact that *the logic cannot test transitions bearing continuously many labels.*

For NLMPs with structure on the labels (S, Σ, Σ_L, T), the notions of bisimulations have to be revisited, most notably the event bisimulation. Also the logic should be updated to cover measurable modalities in the form $\langle A \rangle \psi$, with $A \in \Sigma_L$.

The examples $\mathbf{S_1}$ and $\mathbf{S_2}$ should also be revisited for NLMPs with structure on the labels. First checking if they are conforming to the definition, and then reviewing if they still serve as separating examples. Observe that $\mathbf{S_1}$ could have been used instead of Example 4.10. $\mathbf{S_1}$ conforms to Definition 4.9 but $T : S \to \Sigma_L \otimes \Delta(\Sigma)$ is not measurable, since $T^{-1}(H_{[0,1]\times\{x\}}) = [0,1] \uplus ([2,3] \setminus V)$. This might be showing that we have to strengthen Definition 4.9, to avoid the problematic $\mathbf{S_1}$.

We are also trying to refine the idea of event bisimulation and the logic so that they can distinguish situations like the one shown by the NLMP $\mathbf{S_2}$.

Chapter 7

Schedulers

In this chapter we show how to resolve the nondeterminism (external and internal) by means of history-dependent randomized schedulers. Once the nondeterminism has been resolved by probabilistic choices on an NLMP, the system is purely stochastic and a measure on traces of executions can be readily given.

A trace or path is a sequence of states, labels and measures, i.e. they live in a product space. Therefore, we will make use of different results on product σ-algebras and product measures as well as different results concerning transition measures.

This section extends [67], where the particular case over continuous time Markov decision processes (CTMDP) was developed. The results obtained here are similar, but they differ in the tools used in the proofs.

7.1 Constructing a Path Measure

We take the standard approach of resolving nondeterminism, that is, we resort to policies, adversaries or *schedulers* [54, 64]. Schedulers are functions that, given some information about the evolution of the system to the current state, they choose over all outgoing transitions in a probabilistic measure. The kind and amount of information taken by the function varies, ranging from taking into account just the current state (memoryless) to considering all the previous states, labels and distributions that were traversed up to the current state (history dependent). The applicability of these methods are strongly related to the power of the scheduler [3, 29]. Hence, we define it in the most general way so that the definition subsumes all known types.

First, we define the set of paths and its underlying σ-algebra.

Definition 7.1 (Paths)**.** Let (S, Σ) be the measurable space of states, (L, Σ_L) the measurable space of labels, and $(\Delta(S), \Delta(\Sigma))$ the measurable space of σ-finite measures. The measurable space of *paths of length* n is

$$(Path^1, \Sigma_{Path^1}) \doteq (S, \Sigma)$$
$$(Path^{n+1}, \Sigma_{Path^{n+1}}) \doteq (S \times (L \times \Delta(S) \times S)^n, \Sigma \otimes (\Sigma_L \otimes \Delta(\Sigma) \otimes \Sigma)^n)$$

We can take the (disjoint) sum of all finite paths of length n and define the measurable space of *finite paths*:

$$(Path^\star, \Sigma_{Path^\star}) = (\bigoplus_i Path^i, \bigoplus_i \Sigma_{Path^i})$$

We define the measurable space on *infinite paths* as follows:

$$(Path^\omega, \Sigma_{Path^\omega}) = (S \times (L \times \Delta(S) \times S)^\omega, \Sigma \otimes (\Sigma_L \otimes \Delta(\Sigma) \otimes \Sigma)^\omega)$$

Finally, the set of all *paths* and its σ-algebra are the sum of finite and infinite:

$$(Path, \Sigma_{Path}) = (Path^\star \oplus Path^\omega, \Sigma_{Path^\star} \oplus \Sigma_{Path^\omega})$$

For $\alpha \in Path^n$, $\alpha = s_1 a_2 \mu_2 s_2 \dots a_n \mu_n s_n$, the *last element* is $last(\alpha) \doteq s_n$.

Schedulers are defined on NLMPs with structure on the labels (S, Σ, Σ_L, T) (see Definition 4.9). Since in this chapter we only use NLMPs with structure on the labels, we will shortly refer to them only as NLMPs. Schedulers are functions from $Path^\star$ to a *probability* distribution concentrated on the outgoing transitions.

Definition 7.2 (Scheduler)**.** The function $\eta : Path^\star \to \Delta^{=1}(L \times \Delta(S))$ is a *scheduler* of NLMP (S, Σ, Σ_L, T), if it is a measurable function, and for all $\alpha \in Path^\star$, $\eta(\alpha)((L \times \Delta(S)) \setminus T(last(\alpha))) = 0$.

