The Theory of Traces for Systems with Nondeterminism and Probability

Filippo Bonchi
University of Pisa, Italy

Ana Sokolova
University of Salzburg, Austria

Valeria Vignudelli
CNRS/ENS Lyon, France

***Abstract*—This paper studies trace-based equivalences for systems combining nondeterministic and probabilistic choices. We show how trace semantics for such processes can be recovered by instantiating a coalgebraic construction known as the generalised powerset construction. We characterise and compare the resulting semantics to known definitions of trace equivalences appearing in the literature. Most of our results are based on the exciting interplay between monads and their presentations via algebraic theories.**

1. Introduction

Systems exhibiting both nondeterministic and probabilistic behaviour are abundantly used in verification [1], [2], [3], [4], [5], [6], [7], AI [8], [9], [10], and studied from semantics perspective [11], [12], [13]. Probability is needed to quantitatively model uncertainty and belief, whereas nondeterminism enables modelling of incomplete information, unknown environment, implementation freedom, or concurrency. At the same time, the interplay of nondeterminism and probability has been posing some remarkable challenges [14], [15], [16], [17], [18], [19], [20], [21].

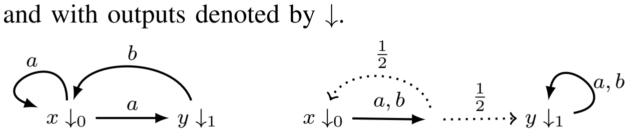
Figure 1 shows a nondeterministic probabilistic system (NPLTS) that we use as a running example.

Traces and trace semantics [22] for nondeterministic probabilistic systems have been studied for several decades within concurrency theory and AI using resolutions or schedulers—entities that resolve the nondeterminism. Most proposals of trace semantics in the literature [23], [24], [25], [26] are based on such auxilary notions of resolutions and differ on how these resolutions are defined and combined. We call such approaches *local-view* approaches.

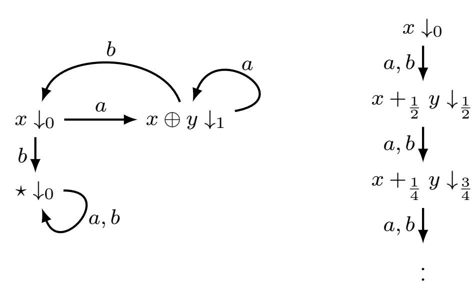
On the other hand, the theory of coalgebra [27], [28] provides uniform generic approaches to trace semantics of various kinds of systems and automata, via Kleisli traces [29] or generalised determinisation [30], providing e.g. an abstract treatment of language equivalence for automata. We use the term *global-view* approaches for the coalgebraic methods via generalised determinisation.

In this paper, we propose a theory of trace semantics for nondeterministic probabilistic systems that unifies the local and the global view. We start by taking the global-view approach founded on algebras and coalgebras and inspired by automata theory, and study determinisation of NPLTS in this framework. Then we find a way to mimic the local-view approach and show that we can recover known trace semantics from the literature. We introduce now the main pieces of our puzzle, and show how everything combines together in the theory of traces for NPLTS.

In order to illustrate our approach, it is convenient to recall nondeterministic automata (NDA) and Rabin probabilistic automata (PA) [31]. Both NDA and PA can be described as maps $\langle o, t\rangle\colon X \to O \times (MX)^A$ where X is a set of states, A is the set of labels, $o\colon X \to O$ is the output function assigning to each state in X an observation, and $t\colon X \to (MX)^A$ is the transition function that assigns to each state x in X and to each letter a of the alphabet A an element of MX that describes the choice of a next state. For NDA, this is a nondeterministic choice; for PA, the choice is governed by a probability distribution. An NDA state observes one of two possible values which qualify the state as accepting or not. A state in a PA observes a real number in $[0, 1]$. Below we depict an example NDA (on the left) and an example PA (on the right) with labels $A = \{a, b\}$ and with outputs denoted by $\downarrow$.



The type of choice, modelled abstractly by a monad M, is often linked to a concrete algebraic theory, the presentation of M. Having such a presentation is a valuable tool, since it provides a finite syntax for describing finite branching. For nondeterministic choice this is the algebraic theory of semilattices (with bottom), for probabilistic choice it is the algebraic theory of convex algebras. Once we have such an algebraic presentation, we have a determinised automaton (as depicted below) and we inductively compute the output value after executing a trace by following the algebraic structure.



Here $x \oplus y$ denotes the nondeterministic choice of x or y, and $x +_p y$ the probabilistic choice where x is chosen with probability p and y with probability $1 - p$.

For example, in the determinised PA we have, since $x \xrightarrow{a} x +_{\frac{1}{2}} y$ and $y \xrightarrow{a} y$:

$$x +_{\frac{1}{2}} y \xrightarrow{a} (x +_{\frac{1}{2}} y) +_{\frac{1}{2}} y = x +_{\frac{1}{4}} y$$

and hence the output of $x +_{\frac{1}{4}} y$ is $o(x) +_{\frac{1}{4}} o(y) = \frac{3}{4}$ giving us the probability of x executing the trace aa. Our computation is enabled by having the right algebraic structure on the set of observations: a semilattice on $\{0, 1\}$ and a convex algebra on $[0, 1]$. The induced semantics is language equivalence and probabilistic language equivalence, respectively.

This is the approach of trace semantics via a determinisation [30], founded in the abstract understanding of automata as coalgebras and computational effects as monads.

We develop a theory of traces for NPLTS using such approach. For this purpose we take the monad for nondeterminism and probability [17] with origins in [14], [18], [19], [20], [21], [32], namely, the monad C of nonempty convex subsets of distributions, and provide all necessary and convenient infrastructure for generalised determinisation. The necessary part is having an algebra of observations, the convenient part is giving an algebraic presentation in terms of *convex semilattices*. These are algebras that are at the same time a semilattice and a convex algebra, with a distributivity axiom distributing probability over nondeterminism. Having the presentation we can write, for example

$$x \xrightarrow{a} x_1 \oplus (x_3 +_{\frac{1}{2}} x_2)$$

for the NPLTS from Figure 1.

The presentation for C is somewhat known, although not explicitly proven, in the community — proving it and putting it to good use is part of our contribution which, in our opinion, drastically clarifies and simplifies the trace theory of systems with nondeterminism and probability.

Remarkably, necessity and convenience go hand in hand on this journey. Having the presentation enables us to clearly identify what are the *interesting* algebras necessary for describing trace and testing semantics (with tests being finite traces). We identify three different algebraic theories: the *theory of pointed convex semilattices*, the *theory of convex semilattices with bottom*, and the *theory of convex semilattices with top*. These theories give rise to three interesting semantics by taking as algebras of observations those freely generated by a singleton set. We prove their concrete characterisations: the free convex semilattice with bottom is carried by $[0, 1]$ with $\max$ as semilattice operation and standard convex algebra operations; the free convex semilattice with top is carried by $[0, 1]$ with $\min$ as semilattice operation; and the pointed convex semilattice freely generated by 1 is carried by the set of closed intervals in $[0, 1]$ where the semilattice operation combines two intervals by taking their minimum and their maximum, and the convex operations are given by Minkowski sum.

We call the resulting three semantics *may trace*, *must trace* and *may-must trace semantics* since there is a close correspondence with *probabilistic testing semantics* [33], [34], [35], [36] when tests are taken to be just the finite traces in A^*. Indeed, the may trace semantics gives the greatest probability with which a state passes a given test;

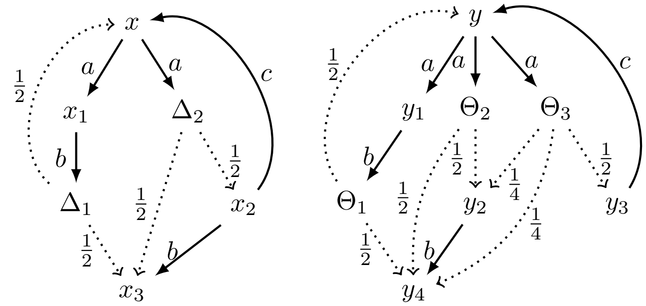


Figure 1. NPLTS

the must trace semantics gives the smallest probability with which a state passes a given test, and the may-must trace semantics gives the closed interval ranging from the smallest to the greatest.

From the abstract theory, we additionally get that:

1) The induced equivalence can be proved coinductively by means of proof-techniques known as *bisimulations up-to* [37]. More precisely, it holds that up-to $\oplus$ and up-to $+_p$ are compatible [38] techniques.
2) The equivalence is implied by the standard branching-time equivalences for NPLTS, namely bisimilarity and convex bisimilarity [7], [39].
3) The equivalence is *backward compatible* w.r.t. trace equivalence for LTS and for reactive probabilistic systems (RPLTS): When regarding an LTS and RPLTS as a nondeterministic probabilistic system, standard trace equivalence coincides with our may trace equivalence and with our three semantics, respectively.

Last but certainly not least, we show that the global view coincides with the local one, namely that our three semantics can be elegantly characterised in terms of resolutions. The may-trace semantics assigns to each trace the greatest probability with which the trace can be performed, with respect to any resolution of the system; the must-trace semantics assigns the smallest one. It is important to remark here that our resolutions differ from those previously proposed in the literature in the fact that they are reactive rather than fully probabilistic. We observe that however this difference does not affect the greatest probability, and we can therefore show that the may-trace coincides with the randomized $\sqcup$-trace equivalence in [25], [26], [40].

Synopsis. We recall monads and algebraic theories in Section 2. We provide a presentation for the monad C in Section 3 (Theorem 4) and combine it with termination in Section 4. We then recall, in Section 5, the generalised determinisation and show an additional useful result (Theorem 16). All these pieces are put together in Section 6, where we introduce our three semantics and discuss their properties. The correspondence of the global view with the local one is illustrated in Section 7 (Theorem 23). The effectiveness of the bisimulation up-to techniques is shown in Appendix A (Example 30). All proofs are in the appendix.