We remark that we are bonding the scheduler to the model, since it can only choose among existing outgoing transitions of $last(\alpha)$. Also notice we are making use of the added structure of the NLMPs, namely the σ-algebra Σ_L on labels and that $T(s) \in \Sigma_L \otimes \Delta(\Sigma)$, so $\eta(\alpha)$ is well defined.

Having resolved the nondeterminism, the quantification of the scheduler and the model can be combined [58].

Definition 7.3 (Combined transition)**.** Let (S, Σ, Σ_L, T) be an NLMP such that T contains only σ-finite measures. Let η be a scheduler on such NLMP. The *combined transition* $\mu_{\eta(\cdot)}(\cdot) : Path^\star \times (\Sigma_L \otimes \Delta(\Sigma) \otimes \Sigma) \to [0, 1]$, is defined:

$$\mu_{\eta(\alpha)}(A \times \xi \times Q) \doteq \int_{A \times \xi} (\lambda \nu : \nu(Q))(\mu)\, \eta(\alpha)(da, d\mu)$$

where $A \in \Sigma_L$, $\xi \in \Delta(\Sigma)$, $Q \in \Sigma$.

Observe that the integral is well-defined since $(\lambda\nu : \nu(Q)) : \Delta(S) \to [0,1]$ is measurable. The next lemma shows that a combined transition is a measure in its second coordinate and a measurable function in its first coordinate, therefore it is a conditional measure.

Lemma 7.1 (Combined transition is a conditional measure)**.** *Let the combined transition μ_η be as in Definition 7.3. For any finite path $\alpha \in Path^\star$, the combined transition $\mu_{\eta(\alpha)}(\cdot)$ extends uniquely to the following σ-finite measure:*

$$\mu_{\eta(\alpha)}(M) = \int (\lambda\nu : \nu(M_{|a,\mu}))(\mu)\, \eta(\alpha)(da, d\mu)$$

where $M \in \Sigma_L \otimes \Delta(\Sigma) \otimes \Sigma$. Moreover, for all $M \in \Sigma_L \otimes \Delta(\Sigma) \otimes \Sigma$, $\mu_{\eta(\cdot)}(M)$ is a measurable function.

Proof. First we prove that given a fixed $\alpha \in Path^\star$, $\mu_{\eta(\alpha)}(\cdot)$ defines a σ-finite measure. We apply the product measure theorem (see Theorem 3.42). The measurable spaces are $(S_1, \Sigma_1) = (L \times \Delta(S), \Sigma_L \otimes \Delta(\Sigma))$, and $(S_2, \Sigma_2) = (S, \Sigma)$. The measure on the first experiment is $\mu_1(A_1) = \eta(\alpha)(A_1)$ with $A_1 \in \Sigma_L \otimes \Delta(\Sigma)$. The conditional measure of the second experiment given the first experiment value is $\mu_2((a,\mu), A_2) = ((\lambda\nu : \nu(A_2)) \circ \pi_2)((a,\mu))$, where $(a,\mu) \in L \times \Delta(S)$ and $A_2 \in \Sigma$.

We have to check that μ_1 is σ-finite, and μ_2 is uniformly σ-finite in its second coordinate. The scheduler is a probability measure, therefore μ_1 is σ-finite. For the conditional probability we develop the expression where $(a,\mu) \in L \times \Delta(S)$, and $A_2 \in \Sigma$:

$$\mu_2((a,\mu), A_2) = ((\lambda\nu : \nu(A_2)) \circ \pi_2)((a,\mu)) = \mu(A_2)$$

Since the pairs (a,μ) come from the NLMPs that is restricted to σ-finite measures (the support of the scheduler is on transitions), μ_2 is uniformly σ-finite in its first coordinate. Besides $\mu_2(\cdot, A_2)$ is measurable for all $A_2 \in \Sigma_2$ since it is a composition of two measurable functions. The projection π_2 is measurable by definition of the product σ-algebra, and $(\lambda\nu : \nu(A_2))$ is measurable by the definition of σ-algebra on measures.

Secondly we prove that $\mu_{\eta(\cdot)}(M)$ is measurable for all $M \in \Sigma_L \otimes \Delta(\Sigma) \otimes \Sigma$. We will use the intermediate result of Fubini theorem (see Theorem 3.44) that establishes that $\int f(x_1, x_2)\mu(x_1, dx_2)$ is well defined and measurable on x_1.

We take $(S_1, \Sigma_1) = (Path^\star, \Sigma_{Path^\star})$ and $(S_2, \Sigma_2) = (L \times \Delta(S), \Sigma_L \otimes \Delta(\Sigma))$. The integrand is $f(\alpha, (a,\mu)) = ((\lambda\nu : \nu(M_{|(a,\mu)})) \circ \pi_2)((a,\mu))$, where $\alpha \in Path^\star$, and $(a,\mu) \in L \times \Delta(S)$. The transition measure is $\mu_2(\alpha, A_2) = \eta(\alpha)(A_2)$, with $\alpha \in Path^\star$ and $A_2 \in \Sigma_L \otimes \Delta(\Sigma)$.

Checking that $\mu_2(\alpha, A_2) = \eta(\alpha)(A_2)$ is a uniformly σ-finite transition measure is direct, given that the scheduler is a transition probability.

We prove now that $f(\alpha,(a,\mu))$ is measurable for all $M \in \Sigma_L \otimes \Delta(\Sigma) \otimes \Sigma$, i.e. that $(\lambda\nu : \nu(M_{|(a,\mu)}))(\mu) : L \times \Delta(S) \to [0,1]$ is measurable. For this, we use the the monotone family theorem (see Theorem 3.10). Let the good sets be

$$\mathcal{G} = \{M \in \Sigma_L \otimes \Delta(\Sigma) \otimes \Sigma \mid ((\lambda\nu : \nu(M_{|(a,\mu)})) \circ \pi_2)((a,\mu)) \text{ is measurable}\}$$

The family of finite disjoint unions of measurable rectangles $\biguplus_{i=1}^n (A_i \times \xi_i \times Q_i)$ generates $\Sigma_L \otimes \Delta(\Sigma) \otimes \Sigma$, besides they form an algebra. However, we still have to check that they are good. We develop the expression for $M = \biguplus_{i=1}^n (A_i \times \xi_i \times Q_i)$:

$$((\lambda\nu : \nu(\biguplus_{i=1}^n (A_i \times \xi_i \times Q_i)_{|(a,\mu)})) \circ \pi_2)((a,\mu)) = \sum_{i=1}^n (\lambda\nu : \nu(Q_i))(\mu)\, \chi_{A_i}(a)\, \chi_{\xi_i}(\mu)$$

The last expression is measurable since it is the sum of products of measurable functions. If we verify that $\mathcal{G}$ is a monotone family, we are done, and hence $f(\alpha,(a,\mu))$ is measurable for all $M \in \Sigma_L \otimes \Delta(\Sigma) \otimes \Sigma$.

Observe that $((\lambda\nu : \nu(M_{|(a,\mu)})) \circ \pi_2)((a,\mu)) = \mu(M_{|(a,\mu)})$. Let $\{M_i\}_i \subseteq \mathcal{G}$. If $M_i \nearrow M$, then by continuity of the measure (see Theorem 3.18), $((\lambda\nu : \nu(M_{i|(a,\mu)})) \circ \pi_2)((a,\mu)) \nearrow ((\lambda\nu : \nu(M_{|(a,\mu)})) \circ \pi_2)((a,\mu))$. Given that for all i, $((\lambda\nu : \nu(M_{i|(a,\mu)})) \circ \pi_2)((a,\mu))$ is measurable, by Theorem 3.26 its limit is also measurable. The case $M_i \searrow M$ is similar, but Theorem 3.18 requires that $((\lambda\nu : \nu(M_{1|(a,\mu)})) \circ \pi_2)((a,\mu)) < \infty$. This can be saved by taking intersections with a partition $S = \biguplus_i B_i$ such that $\mu(B_i) < \infty$, and such a partition exists since all the measures from the NLMPs are σ-finite. □

Combined transition is the main ingredient to define measures on paths.

Definition 7.4 (Finite path measure)**.** Let (S, Σ, Σ_L, T) be an NLMP with σ-finite measures. Let η be a scheduler on such NLMP. If $\nu \in \Delta(S)$ is the σ-finite *initial measure* on states, the *finite path measure* on the measurable rectangles of $((Path^n, \Sigma_{Path^n}))_n$ is defined recursively by:

$$P^1_{\nu,\eta}(M_1) \doteq \nu(M_1)$$

$$P^{n+1}_{\nu,\eta}((\textstyle\prod_{i=1}^n M_i) \times M_{n+1}) \doteq \int_{(\prod_{i=1}^n M_i)} \mu_{\eta(\alpha)}(M_{n+1}) P^n_{\nu,\eta}(d\alpha)$$

where $M_1 \in \Sigma$, for $1 < i \leq n+1$, $M_i \in \Sigma_L \otimes \Delta(\Sigma) \otimes \Sigma$, and μ_η is the combined transition of Definition 7.3.

Observe that this is well defined since $\mu_{\eta(\alpha)}(M)$ is a measurable function in its first coordinate. However it is still needed that $P^n_{\nu,\eta}$, the σ-finite path measure extends uniquely to the whole σ-algebra Σ_{Path^n}.