2. Monads and Algebraic Theories

In this paper, on the algebraic side, we deal with Eilenberg-Moore algebras of a monad on the category $\mathbf{Sets}$ of sets and functions, for which we also give presentations in terms of operations and equations, i.e., algebraic theories.

2.1. Monads

A monad on $\mathbf{Sets}$ is a functor $M\colon \mathbf{Sets} \to \mathbf{Sets}$ together with two natural transformations: a unit $\eta\colon \mathrm{Id} \Rightarrow M$ and multiplication $\mu\colon M^2 \Rightarrow M$ that satisfy the laws $\mu \circ \eta M = \mu \circ M\eta = id$ and $\mu \circ M\mu = \mu \circ \mu M$.

We next introduce several monads on $\mathbf{Sets}$, relevant to this paper. Each monad can be seen as giving side-effects.

Nondeterminism. The finite powerset monad $\mathcal{P}$ maps a set X to its finite powerset $\mathcal{P}X = \{U \mid U \subseteq X,\ U \text{ is finite}\}$ and a function $f\colon X \to Y$ to $\mathcal{P}f\colon \mathcal{P}X \to \mathcal{P}Y$, $\mathcal{P}f(U) = \{f(u) \mid u \in U\}$. The unit η of $\mathcal{P}$ is given by singleton, i.e., $\eta(x) = \{x\}$ and the multiplication μ is given by union, i.e., $\mu(S) = \bigcup_{U \in S} U$ for $S \in \mathcal{P}\mathcal{P}X$. Of particular interest to us in this paper is the submonad $\mathcal{P}_{ne}$ of non-empty finite subsets, that acts on functions just like the (finite) powerset monad, and has the same unit and multiplication. We rarely mention the unrestricted (not necessarily finite) powerset monad, which we denote by $\mathcal{P}_u$. We sometimes write $\overline{f}$ for $\mathcal{P}_u f$ in this paper.

Probability. The finitely supported probability distribution monad $\mathcal{D}$ is defined, for a set X and a function $f\colon X \to Y$, as

$$\mathcal{D}X = \{\varphi\colon X \to [0,1] \mid \sum_{x \in X} \varphi(x) = 1,\ \mathrm{supp}(\varphi) \text{ is finite}\}$$

$$\mathcal{D}f(\varphi)(y) = \sum_{x \in f^{-1}(y)} \varphi(x).$$

The support set of a distribution $\varphi \in \mathcal{D}X$ is $\mathrm{supp}(\varphi) = \{x \in X \mid \varphi(x) \neq 0\}$. The unit of $\mathcal{D}$ is given by a Dirac distribution $\eta(x) = \delta_x = (x \mapsto 1)$ for $x \in X$ and the multiplication by $\mu(\Phi)(x) = \sum_{\varphi \in \mathrm{supp}(\Phi)} \Phi(\varphi) \cdot \varphi(x)$ for $\Phi \in \mathcal{D}\mathcal{D}X$. We sometimes write $\sum_{i \in I} p_i x_i$ for a distribution φ with $\mathrm{supp}(\varphi) = \{x_i \mid i \in I\}$ and $\varphi(x_i) = p_i$.

Termination. The termination monad, also called lift and denoted by $\cdot + 1$ maps a set X to the set $X + 1$, where $+$ denotes the coproduct in $\mathbf{Sets}$, which amounts to disjoint union, and $1 = \{\star\}$. For a coproduct $A + B$ we write $in_l\colon A \to A + B$ and $in_r\colon B \to A + B$ for the left and right coproduct injections, respectively. This monad maps a function $f\colon X \to Y$ to the function $f + 1\colon X + 1 \to Y + 1$ defined, as expected, by $(f + 1)(in_l(x)) = in_l(f(x))$ for $x \in X$ and $(f + 1)(in_r(\star)) = in_r(\star)$. The unit of the termination monad is given by the left injection, $\eta\colon X \to X + 1$ with $\eta(x) = in_l(x)$ and the multiplication by $\mu(in_l \circ in_l(x)) = in_l(x)$ for $x \in X$, $\mu(in_l \circ in_r(\star)) = in_r(\star)$, and $\mu(in_r(\star)) = in_r(\star)$. If clear from the context, we may omit explicit mentioning of the injections, and write for example $(f + 1)(x) = x$ for $x \in X$ and $(f + 1)(\star) = \star$.

2.2. Monad Maps, Quotients and Submonads

A monad map from a monad M to a monad $\hat{M}$ is a natural transformation $\sigma\colon M \Rightarrow \hat{M}$ that makes the following diagrams commute, with η, μ and $\hat{\eta}, \hat{\mu}$ denoting the unit and multiplication of M and $\hat{M}$, respectively, and $\sigma\sigma = \sigma \circ M\sigma = \hat{M}\sigma \circ \sigma_M$.

$$\begin{tikzcd} X \arrow[r, "\eta"] \arrow[dr, "\hat{\eta}"'] & MX \arrow[d, "\sigma"] \\ & \hat{M}X \end{tikzcd} \qquad \begin{tikzcd} MMX \arrow[r, "\sigma\sigma"] \arrow[d, "\mu"'] & \hat{M}\hat{M}X \arrow[d, "\hat{\mu}"] \\ MX \arrow[r, "\sigma"'] & \hat{M}X \end{tikzcd}$$

If $\sigma\colon MX \to \hat{M}X$ is an epi monad map, then $\hat{M}$ is a *quotient* of M. If it is a mono, then M is a *submonad* of $\hat{M}$. If it is an iso, the two monads are isomorphic.

2.3. Distributive Laws

Let (M, η, μ) and $(\hat{M}, \hat{\eta}, \hat{\mu})$ be two monads. A *monad distributive law* of M over $\hat{M}$ is a natural transformation $\lambda\colon M\hat{M} \Rightarrow \hat{M}M$ that commutes appropriately with the units and the multiplications of the monads, see Appendix C.

Given a monad distributive law $\lambda\colon M\hat{M} \Rightarrow \hat{M}M$, we get a composite monad $\bar{M} = \hat{M}M$ with unit $\bar{\eta} = \hat{\eta}\eta$ and multiplication $\bar{\mu} = \hat{\mu}\mu \circ \hat{M}\lambda M$.

For any monad M on $\mathbf{Sets}$, there exists a distributive law $\iota\colon M + 1 \Rightarrow M(\cdot + 1)$ defined as

$$\iota_X = \Big(MX + 1 \xrightarrow{[Mi_l,\, \eta_{X+1} \circ i_r]} M(X + 1) \Big). \quad (1)$$

As a consequence, $M(\cdot + 1)$ is a monad. Moreover, we get the following useful property.

Lemma 1. *Whenever $\sigma\colon M \Rightarrow \hat{M}$ is a monad map, also $\sigma(\cdot + 1)\colon M(\cdot + 1) \Rightarrow \hat{M}(\cdot + 1)$ is a monad map. Injectivity of σ implies injectivity of $\sigma(\cdot + 1)$.* □

2.4. Algebraic Theories

With a monad M one associates the Eilenberg-Moore category $\mathcal{EM}(M)$ of M-algebras. Objects of $\mathcal{EM}(M)$ are pairs $\mathbb{A} = (A, a)$ of a set $A \in \mathbf{Sets}$ and a map $a\colon MA \to A$, making the first two diagrams below commute.

$$\begin{tikzcd} A \arrow[r, "\eta"] \arrow[dr, equal] & MA \arrow[d, "a"] \\ & A \end{tikzcd} \quad \begin{tikzcd} M^2A \arrow[r, "Ma"] \arrow[d, "\mu"'] & MA \arrow[d, "a"] \\ MA \arrow[r, "a"'] & A \end{tikzcd} \quad \begin{tikzcd} MA \arrow[r, "Mh"] \arrow[d, "a"'] & MB \arrow[d, "b"] \\ A \arrow[r, "h"'] & B \end{tikzcd}$$

A homomorphism from an algebra $\mathbb{A} = (A, a)$ to an algebra $\mathbb{B} = (B, b)$ is a map $h\colon A \to B$ between the underlying sets making the third diagram above commute.

In this paper we care for both categorical algebra, algebras of a monad, and their presentations in terms of algebraic theories and their models. An algebraic theory is a pair (Σ, E) of signature Σ (a set of operation symbols) and a set of equations E (a set of pairs of terms). A (Σ, E)-algebra, or a model of the algebraic theory (Σ, E) is an algebra $\mathbb{A} = (A, \Sigma_A)$ with carrier set A and a set of operations Σ_A, one for each operation symbol in Σ, that satisfies the equations in E. A homomorphism from a (Σ, E)-algebra $\mathbb{A} = (A, \Sigma_A)$ to a (Σ, E)-algebra $\mathbb{B} = (B, \Sigma_B)$ is a function $h\colon A \to B$ that commutes with the operations, i.e., $h \circ f_A = f_B \circ h^n$ for all n-ary $f \in \Sigma$, and f_A, f_B its interpretations in $\mathbb{A}, \mathbb{B}$, respectively. (Σ, E)-algebras together with their homomorphisms form a category and a variety.

Definition 2. *A* presentation *of a monad M is an algebraic theory, (Σ, E) such that the category (variety) of (Σ, E)-algebras is isomorphic to* $\mathrm{EM}(M)$.