Lemma 7.2. *Let $P^n_{\nu,\eta}$ be as in Definition 7.4. Then for all n, $P^n_{\nu,\eta}$ extends uniquely to a σ-finite measure on the whole σ-algebra of Σ_{Path^n}.*

Proof. We fix n and apply the finite product measure theorem (see Theorem 3.46) with measurable spaces $((S_i, \Sigma_i))_{i=1}^n = ((Path^i, \Sigma_{Path^i}))_{i=1}^n$. The initial measure on states is $\mu_1(A_1) = \nu(A_1)$, with $A_1 \in \Sigma$. For $1 < i \leq n$, the uniformly σ-finite conditional measures are defined by $\mu_i(s_1, \ldots, s_{i-1}, A_i) = \mu_{\eta(s_1,\ldots,s_{i-1})}(A_i)$, where $s_1 \in S$, for all $1 \leq j < i$, $s_j \in L \times \Delta(S) \times S$, and $A_i \in \Sigma_L \otimes \Delta(\Sigma) \otimes \Sigma$. □

The previous result allows to define the measure spaces of finite paths.

Definition 7.5 (Finite path measure spaces)**.** The *measure space for paths of length n* is $(Path^n, \Sigma_{Path^n}, P^n_{\nu,\eta})$, where $P^n_{\nu,\eta}$ is the σ-finite measure of Definition 7.4. The sum measure space (see Definition 3.35) of all the finite path measures is $(Path^\star, \Sigma_{Path^\star}, P^\star_{\nu,\eta})$, where $P^\star_{\nu,\eta}(A) = P^i_{\nu,\eta}(A)$ if $A \in \Sigma_{Path^i}$.

Notice that if the NLMP is restricted to probability measures, Lemma 7.1 defines a *transition probability*, and Lemma 7.2 defines a probability measure for finite paths. Therefore we can define the probability measure for infinite paths.

Definition 7.6. Let (S, Σ, Σ_L, T) be an NLMP with probability measures. Let η be a scheduler on such NLMP. If $\nu \in \Delta^{=1}(S)$ is the initial probability measure on states, then the probability measure on the measurable rectangles of $(Path^\omega, \Sigma_{Path^\omega})$ (see Definition 3.13) is defined by:

$$P^\omega_{\nu,\eta}((\textstyle\prod_{i=1}^n A_i) \times (\prod_i (L \times \Delta(S) \times S))) \doteq P^n_{\nu,\eta}(\textstyle\prod_{i=1}^n A_i)$$

where $A_1 \in \Sigma$ and for $1 < i \leq n$ $A_i \in \Sigma_L \otimes \Delta(\Sigma) \otimes \Sigma$.

This probability measure extends uniquely to the whole σ-algebra Σ_{Path^ω}.

Lemma 7.3. *Let $P^\omega_{\nu,\eta}$ be as in Definition 7.6. Then, $P^\omega_{\nu,\eta}$ extends uniquely to a probability measure on the whole σ-algebra of Σ_{Path^ω}.*

Proof. Let $(S_1, \Sigma_1) = (S, \Sigma)$ and for $1 < i \leq n$, $(S_i, \Sigma_i) = (L \times \Delta(S) \times S, \Sigma_L \otimes \Delta(\Sigma) \otimes \Sigma)$ be measurable spaces. Let $\mu(\prod_{i=1}^n A_i) = P^n_{\nu,\eta}(\prod_{i=1}^n A_i)$, where $P^n_{\nu,\eta}$ is from Definition 7.4, $A_1 \in \Sigma$, and for $1 < i \leq n$ $A_i \in \Sigma_L \otimes \Delta(\Sigma) \otimes \Sigma$. Then the result follows directly by Theorem 3.47. □

Our previous result is the base to define trace probability. A trace probability on an NLMP is a probability measure on the infinite sequence of labels. It is induced by a scheduler and an initial probability distribution.

Definition 7.7 (Trace probability)**.** Let $\mathbf{S} = (S, \Sigma, \Sigma_L, T)$ be an NLMP with probability measures. Let η be a scheduler on such NLMP, and ν an initial probability on S. If $trace : Path^\omega \to L^\omega$ is the projection function, i.e. $trace(s_1 a_2 \mu_2 s_2 a_3 \mu_3 s_3 \dots) = a_2 a_3 \dots$, we define the *trace probability* $\mathbb{T} \in \Delta(L^\omega)$ of NLMP $\mathbf{S}$ induced by scheduler η and initial distribution ν as follows:

$$\mathbb{T}_{\nu,\eta} = P^\omega_{\nu,\eta} \circ trace^{-1}$$

where $P^\omega_{\nu,\eta}$ is the probability measure on infinite paths of Definition 7.6.

Trace probability is a probability measure by Proposition 3.19, since *trace* is a measurable function. The set of all trace probabilities is the so called trace semantics [32].

Definition 7.8 (Trace semantics)**.** Let $\mathbf{S} = (S, \Sigma, \Sigma_L, T)$ be an NLMP with probability measures. The *trace semantics of* $\mathbf{S}$ is a set of trace probabilities induced by schedulers η and initial probability distributions ν over $\mathbf{S}$:

$$\mathbb{T} = \{\mathbb{T}_{\nu,\eta} \mid \nu \in \Delta(S), \eta \text{ is a scheduler of } \mathbf{S}\}$$

Chapter 8

Conclusions

In this thesis we proposed a definition of labeled nondeterministic conditional measures $T_a : S \to \Delta(\Sigma)$. This definition is sound with respect to Measure Theory, and it is the main ingredient of the NLMP model. The extension from labeled conditional (sub)probabilities $\tau_a : S \to \Delta(S)$ and its related LMP model is far from trivial. This is why the cover of this dissertation is decorated with the definition of the *transition function* for NLMPs (Definition 4.3). The other cornerstones of this thesis are the *hit σ-algebra* (Definition 4.4) and *state bisimulation* (Definition 6.9).

Throughout this thesis we were very careful not to put more structure than needed in the definitions, and the definition of NLMPs itself represents a clear example of this claim. Most of the definitions are in a pure measure theoretical setting, avoiding, whenever possible, any topological structure. Our definitions and results are mostly *topology-free* [65].

We obtained a good amount of results, but a lot is yet to be done. Next we summarize our contributions and what will come in the future.

8.1 Achievements

The following list provides a summary of the contributions of this thesis.

- Measure theoretic topology-free definition of labeled nondeterministic conditional measures, that is, the transition function of NLMPs.
- The use of a hit σ-algebra $H(\Sigma)$ to capture existential quantification.
- Two variations of the definition of NLMPs, one encompassing a less expressive model (LTS with a σ-algebra attached), while the other strengthens the definition of NLMPs including structure to the labels.

- An initial description on how to use the generators of the σ-algebra of measures $\Delta(\Sigma)$ and denumerable set operations as a language to specify continuous nondeterministic choices of continuous probabilities.
- The semantics of three different models that involve nontrivial use of nondeterministic probabilistic choice.
- Comparisons with established similar models that can represent continuous nondeterminism in probabilistic choice.
- Three alternative definitions of bisimulation, proving strict inclusions between them in the general setting.
- An infinitary logic characterizing event bisimulation, and a finitary logic characterizing all three bisimulations on a finitely branching NLMPs restricted to analytic spaces.
- A sound definition of scheduler for NLMPs with structure on the labels, giving rise to probabilistic trace semantics.

8.2 Future Research Directions

Event bisimulation was shown to be too coarse. This was exposed in Example 6.16. Traditional bisimulation on the contrary is too fine, since it distinguishes states over sets that are not measurable (see Example 6.15). It seems that the correct generalization is the intermediate notion of *state bisimulation* for NLMPs, but still more evidence is needed.

Example 4.10 suggests that some σ-algebra structure is also needed in the labels. Moreover, Example 6.15 is still conforming to the definition of NLMPs with structure on the labels (Definition 4.9), and it might be the case it also has to be discarded. If we strengthen the transition relation asking that $T : S \to \Sigma_L \otimes \Delta(\Sigma)$ to be measurable, Example 6.15 is discarded since the transition relation is not measurable; however, Example 6.16 is still conforming. If this research direction is taken, new definitions of all three bisimulations and a new logic should be given, and there, Example 6.16 should be revisited. For the traditional bisimulations the modifications are minor since they do not take into account the structure on the labels, and $T_{|a}$ has the desired properties. The state bisimulation situation with respect to this new structure is somehow similar. The event bisimulation will differ since the measurability of T now deals with a more complex hit σ-algebra. The modifications to be done in the logic are not minor. The modal operator should be able to capture measurability of the new transition function, i.e.

$T^{-1}(H_A)$ where $A \in \Sigma_L \otimes \Delta(\Sigma)$. The second level of the logic should be restructured, and it is not clear how to do that.

Currently we are busy working on the definition of NLMPs with structure on the labels $T : S \to \Sigma_L \otimes \Delta(\Sigma)$, where the hit σ-algebra attached to the image is generated by the smaller family $\{H_{A\times\xi} \mid A \in \Sigma_L, \xi \in \Delta(\Sigma)\}$. In this case the logic is easier to give: the modality uses measurable sets of labels instead of individual labels. The syntax of modality is $\langle A\rangle\psi$, with $A \in \Sigma_L$, while the semantics is defined as $[\![\langle A\rangle\psi]\!] = T^{-1}(H_{A\times[\![\psi]\!]})$. However it is not clear if this coarser version has enough power to differentiate some continuous labeled systems, like "lines" in the form $T(s) = \{(t, \delta_t) \mid t \in [0, 1]\}$. Notice that examples of this kind already appeared in Section 5.2. We already proved that under this definition of NLMPs, related hit sets, and previously discussed logic, it its not the case that $s \sim_e t$ in Example 6.16. The formula $\langle a\rangle\neg[\langle[0, 1]\rangle[\top]_{>0}]_{>0}$ differentiates s and t. These results are partially explored in [10].