Given a presentation (Σ, E) of a monad M, M is isomorphic to the monad $M_{\Sigma,E}$ of Σ-terms modulo E-equations, i.e., there is an isomorphism monad map between them. Given a signature Σ, the free monad $T_\Sigma = T_{\Sigma,\emptyset}$ of terms over Σ maps a set X to the set of all Σ-terms with variables in X, and $f\colon X \to Y$ to the function that maps a term over X to a term over Y obtained by substitution according to f. The unit maps a variable X to itself, and the multiplication is term composition. We have that $T_{\Sigma,E}$ is a quotient of T_Σ. Moreover, for two sets of equations $E_1 \subseteq E_2$ we have that the monad T_{Σ,E_2} is a quotient of T_{Σ,E_1}. In the sequel we present several algebraic theories that give presentations to the monads of interest.

Presenting the monad $\mathcal{P}_{ne}$. Let Σ_N be the signature consisting of a binary operation $\oplus$. Let E_N be the following set of axioms.

$$\begin{array}{ccc} (x \oplus y) \oplus z & \stackrel{(A)}{=} & x \oplus (y \oplus z) \\ x \oplus y & \stackrel{(C)}{=} & y \oplus x \\ x \oplus x & \stackrel{(I)}{=} & x \end{array}$$

The algebraic theory (Σ_N, E_N) of *semilattices* provides a presentation for the monad $\mathcal{P}_{ne}$. We refer to this theory as the theory of nondeterminism. To avoid confusion later, it is convenient to fix here the interpretation of $\oplus$ as a join (rather than a meet) and, thus, to think of the induced order as $x \sqsubseteq y$ iff $x + y = y$.

Presenting the monad $\mathcal{D}$. Let Σ_P be the signature consisting of binary operations $+_p$ for all $p \in (0, 1)$. Let E_P be the following set of axioms.

$$\begin{array}{ccc} (x +_q y) +_p z & \stackrel{(A_p)}{=} & x +_{pq} (y +_{\frac{p(1-q)}{1-pq}} z) \\ x +_p y & \stackrel{(C_p)}{=} & y +_{1-p} x \\ x +_p x & \stackrel{(I_p)}{=} & x \end{array}$$

Here, (A_p), (C_p), and (I_p) are the axioms of parametric associativity, commutativity, and idempotence. The algebraic theory (Σ_P, E_P) of *convex algebras*, see [41], [42], [43], [44], [45], provides a presentation for the monad $\mathcal{D}$.

Another presentation of convex algebras is given by the algebraic theory with infinitely many operations denoting arbitrary (and not only binary) convex combinations — see Appendix C for more details. This allows us to interchangeably use binary convex combinations or arbitrary convex combinations whenever more convenient. Moreover, we can write binary convex combinations $+_p$ for $p \in [0, 1]$ and not just $p \in (0, 1)$. We refer to the theory of convex algebras as the algebraic theory for probability.

Presenting $\cdot + 1$. The algebraic theory (Σ_T, E_T) for the termination monad consists of a single constant (nullary operation symbol) $\Sigma_T = \{\star\}$ and no equations $E_T = \emptyset$. This is called the theory of *pointed sets*.

Combining Algebraic Theories. Algebraic theories can be combined in a number of general ways: by taking their coproduct, their tensor, or by means of distributive laws (see e.g. [46]). Unfortunately, these abstract constructions do not lead to a presentation for the monad we are interested in. We will thus devote the next section to show a "hand-made" presentation for this monad.

We conclude this section with a well known fact that can be easily proved, for instance by taking the distributive law in (1): given a presentation (Σ, E) for a monad M, the monad $M(\cdot + 1)$ is presented by the theory (Σ', E) where Σ' is Σ together with an extra constant $\star$. For instance, the subdistributions monad $\mathcal{D}(\cdot + 1)$ is presented by the theory $(\Sigma_P \cup \Sigma_T, E_P)$ of *pointed convex algebras*, also known as *positive convex algebras*. The theory $(\Sigma_N \cup \Sigma_T, E_N)$ of *pointed semilattices* provides instead a presentation for the monad $\mathcal{P}_{ne}(\cdot + 1)$. It is interesting to observe that the powerset monad $\mathcal{P}$ is presented by adding to $(\Sigma_N \cup \Sigma_T, E_N)$ the equation

$$x \oplus \star \stackrel{(B)}{=} x$$

leading to the theory of *semilattices with bottom*. The theory of *semilattices with top* can be obtained by adding instead the following equation:

$$x \oplus \star \stackrel{(T)}{=} \star.$$

Similar axioms can be added to the theory of pointed convex algebras $(\Sigma_P \cup \Sigma_T, E_P)$. The axiom

$$x +_p \star \stackrel{(B_p)}{=} x$$

makes the probabilistic structure collapse, see Figure 6 in Appendix C for the details. On the other hand, the axiom

$$x +_p \star \stackrel{(T_p)}{=} \star$$

quotients the monad $\mathcal{D}(\cdot + 1)$ into $\mathcal{D} + 1$: intuitively, each term of this theory is either a sum of only variables (a distribution) or an extra element ($\star$). This axiom describes the unique functorial way of adding termination to a convex algebra, the so-called black-hole behaviour of $\star$, cf. [47].

3. Algebraic Theory for Nondeterminism and Probability

In this section we recall the definition of the monad C for probability and nondeterminism, give its presentation via *convex semilattices*, and present examples of C-algebras.

3.1. The monad C of convex subsets of distributions

The monad C origins in the field of domain theory [19], [20], [21], and in the work of Varacca and Winskel [14], [18], [32]. Jacobs [17] gives a detailed study of (a generalisation of) this monad.

For a set X, CX is the set of non-empty, finitely-generated convex subsets of distributions on X, i.e.,

$$CX = \{S \subseteq \mathcal{D}X \mid S \neq \emptyset, \mathrm{conv}(S) = S, \\ S \text{ is finitely generated}\}.$$

Recall that, for a subset S of a convex algebra, $\mathrm{conv}(S)$ is the convex closure of S, i.e., the smallest convex set that contains S, i.e.,

$$\mathrm{conv}(S) = \{\sum p_i x_i \mid p_i \in [0,1], \sum p_i = 1, x_i \in S\}.$$

We say that a convex set S is generated by its subset B if $S = \mathrm{conv}(B)$. In such a case we also say that B is a basis for S. A convex set S is finitely generated if it has a finite basis.

For a function $f\colon X \to Y$, $Cf\colon CX \to CY$ is given by

$$Cf(S) = \{\mathcal{D}f(d) \mid d \in S\} = \overline{\mathcal{D}f}(S).$$

The unit of C is $\eta\colon X \to CX$ given by $\eta(x) = \{\delta_x\}$.

The multiplication of C, $\mu\colon CCX \to CX$ can be expressed in concrete terms as follows [17]. Given $S \in CCX$,

$$\mu(S) = \bigcup_{\Phi \in S} \{ \sum_{U \in \mathrm{supp}\,\Phi} \Phi(U) \cdot d \mid d \in U\}.$$

3.2. The presentation of C

We now introduce the algebraic theory (Σ_{NP}, E_{NP}) of *convex semilattices*, that gives us the presentation of C and thus provides an algebraic theory for nondeterminism and probability.

A convex semilattice $\mathbb{A}$ is an algebra $\mathbb{A} = (A, \oplus, +_p)$ with a binary operation $\oplus$ and for each $p \in (0,1)$ a binary operation $+_p$ satisfying the axioms $(A), (C), (I)$ of a semilattice, the axioms $(A_p), (C_p), (I_p)$ for a convex algebra, and the following distributivity axiom:

$$(x \oplus y) +_p z \overset{(D)}{=} (x +_p z) \oplus (y +_p z)$$

Hence, (Σ_{NP}, E_{NP}) for $\Sigma_{NP} = \Sigma_N \cup \Sigma_P$ and $E_{NP} = E_N \cup E_P \cup \{(D)\}$.

In every convex semilattice there also holds a convexity law, of which we directly present the generalized version in the following lemma.

Lemma 3. *Let $\mathbb{A} = (A, \oplus, +_p)$ be a convex semilattice. Then for all $n \in \mathbb{N}$, all $a_1, \ldots a_n \in A$ and all $p_1, \ldots, p_n \in [0,1]$ with $\sum_{i=1}^n p_i = 1$ we have*

$$a_1 \oplus \ldots \oplus a_n \oplus \sum_{i=1}^{n} p_i a_i \overset{(C)}{=} a_1 \oplus \ldots \oplus a_n. \quad \square$$

For $p \in [0,1]$ we set $\overline{p} = 1 - p$. Let X be an arbitrary set. We define Σ_{NP}-operations on CX by

$$S_1 \oplus S_2 = \mathrm{conv}(S_1 \cup S_2)$$

and for $p \in (0,1)$

$$S_1 +_p S_2 = \{\varphi \mid \varphi = p\varphi_1 + \overline{p}\varphi_2 \text{ for some } \varphi_1 \in S_1, \varphi_2 \in S_2\}$$

where $p\varphi_1 + \overline{p}\varphi_2 = \varphi_1 +_p \varphi_2$ is the binary convex combination of φ_1 and φ_2 in $\mathcal{D}X$, defined point-wise. Note that $S_1 +_p S_2$ is the Minkowski sum of two convex sets. If convenient, we may sometimes also write, as usual, $pS_1 + \overline{p}S_2$ for the Minkowski sum $S_1 +_p S_2$.

To prove the presentation theorem, we identify a generic proof method that we only present in the appendix for lack of space. We encourage the reader to read the appendix, also for many other useful properties that deepen the understanding of convex semilattices.