It would be interesting to draw conclusions whether NLMPs can capture infLMPs or not. We have seen that under reasonable restrictions of the σ-algebra the realizations of a super-additive functions are $\Delta(\Sigma)$ measurable. It remains to be proven that the measurability of the super-additive transition function is inherited by the NLMP translation. In general the infLMP model is very promising, although it lacks a definition of a scheduler. Once Desharnais et al. obtain a definition of scheduler in the infLMP model, this scheduler should be compared to ours.

We aim to obtain a result on probabilistic trace equivalence for bisimilar states, the so called *execution correspondence lemma* [58]. Namely, if $s \sim t$ for some notion of bisimilarity $\sim \in \{\sim_t, \sim_s, \sim_e\}$, then the probabilistic trace semantics with initial probability δ_s, is equal to the probabilistic trace semantics with initial probability δ_t. Although this is a desired result, it has eluded us so far.

One advantage of NLMPs with respect to LMPs, besides the obvious incorporation of internal nondeterminism, is that the former do not restrict to (sub)probability measures. The definition includes every measure, from general ones to probabilities. We carefully pointed out where some restrictions (usually σ-finite measures) were needed. However in the proofs involving the logic, finiteness of the measures is strongly used. It would be desirable to avoid this restriction, and in doing so, perhaps we have to move to a slightly more complex logic also capturing complements at the state level (technically speaking, we need an algebra). Moreover, it would be important to enhance those results providing examples on the fields of Economics, Physics

or Biology that require the use of NLMPs with σ-finite measures beyond (sub)probabilities.

The aspect of using $\Delta^{\bowtie q}(Q)$ and countable set operations as a specification language for nondeterministic probabilistic choices, is promising, however more work is needed. Not only Example 4.7 is important on showing an NLMP with a continuous nondeterminism that depends on the current state. It also shows a promising line to explore in the field of Model Checking. Systems in this form can be checked against formulas of the logic $\mathcal{L}$. Take an NLMP $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), \{T_a \mid a \in L\})$, where $T_a(x)$ is given by means of set operations on generators $\Delta^{\bowtie q(x)}([p_1(x), p_2(x)])$, and the functions q, p_1, p_2 are restricted to be linear. Then the satisfiability relation between formulas in $\mathcal{L}$ and NLMP models, reduces to the satisfiability of a Boolean expression on linear inequalities over interval endpoints. From this point on, a satisfiability modulo theories (SMT) solver could resolve the model checking problem. At first sight the only drawback is that model and formulas basically use the same language. This language is good for specification of properties, since very general properties can be captured by small expressions. However for concrete model description the set expression would need many $\Delta^{\bowtie q(x)}([p_1(x), p_2(x)])$, or even denumerably many as in the measurable expression for $\delta(Q)$, rendering the model checking procedure noncomputable. Approximation and abstraction techniques should be developed in order to alleviate this problem.

An NLMP plus a scheduler defines a path probability. By the Radon-Nikodym theorem [2] we can do the converse, namely if a path probability is "compatible" with an NLMP [67], then there is a scheduler that gives rise to this path probability. This would show a kind of completeness result for schedulers, since there is no path probability out of a given NLMP that is not generated by a scheduler. A result like this would extend [67] from CTMDP to much more general systems.

Finally this thesis can be considered as an important first step towards a modeling formalism to capture continuous systems *as is*. Once we establish simulation and approximation results, discrete approximated systems that are sound with respect to a subset of formulas could be verified.

Bibliography

[1] C. D. Aliprantis and K. C. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer, third edition, 2006.

[2] R. Ash and C. Doléans-Dade. *Probability & Measure Theory*. Academic Press, 2000.

[3] C. Baier, B. Haverkort, H. Hermanns, and J.-P. Katoen. Efficient computation of time-bounded reachability probabilities in uniform continuous-time markov decision processes. In *TACAS'04*, volume 2988 of *LNCS*, pages 61–76. Springer, 2004.

[4] P. Billingsley. *Probability and Measure*. Wiley-Interscience, second edition, 1986.

[5] A. Bouchard-Cote, N. Ferns, P. Panangaden, and D. Precup. An approximation algorithm for labelled markov processes: towards realistic approximation. In *QEST'05*, pages 54–62. IEEE Computer Society, 2005.

[6] M. S. Branicky. Introduction to hybrid systems. In *Handbook of Networked and Embedded Control Systems*, pages 91–116. Birkhäuser, 2005.

[7] M. Bravetti. *Specification and Analysis of Stochastic Real-Time Systems*. PhD thesis, Università di Bologna, Padova, Venezia, 2002.

[8] M. Bravetti and P. D'Argenio. Tutte le algebre insieme: Concepts, discussions and relations of stochastic process algebras with general distributions. In *Validation of Stochastic Systems*, volume 2925 of *LNCS*, pages 44–88. Springer, 2004.

[9] L. Breiman. *Probability*. Addison-Wesley, 1968.

[10] C. E. Budde. No determinismo completamente medible en procesos probabilísticos continuos. Master's thesis, FaMAF, Universidad Nacional de Córdoba, 2012.

[11] S. Cattani. *Trace-based Process Algebras for Real-Time Probabilistic Systems*. PhD thesis, University of Birmingham, 2005.

[12] S. Cattani, R. Segala, M. Kwiatkowska, and G. Norman. Stochastic transition systems for continuous state spaces and non-determinism. In *FoSSaCS'05*, volume 3441 of *LNCS*, pages 125–139. Springer, 2005.

[13] P. Celayes. Procesos de Markov etiquetados sobre espacios de Borel estándar. Master's thesis, FaMAF, Universidad Nacional de Córdoba, 2006.

[14] V. Danos, J. Desharnais, F. Laviolette, and P. Panangaden. Bisimulation and cocongruence for probabilistic systems. *Inf. & Comp.*, 204:503–523, 2006.

[15] P. D'Argenio. *Algebras and Automata for Timed and Stochastic Systems*. PhD thesis, Department of Computer Science, University of Twente, 1999.

[16] P. D'Argenio and J.-P. Katoen. A theory of stochastic systems, Part I: Stochastic automata, *and* Part II: Process algebra. *Inf. & Comp.*, 203(1):1–38, 39–74, 2005.

[17] P. R. D'Argenio, P. S. Terraf, and N. Wolovick. Bisimulations for nondeterministic labeled Markov processes. *Math. Struct. in Comp. Science*, 22(1):43–68, 2012.

[18] P. R. D'Argenio, N. Wolovick, P. S. Terraf, and P. Celayes. Nondeterministic labeled Markov processes: Bisimulations and logical characterization. In *QEST'09*, pages 11–20. IEEE Computer Society, 2009.

[19] B. Delahaye, J.-P. Katoen, K. G. Larsen, A. Legay, M. L. Pedersen, F. Sher, and A. Wasowski. Abstract probabilistic automata. In *VMCAI'11*, volume 6538 of *LNCS*, pages 324–339. Springer, 2011.

[20] J. Desharnais. *Labeled Markov Processes*. PhD thesis, McGill University, 1999.

[21] J. Desharnais, A. Edalat, and P. Panangaden. Bisimulation for labelled Markov processes. *Inf. & Comp.*, 179(2):163–193, 2002.

[22] J. Desharnais, F. Laviolette, and A. Turgeon. A logical duality for underspecified probabilistic systems. *Inf. & Comp.*, 209(5):850–871, 2011.

[23] J. Desharnais and P. Panangaden. Continuous stochastic logic characterizes bisimulation of continuous-time markov processes. *J. Log. Algebr. Program.*, 56(1–2):99–115, 2003.

[24] E. W. Dijkstra. *A Discipline of Programming*. Prentice-Hall, 1976.

[25] E. Doberkat. *Stochastic relations: foundations for Markov transition systems*. Chapman & Hall/CRC studies in informatics. Chapman & Hall/CRC, 2007.

[26] L. Dubins and D. Freedman. Measurable sets of measures. *Pacific J. Math.*, 14(4):1211–1222, 1964.

[27] E. Dynkin. *Markov Processes*. Number 121-122 in Die Grundlehren der Math. Wissenschaften. Springer-Verlag, 1965.

[28] M. Fränzle, E. M. Hahn, H. Hermanns, N. Wolovick, and L. Zhang. Measurability and safety verification for stochastic hybrid systems. In *HSCC'11*, pages 43–52. ACM, 2011.

[29] S. Giro and P. R. D'Argenio. On the expressive power of schedulers in distributed probabilistic systems. *Electr. Notes Theor. Comput. Sci.*, 253(3):45–71, 2009.

[30] M. Giry. A categorical approach to probability theory. In *Categorical Aspects of Topology and Analysis*, pages 68–85. Springer, 1982.

[31] R. v. Glabbeek. The linear time – branching time spectrum II; the semantics of sequential systems with silent moves (extended abstract). In *CONCUR'93*, volume 715 of *LNCS*, pages 66–81. Springer, 1993.

[32] R. v. Glabbeek. The linear time – branching time spectrum I; The semantics of concrete, sequential processes. In *Handbook of Process Algebra*, chapter 1, pages 3–99. Elsevier, 2001.

[33] E. M. Hahn. Personal communication. Saarland University, 2011.

[34] D. J. Hartfiel. *Markov Set-Chains*, volume 1695 of *LNM*. Springer, 1998.

[35] T. A. Henzinger, P. W. Kopke, A. Puri, and P. Varaiya. What's decidable about hybrid automata? *Journal of Computer and System Sciences*, 57:94–124, 1998.