Theorem 4. *The theory for nondeterminism and probability (Σ_{NP}, E_{NP}), i.e., the theory of convex semilattices, is a presentation for the monad C.* $\square$

Remark 5. *Theorem 4 is to some extent known*[1] *but we could not find a proof of it in the literature. In [14], [18] a monad for probability and nondeterminism is given starting from a similar algebraic theory (with somewhat different basic algebraic structure). There is also another possible way of combining probability with nondeterminism, by distributing $\oplus$ over $+_p$ (see e.g. [15], [48]).*

Remark 6. *Having the presentation enables us to identify and interchangeably use convex subsets of distributions and terms in Σ_{NP} modulo equations in E_{NP}. This is particularly useful in examples and our further developments. Note that in the syntactic view $\eta(x)$ is identified with the term x.*

The presentation is a valuable tool in many situations where reasoning with algebraic theories is more convenient than reasoning with monads. For instance, it is much easier to check whether a certain algebra is a (Σ_{NP}, E_{NP})-model, than to check that it is an algebra for the monad C. We illustrate this with three (Σ_{NP}, E_{NP}) models that play a key role in our further results and exposition.

The max convex semilattice. $\mathbb{M}\mathrm{ax} = ([0,1], \max, +_\mathrm{p})$ is a (Σ_{NP}, E_{NP})-algebra when taking $\oplus$ to be $\max\colon [0,1] \times [0,1] \to [0,1]$ and $+_p$ the standard convex combination $+_p\colon [0,1] \times [0,1] \to [0,1]$ with $x +_p y = p \cdot x + \overline{p} \cdot y$ for $x, y \in [0,1]$. To check that this is a (Σ_{NP}, E_{NP}) model, it is enough to prove that $\max$ satisfies the axioms in E_N, that $+_p$ satisfies the axioms in E_P, and that they satisfy the axiom (D), namely that $\max(x,y) +_p z = \max(x +_p z, y +_p z)$.

The min convex semilattice. $\mathbb{M}\mathrm{in} = ([0,1], \min, +_\mathrm{p})$ is obtained similarly by taking $\oplus$ to be $\min\colon [0,1] \times [0,1] \to [0,1]$ rather than $\max$, and gives another example of a (Σ_{NP}, E_{NP})-algebra. It is indeed very simple to check that $([0,1], \min)$ forms a semilattice and that the distributivity law holds.

1. Personal communication with Gordon Plotkin.

The min-max interval convex semilattice. We consider the algebraic structure $\mathbb{M}_{\mathcal{I}} = (\mathcal{I}, \text{min-max}, +_p^{\mathcal{I}})$ for $\mathcal{I}$ the set of intervals on $[0,1]$, i.e.,

$$\mathcal{I} = \{[x,y] \mid x, y \in [0,1] \text{ and } x \leq y\}.$$

For $[x_1,y_1],[x_2,y_2] \in \mathcal{I}$, we define $\text{min-max}\colon \mathcal{I} \times \mathcal{I} \to \mathcal{I}$ as

$$\text{min-max}([x_1,y_1],[x_2,y_2]) = [\min(x_1,x_2), \max(y_1,y_2)]$$

and $+_p^{\mathcal{I}}\colon \mathcal{I} \times \mathcal{I} \to \mathcal{I}$ by

$$[x_1,y_1] +_p^{\mathcal{I}} [x_2,y_2] = [x_1 +_p x_2,\ y_1 +_p y_2].$$

The fact that this is a model for (Σ_{NP}, E_{NP}) follows easily from the fact that $\mathbb{M}\mathrm{ax}$ and $\mathbb{M}\mathrm{in}$ are models for (Σ_{NP}, E_{NP}).

Remark 7. *The fact that* $\mathbb{M}\mathrm{ax}$ *and* $\mathbb{M}\mathrm{in}$ *are* C*-algebras on* $[0,1]$ *was already proven in [49], without an algebraic presentation. Having the algebraic presentation significantly simplifies the proofs.*

4. Adding termination

So far, we have provided a presentation for the monad C which combines probability and nondeterminism. In order to properly model NPLTS, we need a last ingredient: termination. As discussed in Section 2, termination is given by the monad $\cdot + 1$ which can always be safely combined with any monad. Following the discussion at the end of Section 2, the theory $\mathcal{PCS} = (\Sigma_{NP} \cup \Sigma_T, E_{NP})$ presents the monad $C(\cdot + 1)$ which is the monad of finitely generated non empty convex sets of *sub*distributions.

We call this theory $\mathcal{PCS}$ since algebras for this theory are *pointed convex semilattices*, namely convex semilattices with a pointed element denoted by $\star$. A noteworthy example is $\mathbb{M}_{\mathcal{I},[0,0]} = (\mathcal{I}, \text{min-max}, +_p^{\mathcal{I}}, [0,0])$ where $\mathbb{M}_{\mathcal{I}} = (\mathcal{I}, \text{min-max}, +_p^{\mathcal{I}})$ is the convex semilattice of intervals from Section 3 and $[0,0]$ is the pointed element. Moreover, this is not just any pointed convex semilattice:

Proposition 8. $\mathbb{M}_{\mathcal{I},[0,0]} = (\mathcal{I}, \text{min-max}, +_p^{\mathcal{I}}, [0,0])$ *is the free pointed convex semilattice generated by a singleton set.*

Like for the monad $\mathcal{P}_{ne}$, there exist more than one interesting way of combining C with $\cdot + 1$. Rather than pointed convex semilattices, one can consider *convex semilattices with bottom*, namely algebras for the theory $\mathcal{CSB} = (\Sigma_{NP} \cup \Sigma_T, E_{NP} \cup \{(B)\})$ obtained by adding (B) to $\mathcal{PCS}$. Otherwise, one can add the axiom (T) and obtain the theory $\mathcal{CST} = (\Sigma_{NP} \cup \Sigma_T, E_{NP} \cup \{(T)\})$ of *convex semilattices with top*. We denote by $T_{\mathcal{CSB}}$ and $T_{\mathcal{CST}}$ the corresponding monads.

As we will illustrate in Section 5, particularly relevant for defining trace semantics is the free algebra $\mu\colon MM\{\bullet\} \to M\{\bullet\}$ generated by a singleton $\{\bullet\}$. In the next two propositions we identify these algebras for the monads $T_{\mathcal{CSB}}$ and $T_{\mathcal{CST}}$ in concrete terms.

Proposition 9. $\mathbb{M}\mathrm{ax}_B = ([0,1], \max, +_p, 0)$ *is the free convex semilattice with bottom generated by* $1 = \{\bullet\}$. □

Proposition 10. $\mathbb{M}\mathrm{in}_T = ([0,1], \min, +_p, 0)$ *is the free convex semilattice with top generated by* $1 = \{\bullet\}$. □

At this point the reader may wonder what happens when one considers the axioms (B_p) and (T_p) in place of (B) and (T). We have already shown at the end of Section 2.4, that the axiom (B_p) makes the probabilistic structure collapse. When focussing on the free algebra generated by $\{\bullet\}$, also quotienting by (T_p) is not really interesting: one can show by induction on the terms in $T_{\Sigma_{NP} \cup \Sigma_T}(\{\bullet\})$ that every term is equal via $E_{NP} \cup \{(T_p)\}$ to either $\bullet$ or $\star$ or $\bullet \oplus \star$.

So, we have found three interesting ways of combining termination with probability and nondeterminism. Table 1 summarises these theories, their monads, and their algebras.

5. Coalgebras and Determinisation

In this section, we briefly introduce coalgebra and the generalized determinization [30] construction, as well as trace semantics by determinization. We present some simple properties and a new important result concerning the semantics.

5.1. Coalgebra

The theory of coalgebra provides an abstract framework for state-based transition systems and automata. Let $\mathbf{Sets}$ be the category of sets and functions. A *coalgebra* in $\mathbf{Sets}$ is a pair (S,c) of a state space S and a function $c\colon S \to FS$ where $F\colon \mathbf{Sets} \to \mathbf{Sets}$ is a functor that specifies the type of transitions. Sometimes we say the coalgebra $c\colon S \to FS$, meaning the coalgebra (S,c).

A *coalgebra homomorphism* from a coalgebra (S,c) to a coalgebra (T,d) is a function $h\colon S \to T$ that satisfies $d \circ h = Fh \circ c$. Coalgebras of a functor F and their coalgebra homomorphisms form a category, denoted by $\mathrm{Coalg}\,(F)$.

The final object in $\mathrm{Coalg}\,(F)$, when it exists, is the *final* F*-coalgebra*. We write $\zeta\colon Z \xrightarrow{\cong} FZ$ for the final F-coalgebra. For every coalgebra $c\colon S \to FS$, there is a unique homomorphism $[\![\cdot]\!]_c$ to the final one, the *final coalgebra map*, making the diagram below commute:

$$\begin{tikzcd} FS \arrow[r, dashed, "F[\![\cdot]\!]_c"] & FZ \\ S \arrow[u, "c"] \arrow[r, dashed, "\exists!\,[\![\cdot]\!]_c"] & Z \arrow[u, "\cong"', "\zeta"] \end{tikzcd}$$

The *final coalgebra semantics* $\sim$ is the kernel of the final coalgebra map, i.e., two states s and t are equivalent in the final coalgebra semantics iff $[\![s]\!]_c = [\![t]\!]_c$.

Even without a final coalgebra, coalgebras over a concrete category are equipped with a generic behavioural equivalence. Let (S,c) be an F-coalgebra on $\mathbf{Sets}$. An equivalence relation $R \subseteq S \times S$ is a kernel bisimulation (synonymously, a cocongruence) [50], [51], [52] if it is the kernel of a homomorphism, i.e., $R = \ker h = \{(s,t) \in S \times S \mid h(s) = h(t)\}$ for some coalgebra homomorphism $h\colon (S,c) \to (T,d)$ to some F-coalgebra (T,d). Two states s,t of a coalgebra are behaviourally equivalent (notation:

Theory (Σ, E)	Monad M	free algebra $\mu_1\colon MM1 \to M1$
$\mathcal{PCS} = (\Sigma_{NP} \cup \Sigma_T, E_{NP})$	$C(\cdot+1) = T_{\mathcal{PCS}}$	$\mathbb{M}_{\mathcal{I},[0,0]} = (\mathcal{I}, \text{min-max}, +^{\mathcal{I}}_p, [0,0])$
$\mathcal{CSB} = (\Sigma_{NP} \cup \Sigma_T, E_{NP} \cup \{(B)\})$	$T_{\mathcal{CSB}}$	$\mathbb{M}\mathrm{ax}_B = ([0,1], \max, +_p, 0)$
$\mathcal{CST} = (\Sigma_{NP} \cup \Sigma_T, E_{NP} \cup \{(T)\})$	$T_{\mathcal{CST}}$	$\mathbb{M}\mathrm{in}_T = ([0,1], \min, +_p, 0)$

TABLE 1. THE THEORIES OF POINTED CONVEX SEMILATTICES, WITH BOTTOM, AND WITH TOP.

$s \approx t$) iff there is a kernel bisimulation R with $(s,t) \in R$. If a final coalgebra exists, then the behavioural equivalence and the final coalgebra semantics coincide, i.e., $\approx\, = \,\sim$.

The following are well-known examples of F-coalgebras on **Sets**:

1) Labelled transition systems, LTS, are coalgebras for the functor $F = \mathcal{P}^A$. Behavioural equivalence coincides with strong bisimilarity.
2) Nondeterministic automata, NA, are coalgebras for $F = 2 \times \mathcal{P}^A$ where $2 = \{0,1\}$ is needed to differentiate whether a state is accepting or not.
3) Deterministic automata, DA, are coalgebras for $F = 2 \times (\cdot)^A$. The final coalgebra is carried by the set of all languages 2^{A^*}.
4) Moore automata, MA, are a slight generalisation of deterministic automata with observations O: they are coalgebras for $F = O \times (\cdot)^A$. The final coalgebra is carried by the set of all O-valued languages O^{A^*}.

Systems and Automata with M-effects. In general, for a monad M, we call an M^A-coalgebra a *system with M-effects*, and we call an $O \times M^A$-coalgebra an *automaton with M-effects and observations in O*. We write $c = \langle o, t\rangle$ for an automaton with M-effects and observations in O, where $o\colon X \to O$ is the observation map assigning observations to states, and $t\colon X \to (MX)^A$ is the transition structure.

For instance, an LTS is a system with $\mathcal{P}$-effects, and a nondeterministic automaton is an automaton with $\mathcal{P}$-effects and observations in 2. We now introduce the systems and automata that we focus on in this paper.

Nondeterministic probabilistic labelled transition systems, NPLTS,. also known as simple Segala systems, are coalgebras for the functor $F = (\mathcal{PD})^A$. Behavioural equivalence coincides with strong probabilistic bisimilarity [53], [54]. Special cases of NPLTS are LTS, when all distributions are Dirac distributions, and reactive probabilistic labelled transition systems (RPLTS), when all subsets are at most singletons. An RPLTS is a coalgebra of the functor $(\mathcal{D}+1)^A$.

Convex NPLTS. are coalgebras for $(C+1)^A$. Behavioural equivalence coincides with convex probabilistic bisimilarity [16]. The move from NPLTS to convex NPLTS is given by a natural transformation $\mathrm{conv}\colon \mathcal{PD} \Rightarrow C+1$ with $\mathrm{conv}(X)$ the convex hull for $X \subseteq \mathcal{D}S$, $X \neq \emptyset$, and $\mathrm{conv}(\emptyset) = \star$. Therefore, $\mathrm{conv}^A\colon (\mathcal{PD})^A \Rightarrow (C+1)^A$ defined pointwise is natural as well. As a consequence [27], [53], we get a translation functor from NPLTS to convex NPLTS, and hence bisimilarity implies convex bisimilarity for NPLTS.

Nondeterministic Probabilistic automata, NPA,. with observations in O are (for us in this paper) coalgebras for $F = O \times (C(\cdot+1))^A$. We explain in Section 5.3 below how to move from (convex) NPLTS to NPA, which involves two steps: (1) Adding observations and (2) Dealing with termination.

We write $x \xrightarrow{a} m$ for $t(x)(a) = m$ with $a \in A, x \in X, m \in MX$ in a system or automaton with M-effects. For an LTS $t\colon X \to (\mathcal{P}X)^A$ we also write, as usual, $x \xrightarrow{a} y$ for $y \in t(x)(a)$ and $x \not\xrightarrow{a}$ if $t(x)(a) = \emptyset$; for an RPLTS $t\colon X \to (\mathcal{D}X+1)^A$, we may also write $x \xrightarrow{a}_p y$ for $t(x)(a)(y) = p$ and again $x \not\xrightarrow{a}$ if $t(x)(a) = \star$. Note that in all our examples of systems and automata there is an implicit finite branching property ensured by the use of $\mathcal{P}$, $\mathcal{D}$ and C involving only finite subsets, finitely supported distributions, and finitely generated convex sets.

5.2. Generalised Determinisation

The construction of generalised determinisation was originally discovered in [30]. It enables us to obtain trace semantics for coalgebras of type $c\colon X \to FMX$ where F is a functor and M a monad. The result is a determinised coalgebra $c^{\#}\colon MX \to FMX$ and the semantics is derived from behavioural equivalence for F-coalgebras.

Let $c\colon X \to FMX$ be a coalgebra and $\lambda\colon MF \Rightarrow FM$ a distributive law of the monad M over the functor F. Such a λ is a natural transformation that commutes appropriately with the unit and the multiplication of M, i.e., $\lambda \circ \eta = F\eta$ and $\lambda \circ \mu = F\lambda \circ \lambda \circ M\lambda$. Then the determinisation is the coalgebra

$$c^\sharp = F\mu \circ \lambda \circ Mc. \tag{2}$$

It is easy to show that $c^\sharp \circ \eta = c$ which justifies the notation $c^\sharp$: The carrier MX carries an M-algebra, the free one generated by X, FMX also does, $F\mu \circ \lambda$ is an M-algebra, and $c^\sharp$ is the unique extension of c to a homomorphism from the free M-algebra (MX, μ) to the M-algebra $(FMX, F\mu \circ \lambda)$.

We obtain behavioral equivalence on MX via the final coalgebra morphism $[\![\cdot]\!]_{c^\sharp}$ into the final coalgebra for F: for m, n in MX, $m \sim n$ iff $[\![m]\!]_{c^\sharp} = [\![n]\!]_{c^\sharp}$. This in turn induces an equivalence on X, via the unit of the monad η: for $x, y \in X$, $x \equiv y$ iff $\eta(x) \sim \eta(y)$. If F is such that a final F-coalgebra does not exist, we can still define $\equiv$ via behavioural equivalence by: for $x, y \in X$, $x \equiv y$ iff $\eta(x) \approx \eta(y)$. This induced semantics $\equiv$ on X is what we call the *trace semantics via determinisation*.

Determinizing automata with M-effects and observations in O. In this paper, we only consider determinisation of automata with M-effects and observations in O. Hence, FM-coalgebras for the Moore-automata functor $F = O \times (\cdot)^A$, where O is a set of observations. The following proposition shows that determinising automata with M-effects and observations in O is always possible when the observations carry an M-algebra [30], [55].

Proposition 11. *For an Eilenberg-Moore algebra $a\colon MO \to O$, for $F = O\times(\cdot)^A$ and any monad M on $\mathbf{Sets}$ there is a canonical distributive law $\lambda_X : MF \Rightarrow FM$ given by*

$$M\big(O\times X^A\big) \xrightarrow{\langle M\pi_1, M\pi_2\rangle} MO\times M(X^A) \xrightarrow{a\times \mathsf{st}} O\times(MX)^A$$

where st is the map $\mathsf{st}\colon M(X^A) \to (MX)^A$ defined by $\mathsf{st}(\varphi) = (a \mapsto M\mathsf{ev}_a(\varphi))$ with $\mathsf{ev}_a\colon X^A \to X$ the evaluation map defined as $\mathsf{ev}_a(\varphi) = \varphi(a)$. $\square$

As a consequence, we can determinise $c = \langle o, t\rangle\colon X \to O \times (MX)^A$ to $c^\sharp = \langle o^\sharp, t^\sharp\rangle$ where $o^\sharp = a \circ Mo$ and $t^\sharp = \mu_X^A \circ \mathsf{st} \circ Mt$. The final coalgebra for the determinization of automata with M-effects and observations in O is carried by the O-weighted languages over alphabet A, i.e., maps $A^* \to O$. Unfolding the inductive definition of the final coalgebra semantics for automata with M-effects and observations in O, see e.g. [28], gives $[\![\eta(x)]\!]_{c^\sharp}(\varepsilon) = o^\sharp(x)$ and $[\![\eta(x)]\!]_{c^\sharp}(aw) = [\![t^\sharp(x)(a)]\!]_{c^\sharp}(w)$.