[36] A. Kechris. *Classical Descriptive Set Theory*, volume 156 of *Graduate Texts in Mathematics*. Springer, 1995.

[37] R. M. Keller. Formal verification of parallel programs. *Communications of the ACM*, 19(7):371–384, 1976.

[38] J. G. Kemeny and J. L. Snell. *Finite Markov Chains*. Van Nostrand, 1960.

[39] M. Kwiatkowska, G. Norman, R. Segala, and J. Sproston. Verifying quantitative properties of continuous probabilistic timed automata. In *CONCUR'00*, volume 1877 of *LNCS*, pages 123–137. Springer, 2000.

[40] M. Kwiatkowska, G. Norman, R. Segala, and J. Sproston. Automatic verification of real-time systems with discrete probability distributions. *TCS*, 282:101–150, 2002.

[41] K. Larsen and A. Skou. Bisimulation through probabilistic testing. *Inf. & Comp.*, 94(1):1–28, 1991.

[42] F. W. Lawvere. The category of probabilistic mappings. Unpublished notes, 1962.

[43] R. Milner. A calculus of communicating systems. *LNCS*, 92, 1980.

[44] R. Milner. *Communication and Concurrency*. Prentice Hall, 1989.

[45] R. Milner. *Communicating and Mobile Systems: The π-calculus*. Cambridge University Press, 1999.

[46] C. Morgan and A. McIver. *Abstraction, Refinement and Proofs for Probabilistic Programs*. Monographs in Computer Science. Springer, 2004.

[47] L. S. Moss and I. D. Viglizzo. Final coalgebras for functors on measurable spaces. *Inf. & Comp.*, 204(4):610–636, 2006.

[48] S. Naimpally. What is a hit-and-miss topology? *Topological Comment.*, 8(1), 2003.

[49] P. Panaganden. Stochastic techniques in concurrency. Unpublished lecture notes from a course given at BRICS, 2000.

[50] P. Panangaden. *Labelled Markov Processes*. Imperial College Press, 2009.

[51] D. Park. Concurrency and automata on infinite sequences. In *LNCS*, volume 104, pages 167–183. Springer-Verlag, 1981.

[52] A. Parma and R. Segala. Logical characterizations of bisimulations for discrete probabilistic systems. In *FoSSaCS'07*, volume 4423 of *LNCS*, pages 287–301. Springer, 2007.

[53] R. Penrose. *Techniques of Differential Topology in Relativity*. Society for Industrial Mathematics, 1972.

[54] M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley-Interscience, 1994.

[55] S. Ratschan and Z. She. Safety verification of hybrid systems by constraint propagation based abstraction refinement. *ACM Transactions on Embedded Computing Systems*, 6(1), 2007.

[56] P. Sánchez Terraf. Unprovability of the logical characterization of bisimulation. *Inf. & Comp.*, 209(7):1048–1056, 2011.

[57] R. Segala. A compositional trace-based semantics for probabilistic automata. In *CONCUR'95*, volume 962 of *LNCS*, pages 234–248, 1995.

[58] R. Segala. *Modeling and verification of randomized distributed real-time systems*. PhD thesis, Massachusetts Institute of Technology, Dept. of Electrical Engineering and Computer Science, 1995.

[59] R. Segala and N. Lynch. Probabilistic simulations for probabilistic processes. *Nordic Journal of Computing*, 2(2):250–273, 1995.

[60] C. P. Stirling. *Modal and Temporal Properties of Processes*. Texts in Computer Science. Springer, 2001.

[61] M. I. A. Stoelinga. *Alea Jacta Est: Verification of Probabilistic, Real-Time and Parametric Systems*. PhD thesis, University of Nijmegen, 2002.

[62] B. Strulo. *Process Algebra for Discrete Event Simulation*. PhD thesis, Department of Computing, Imperial College, University of London, 1993.

[63] D. Varacca. Probabilistic models for concurrency. Notes for a FIRST minicourse, ITU Copenhagen, 2005.

[64] M. Y. Vardi. Automatic verification of probabilistic concurrent finite-state programs. In *Foundations of Computer Science*, pages 327–338, 1985.

[65] I. D. Viglizzo. *Coalgebras on Measurable Spaces*. PhD thesis, Indiana University, 2005.

[66] R. Wheeden and A. Zygmund. *Measure and Integral*. Dekker, 1977.

[67] N. Wolovick and S. Johr. A characterization of meaningful schedulers for continuous-time markov decision processes. In *FORMATS'06*, volume 4202 of *LNCS*, pages 352–367. Springer, 2006.

[68] L. Zhang, Z. She, S. Ratschan, H. Hermanns, and E. M. Hahn. Safety verification for probabilistic hybrid systems. In *CAV'10*, volume 6174 of *LNCS*, pages 196–211, 2010.

Revision: 3626, (2012-06-21)