Knowing that (Σ, E) is a presentation for the monad M, we can write the algebraic structure, and hence the determinisation concretely as follows. For an n-ary operation symbol $f \in \Sigma$ and a (Σ, E)-algebra $\mathbb{A} = (A, \Sigma_A)$ we write f_A for the n-ary operation on A that is the interpretation of f. We have

$$\begin{aligned} &f_{FMX}(\langle o_1, f_1\rangle, \ldots, \langle o_n, f_n\rangle) = \\ &\quad \langle f_O(o_1, \ldots, o_n), (a \mapsto f_{MX}(f_1(a), \ldots, f_n(a)))\rangle. \end{aligned}$$

Therefore, for a coalgebra $c\colon X \to FMX$, we have that $c^\sharp = \langle o^\sharp, t^\sharp\rangle$ is inductively defined on the structure of the Σ-terms by $o^\sharp(x) = o(x)$, $t^\sharp(x) = t(x)$ and

$$\begin{aligned} &o^\sharp(f_{MX}(t_1, \ldots, t_n)) = f_O(o^\sharp(t_1), \ldots, o^\sharp(t_n)) \qquad (3)\\ &t^\sharp(f_{MX}(t_1, \ldots, t_n))(a) = f_{MX}(t^\sharp(t_1)(a), \ldots, t^\sharp(t_n)(a)) \end{aligned}$$

Example 12. *Applying this construction to $F = 2 \times (\cdot)^A$ and $M = \mathcal{P}$, one transforms $c\colon X \to 2 \times (\mathcal{P}X)^A$ into $c^\sharp\colon \mathcal{P}X \to 2 \times (\mathcal{P}X)^A$. The former is a nondeterministic automaton and the latter is a deterministic automaton which has $\mathcal{P}X$ as states space. In [30], see also [55], it is shown that, using the distributive law from Proposition 11, as $2 = \mathcal{P}1$ is the carrier of the free $\mathcal{P}$-algebra, this amounts exactly to the standard determinisation from automata theory and justifies the term* generalised determinisation. *The obtained semantics is language equivalence.*

It is worth to mention that both the determinised coalgebra $c^\sharp\colon MX{\to}FMX$ and the final F-coalgebra are actually *bialgebras* [56], [57], roughly they are both an M-algebra and an F-coalgebra. Moreover, the unique coalgebra morphism $[\![\cdot]\!]_{c^\sharp}\colon MX \to O^{A^*}$ is also an M-algebra homomorphism. The latter entails the first item of the following.

Theorem 13 ([30], [58])**.** *The following properties hold for any coalgebra $c\colon X{\to}FMX$ and its determinisation $c^\sharp\colon MX{\to}FMX$:*

1) *Behavioural equivalence for $(MX, c^\sharp)$ is a congruence w.r.t. the algebraic structure of M.*
2) *Behavioural equivalence for (X, c) implies trace semantics via determinisation.*
3) *Up-to context is a compatible [38] proof technique.*

The second item will be used later in Section 6 to show that convex bisimilarity implies trace equivalence for NPLTS. The third item will be better explained in Section A.

5.3. From Systems to Automata

Dealing with automata, i.e., having observations, is crucial for determinisation. Starting from an LTS $t\colon X \to (\mathcal{P}X)^A$, we can add observations in $2 = \mathcal{P}1$ in the simplest possible way, making every state an accepting state:

$$o = (X \xrightarrow{!} 1 \xrightarrow{\eta_1} \mathcal{P}1 = 2)$$

and determinise the NA $\langle o, t\rangle\colon X \to 2 \times (\mathcal{P}X)^A$. The induced semantics $\equiv^{LTS}$ on the state space X is the standard trace semantics for LTS [59].

This same approach can be applied in the case of any system with M-effects $t\colon X \to (MX)^A$. We can add observations in $O = M1$ by

$$o = (X \xrightarrow{!} 1 \xrightarrow{\eta_1} M1),$$

determinise the automaton $\langle o, t\rangle$ with M-effects using the free algebra on $M1$, and obtain the trace semantics after determinisation $\equiv$.

From NPLTS to NPA. In order to define trace semantics for NPLTS via generalised determinisation, we need to transform them into NPA which are automata with $C(\cdot{+}1)$-effects. We proceed in two steps: we transform an NPLTS to a system with $C(\cdot{+}1)$-effects, and then add observations via the general recipe of this section. Given an NPLTS $t\colon X \to (\mathcal{P}\mathcal{D}X)^A$ we first transform it into the convex NPLTS $X \xrightarrow{t} (\mathcal{P}\mathcal{D}X)^A \xrightarrow{\mathrm{conv}^A} (CX{+}1)^A$ and then employ the distributive law ι from Section 2.3 to obtain

$$\begin{aligned} \bar{t} = \big(X \xrightarrow{t} (\mathcal{P}\mathcal{D}X + 1)^A \xrightarrow{\mathrm{conv}^A} (CX+1)^A \qquad (4)\\ \xrightarrow{\iota^A} (C(X+1))^A\big) \end{aligned}$$

Note that $\bar{t}$ is a system with $C(\cdot+1)$-effects. Moreover, by construction, NPLTS-bisimilarity for t implies convex bisimilarity, and further convex bisimilarity implies behavioural equivalence for the resulting system with $C(\cdot{+}1)$-effects $\bar{t}$. Finally, we add observations as prescribed above:

$$\bar{o} = \big(X \xrightarrow{!} 1 \xrightarrow{\eta_1} C(1+1)\big) \qquad (5)$$

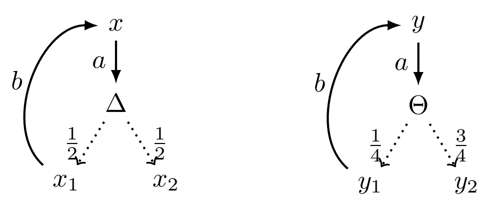
and get the desired automaton with $C(\cdot+1)$-effects and observations in $C(1+1)$. Adding such observations again preserves behavioural equivalence.

Why is termination inside?. We have seen that, when moving from NPLTS to NPA, in particular when moving from convex NPLTS to NPA, we are not just adding an observation. We are also moving, via the ι distributive law, from the functor $C+1$ to the functor $C(\cdot+1)$. The reason why we do this can already be understood in the simpler case of RPLTS, where the monad $\mathcal{D}$ is used instead of C. We have that $\mathcal{D}+1$ is already a monad, and there is a monad map in both directions between $\mathcal{D}(\cdot+1)$ and $\mathcal{D}+1$. So we could take a $\mathcal{D}+1$-algebra and perform a determinisation with respect to $\mathcal{D}+1$. There is however an undesired consequence of doing so, as illustrated by the following example.

Example 14. *Trace semantics for RPLTS is defined in a similar way, see the construction in [60], [61]. An RPLTS $t_*\colon X \to (\mathcal{D}X+1)^A$ can similarly be transformed to a system with $\mathcal{D}(\cdot+1)$-effects using the distributive law ι:*

$$t = X \xrightarrow{t_*} (\mathcal{D}X+1)^A \xrightarrow{\iota^A} (\mathcal{D}(X+1))^A.$$

Consider the following RPLTS.



The states x and y should not be trace equivalent, since x has probability $\frac{1}{2}$ of performing trace ab, and y has probability $\frac{1}{4}$ of performing trace ab. Let us look at what happens, however, if we determinize this system (seen as the $(\mathcal{D}+1)^A$ coalgebra t_*) with respect to the monad $\mathcal{D}+1$. The determinised transition function $t_*^\sharp$ will give us states in $\mathcal{D}X+1$, i.e., states that are either full distributions or the element $\star \in 1$ and we have

$$t_*^\sharp(x)(a) = x_1 +_{\frac{1}{2}} x_2 \qquad t_*^\sharp(y)(a) = y_1 +_{\frac{1}{4}} y_2$$

However,
$$t_*^\sharp(x_1 +_{\frac{1}{2}} x_2)(b) = t_*(x_1)(b) +_{\frac{1}{2}} t_*(x_2)(b) = \star$$
$$t_*^\sharp(y_1 +_{\frac{1}{4}} y_2)(b) = t_*(y_1)(b) +_{\frac{1}{4}} t_*(y_2)(b) = \star$$

Hence, whatever $(\mathcal{D}+1)$-algebra of observation we take, these states in the lifted system will return the same observation, i.e., $o^\sharp(x)(ab) = o^\sharp(y)(ab)$. As a consequence, x and y will be equivalent.

Hence, moving to a monad with termination inside is a fundamental step in our construction, if we want to distinguish processes such as those in the previous example.

However, there are cases in which determinizing with respect to two different monads and algebras leads to the same semantics, as shown in the next example.

Example 15. *As described above, we turn an RPLTS into an automaton with $\mathcal{D}(\cdot+1)$-effects with observations in $[0,1] = \mathcal{D}(1+1)$ equipped with the the free algebra generated by 1. The observation function $o\colon X \to [0,1]$ maps every state $x \in X$ into the element $1 \in [0,1]$. The function $[\![\cdot]\!]_{c^\sharp} \circ \eta\colon X \to [0,1]^{A^*}$ obtained via the generalised determinisation of $c = \langle o, t\rangle$ assigns to each state $x \in X$ and trace $w \in A^*$ the probability of reaching from x any other state via w. We write $\equiv^{RP}$ for the induced trace equivalence.*

Interestingly, (Rabin) probabilistic automata *[31] are defined slightly differently: these are automata with $\mathcal{D}$-effects and observations in $[0,1]$, $\langle o,t\rangle\colon X \to [0,1]\times(\mathcal{D}X)^A$ (see [30]). The set of observations is the same, but transitions go in distributions rather than in subdistributions. The theorem of the next section guarantees that* only *the algebra of observations matters for the resulting semantics, so using $\mathcal{D}$ in place of $\mathcal{D}(\cdot+1)$ does not change the obtained equivalence which in both cases coincides with the probabilistic language equivalence of [31].*

5.4. Invariance of the Semantics

We next state a theorem that guarantees invariance of the trace semantics via determinisation for automata with M-effects and observations in O, under controlled changes of the monad or the algebra of observations. The proofs of the invariance theorem and its corollary are in Appendix F.

Theorem 16 (Invariance Theorem)**.** *Let (M,η,μ) be a monad and $a\colon MO \to O$ an M-algebra. Let $c = \langle o,t\rangle\colon X \to O \times (MX)^A$ be an automaton with M-effects and observations in O and $[\![\cdot]\!]\colon MX \to O^{A^*}$ the semantic map induced by the generalised determinisation wrt. a, i.e., $[\![\cdot]\!] = [\![\cdot]\!]_{c^\sharp}$*

1) **Transitions:** *Let $(\hat{M},\hat{\eta},\hat{\mu})$ be a monad and $\sigma\colon M \Rightarrow \hat{M}$ a monad map. Let $\hat{a}\colon \hat{M}O \to O$ be an $\hat{M}$-algebra. Consider the coalgebra*
$$\hat{c} = \langle o,\hat{t}\rangle = \langle o, \sigma_X^A \circ t\rangle\colon X \to O \times (\hat{M}X)^A$$
and let $\hat{[\![\cdot]\!]}\colon \hat{M}X \to O^{A^*}$ be the semantic map induced by its generalised determinisation wrt. $\hat{a}$. If $a = \hat{a}\circ\sigma_O$, then $[\![\cdot]\!]\circ\eta_X = \hat{[\![\cdot]\!]}\circ\hat{\eta}_X$.

2) **Observations:** *Let $\hat{a}\colon M\hat{O} \to \hat{O}$ be an M-algebra and let $h\colon (O,a) \to (\hat{O},\hat{a})$ be an M-algebra morphism. Consider the coalgebra*
$$\hat{c} = \langle \hat{o},t\rangle = \langle h\circ o, t\rangle\colon X \to \hat{O}\times(MX)^A$$
and let $\hat{[\![\cdot]\!]}\colon TX \to \hat{O}^{A^*}$ be induced by the generalised determinisation wrt. $\hat{a}$. Then $\hat{[\![\cdot]\!]} = h^{A^*}\circ[\![\cdot]\!]$. □

Corollary 17. *Let (M,η,μ) be a submonad of $(\hat{M},\hat{\eta},\hat{\mu})$ via an injective monad map $\sigma\colon M \Rightarrow \hat{M}$. Let $t\colon X \to (MX)^A$ be a system with M-effects and let $\hat{t}$ be the system with $\hat{M}$-effects $\sigma_X^A \circ t\colon X \to (\hat{M}X)^A$. Let $o = \big(X \xrightarrow{!} 1 \xrightarrow{\eta_1} M1\big)$ and $\hat{o} = \big(X \xrightarrow{!} 1 \xrightarrow{\hat{\eta}_1} \hat{M}1\big)$ and $\equiv, \hat{\equiv} \subseteq X \times X$ be the corresponding trace equivalences after determinisation of $\langle o,t\rangle$, $\langle \hat{o},\hat{t}\rangle$, respectively. Then $\equiv = \hat{\equiv}$.* □

6. May / Must Traces for NPLTS

In this section, we put all the pieces together and give the definitions of may, must, and may-must trace semantics for NPLTS using generalised determinisation. We work with the monad $T_{\mathcal{PCS}} = C(\cdot + 1)$ and consider its two quotients $T_{\mathcal{CSB}}$ and $T_{\mathcal{CST}}$. Each of these choices gives us a trace equivalence via determinisation. We start with the notion of may-must traces.

May-must trace equivalence. Given an NPLTS $t\colon X \to (\mathcal{PD}X)^A$, let $(\bar{o}, \bar{t}\,)$ be the automaton with $T_{\mathcal{PCS}}$-effects and observations in $T_{\mathcal{PCS}}1$ as in Equation (4) and Equation (5). Let $\langle \bar{o}^\sharp, \bar{t}^\sharp \rangle$ be the determinisation of $\langle o, t \rangle$ using the free $T_{\mathcal{PCS}}$-algebra, i.e., by Proposition 8, the min-max interval pointed convex semilattice $\mathbb{M}_{\mathcal{I},[0,0]}$, on $T_{\mathcal{PCS}}1$. We write $[\![\cdot]\!]$ for the semantics map from $T_{\mathcal{PCS}}X \to (T_{\mathcal{PCS}}1)^{A^*}$ and $\equiv$ for the corresponding trace equivalence on X. We call this equivalence *may-must trace equivalence* for the original NPLTS.

Using the presentation of the monad, as in Equation (3), recalling that $\bar{o}(x) = [1,1]$ we can spell out the inductive definition of the determinisation:

$$\bar{o}^\sharp(S) \quad = \quad \begin{cases} [1,1] & \text{if } S = x; \\ [0,0] & \text{if } S = \star; \\ S_1 \text{ min-max}\, S_2 & \text{if } S = S_1 \oplus S_2; \\ S_1 +_p S_2 & \text{if } S = S_1 +_p S_2. \end{cases}$$

$$\bar{t}^\sharp(S)(a) \quad = \quad \begin{cases} \bar{t}(x)(a) & \text{if } S = x; \\ \star & \text{if } S = \star; \\ \bar{t}^\sharp(S_1)(a) \oplus \bar{t}^\sharp(S_2)(a) & \text{if } S = S_1 \oplus S_2; \\ \bar{t}^\sharp(S_1)(a) +_p \bar{t}^\sharp(S_2)(a) & \text{if } S = S_1 +_p S_2. \end{cases}$$

May trace equivalence and must trace equivalence. Now one may want to treat termination in a different way and exploit the monads $T_{\mathcal{CSB}}$ and $T_{\mathcal{CST}}$ discussed in Section 4. Given the monad morphisms $q_B\colon T_{\mathcal{PCS}} \Rightarrow T_{\mathcal{CSB}}$ and $q_T\colon T_{\mathcal{PCS}} \Rightarrow T_{\mathcal{CST}}$ quotienting $T_{\mathcal{PCS}}$ by (B) and (T), respectively, one can construct the transition functions

$$\bar{t}_B = q_B^A \circ \bar{t}\colon X \to (T_{\mathcal{CSB}}X)^A$$
$$\bar{t}_T = q_T^A \circ \bar{t}\colon X \to (T_{\mathcal{CST}}X)^A.$$

For the observations, we always use the general recipe of Section 5.3 and take the observation functions:

$$\bar{o}_B = \big(X \xrightarrow{!} 1 \xrightarrow{\eta_1} T_{\mathcal{CSB}}1\big) \quad \bar{o}_T = \big(X \xrightarrow{!} 1 \xrightarrow{\eta_1} T_{\mathcal{CST}}1\big).$$

Recall from Proposition 9 and Proposition 10 that $\mathbb{M}\text{ax}_B = ([0,1], \max, +_p, 0)$ and $\mathbb{M}\text{in}_T = ([0,1], \min, +_p, 0)$ are, the free convex semilattice with bottom and, respectively, with top, generated by the singleton set 1. Therefore these algebraic structures will be used for the determinisations. We have, $\bar{o}_B^\sharp\colon T_{\mathcal{CSB}}X \to [0,1]$ and $\bar{o}_T^\sharp\colon T_{\mathcal{CST}}X \to [0,1]$ are given as follows, since $\bar{o}_B(x) = 1$ and $\bar{o}_T(x) = 1$:

$$\bar{o}_B^\sharp(S) \quad = \quad \begin{cases} 1 & \text{if } S = x; \\ 0 & \text{if } S = \star; \\ S_1 \max S_2 & \text{if } S = S_1 \oplus S_2; \\ S_1 +_p S_2 & \text{if } S = S_1 +_p S_2. \end{cases}$$

$$\bar{o}_T^\sharp(S) \quad = \quad \begin{cases} 1 & \text{if } S = x; \\ 0 & \text{if } S = \star; \\ S_1 \min S_2 & \text{if } S = S_1 \oplus S_2; \\ S_1 +_p S_2 & \text{if } S = S_1 +_p S_2. \end{cases}$$

The determinisation of the transition function $\bar{t}_B^\sharp\colon T_{\mathcal{CSB}}X \to (T_{\mathcal{CSB}}X)^A$ and $\bar{t}_T^\sharp\colon T_{\mathcal{CST}}X \to (T_{\mathcal{CST}}X)^A$ are defined in the same way like $\bar{t}^\sharp$ above.

The coalgebras $\langle \bar{o}_B^\sharp, \bar{t}_B^\sharp \rangle$ and $\langle \bar{o}_T^\sharp, \bar{t}_T^\sharp \rangle$ give rise to morphisms $[\![\cdot]\!]_B\colon T_{\mathcal{CSB}}X \to [0,1]^{A^*}$ and $[\![\cdot]\!]_T\colon T_{\mathcal{CST}}X \to [0,1]^{A^*}$ and corresponding behavioural equivalences: $\equiv_B$ and $\equiv_T$. We call $\equiv_B$ the *may trace equivalence* for the NPLTS, and $\equiv_T$ the *must trace equivalence*.

Example 18. *Consider the convex closure of the NPLTS from Figure 1. We can syntactically describe the sets of subdistributions reached by a state when performing a transition as follows:*

$$x \xrightarrow{a} x_1 \oplus (x_3 +_{\frac{1}{2}} x_2)$$
$$y \xrightarrow{a} y_1 \oplus (y_4 +_{\frac{1}{2}} y_2) \oplus ((y_2 +_{\frac{1}{2}} y_4) +_{\frac{1}{2}} y_3)$$
$$x_1 \xrightarrow{b} x +_{\frac{1}{2}} x_3 \qquad y_1 \xrightarrow{b} y +_{\frac{1}{2}} y_4$$
$$x_2 \xrightarrow{b} x_3 \qquad x_2 \xrightarrow{c} x \qquad y_2 \xrightarrow{b} y_4 \qquad y_3 \xrightarrow{c} y$$

In the determinised system, we have

$$x \xrightarrow{a} S_1 \xrightarrow{b} S_2 \qquad y \xrightarrow{a} S_1' \xrightarrow{b} S_2'$$

for $\quad S_1 = x_1 \oplus (x_3 +_{\frac{1}{2}} x_2) \qquad S_2 = (x +_{\frac{1}{2}} x_3) \oplus (\star +_{\frac{1}{2}} x_3)$

$$S_1' = y_1 \oplus (y_4 +_{\frac{1}{2}} y_2) \oplus ((y_2 +_{\frac{1}{2}} y_4) +_{\frac{1}{2}} y_3)$$
$$S_2' = (y +_{\frac{1}{2}} y_4) \oplus (\star +_{\frac{1}{2}} y_4) \oplus ((y_4 +_{\frac{1}{2}} \star) +_{\frac{1}{2}} \star)$$

Consider now the observations associated to the terms in the may-must semantics. We have $\bar{o}^\sharp(x) = [1,1] = \bar{o}^\sharp(y)$ *and hence*

$$\bar{o}^\sharp(S_1) = [1,1] \text{ min-max}([1,1] +_{\frac{1}{2}} [1,1]) = [1,1].$$

Analogously, $\bar{o}^\sharp(S_1') = [1,1]$. *Furtheron*

$$\bar{o}^\sharp(S_2) = ([1,1] +_{\frac{1}{2}} [1,1]) \text{ min-max}([0,0] +_{\frac{1}{2}} [1,1]) = [\frac{1}{2}, 1]$$

and in the same way we derive $\bar{o}^\sharp(S_2') = [\frac{1}{4}, 1]$. *Hence, x and y are not may-must trace equivalent:* $[\![x]\!](ab) = \bar{o}^\sharp(S_2) \neq \bar{o}^\sharp(S_2') = [\![y]\!](ab)$.

However, using $\mathbb{M}\text{ax}_B$, *we get* $\bar{o}_B^\sharp(S_2) = \bar{o}_B^\sharp(S_2')$ *as the intervals obtained via the may-must observation over* S_2, S_2' *have the same upper bound* 1, *which is the value returned by both* $\bar{o}_B^\sharp(S_2)$ *and* $\bar{o}_B^\sharp(S_2')$. *Hence,* $[\![x]\!]_B(ab) = \bar{o}_B^\sharp(S_2) =$

$\bar{o}_B^\sharp(S_2') = [\![y]\!]_B(ab)$. *More generally, it holds that x and y are may trace equivalent. We can elegantly prove this by using up-to techniques, as shown in Appendix A.*

The following properties follow automatically from our abstract construction: see Theorem 13 and the discussions in Section 5.3.

Theorem 19. *The following properties hold for NPLTS:*

1) *Each of the three trace equivalences is a congruence w.r.t. $+_p$, $\oplus$ and $\star$.*
2) *Both bisimilarity and convex bisimilarity imply each of the three trace equivalences.*
3) *Up-to context is compatible (see Appendix A) for each of the three equivalences.* $\square$

We might have performed the generalised determinisation in a number of different ways, for instance by eliminating conv from the definition of $\bar{t}$. In Appendix G, we show that Theorem 16 guarantees that many different construction always lead to our semantics. In the same appendix we give also a simple concrete description of the final-coalgebra bialgebra of probabilistic traces.

Backward compatibility. We now state the backward compatibility of our semantics with the corresponding trace semantics for LTS and RPLTS. The proof follows from Corollary 17, since: (1) $\mathcal{P} \cong T_{\mathcal{SB}}$ for the theory $\mathcal{SB}$ of semilattices with bottom and we show that there is an injective monad map $T_{\mathcal{SB}} \Rightarrow T_{\mathcal{CSB}}$; and (2) The natural transformation $\mathrm{conv} \circ \eta^{\mathcal{P}_{ne}} \colon \mathcal{D} \Rightarrow C$ is an injective monad map and hence, by Lemma 1, there is an injective monad map $\mathcal{D}(\cdot + 1) \Rightarrow C(\cdot + 1)$.

Theorem 20. *Trace semantics $\equiv^{LTS}$ for LTS coincides with may trace semantics after determinisation $\equiv_B$ of the LTS seen as NPLTS. Trace semantics $\equiv^{RP}$ for RPLTS coincides with each of the three (may, must, and may-must) trace semantics $\equiv$ of the RPLTS seen as NPLTS.* $\square$

For LTS, one can also study the variants corresponding to must and may-must trace semantics, that have not been studied in the literature. We define them in Appendix G, and show backward compatibility results for them as well.

7. From the global to the local perspective

Usually trace semantics for NPLTS is defined in terms of *schedulers*, or *resolutions*: intuitively, a scheduler resolves the nondeterminism by choosing, at each step of the execution of an NPLTS, one of its possible transitions; the transition systems resulting from these choices are called resolutions.

This perspective on trace semantics is somehow opposed to ours, where the generalised determinisation keeps track of all possible executions at once. In this sense, the determinisation provides a perspective which is *global*, opposite to those of resolutions that are *local*. In this section, we show that our semantics can be characterised through such local views, by means of resolutions, defined as follows.

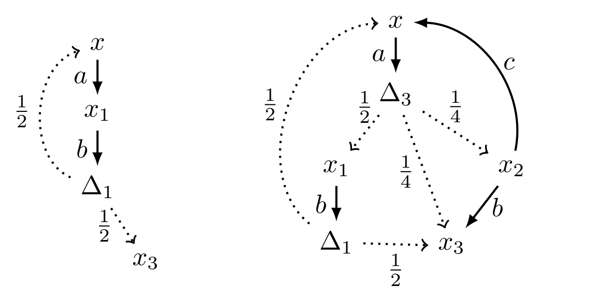


Figure 2. The resolutions $\mathcal{R}_1$ (left) and $\mathcal{R}_2$ (right)

Definition 21. *Let $t \colon X \to (\mathcal{PD}X)^A$ be an NPLTS. A* (randomized) resolution *for t is a triple $\mathcal{R} = (Y, \mathrm{corr}, r)$ where Y is a set of states,* $\mathrm{corr} \colon Y \to X$ *is* the correspondence function, *and $r \colon Y \to (\mathcal{D}Y + 1)^A$ is an RPLTS such that for all $y \in Y$ and $a \in A$,*

1) *$r(y)(a) = \star$ iff $t(\mathrm{corr}(y))(a) = \star$,*
2) *if $r(y)(a) \neq \star$ then $\mathcal{D}(\mathrm{corr})(r(y)(a)) \in \mathrm{conv}(t(\mathrm{corr}(y))(a))$.*

Intuitively, this means that a resolution of an NPLTS is built from the original system by discarding internal nondeterminism (the possibility to perform multiple transitions labelled with the same action) and in such a way that the structure of the original system is preserved.

Example 22. *Consider the NPLTS on the left of Figure 1. Figure 2 illustrates two resolutions for it, both having the identity as correspondence function. In the resolution $\mathcal{R}_1$, the nondeterministic choice of x is resolved by choosing the leftmost a-transition. Instead, the resolution $\mathcal{R}_2$ is obtained by taking a convex combination of the two distributions δ_{x_1} and Δ_1, assigning one half probability to each of them.*

The reason why we take arbitrary corr functions, rather than just injective ones, is that the original NPLTS might contain cycles, in which case we want to allow the resolution to take different choices at different times (see Appendix B.1).

Given a resolution $\mathcal{R} = (Y, \mathrm{corr}, r)$, we define the function $\mathrm{prob}_{\mathcal{R}} \colon Y \to [0,1]^{A^*}$ inductively for all $y \in Y$ and all $w \in A^*$ as

$$\mathrm{prob}_{\mathcal{R}}(y)(\varepsilon) = 1;$$
$$\mathrm{prob}_{\mathcal{R}}(y)(aw) =$$
$$\begin{cases} 0 & \text{if } r(y)(a) = \star; \\ \sum_{y' \in \mathrm{supp}(\Delta)} \Delta(y') \cdot \mathrm{prob}_{\mathcal{R}}(y')(w) & \text{if } r(y)(a) = \Delta. \end{cases}$$

Intuitively, for all states $y \in Y$, $\mathrm{prob}_{\mathcal{R}}(y)(w)$ gives the probability of y performing the trace w. For instance, in the resolutions in Figure 2, $\mathrm{prob}_{\mathcal{R}_1}(x)(abab) = \frac{1}{4}$ and $\mathrm{prob}_{\mathcal{R}_2}(x)(abab) = \frac{3}{16}$.

Now, given an NPLTS (X, t), define $\lfloor\!\lfloor \cdot \rfloor\!\rfloor \colon X \to [0,1]^{A^*}$ with, for all $x \in X$ and $w \in A^*$, $\lfloor\!\lfloor x \rfloor\!\rfloor(w)$ equal to

$$\bigsqcup \{\mathrm{prob}_{\mathcal{R}}(y)(w) \mid \mathcal{R} = (Y, \mathrm{corr}, r) \text{ is a resolution of } (X,t) \text{ and } \mathrm{corr}(y) = x\}.$$

Similarly, we define $\lceil\!\lceil x\rceil\!\rceil(w)$ as

$$\bigsqcap\{\mathrm{prob}_{\mathcal{R}}(y)(w)\ \mid \mathcal{R}=(Y,\mathrm{corr},r) \text{ is a resolution of } (X,t) \text{ and } \mathrm{corr}(y)=x\}.$$

The following theorem states that the global view of trace semantics developed in Section 6 coincides with the trace semantics defined locally via resolutions.

Theorem 23 (Global/local correspondence)**.** *Let (X,t) be an NPLTS. For all $x\in X$ and $w\in A^*$, it holds that*

$$[\![x]\!](w)=[\,\lceil\!\lceil x\rceil\!\rceil(w),\lfloor\!\lfloor x\rfloor\!\rfloor(w)\,].$$

Corollary 24. *Let (X,t) be an NPLTS. For all $x\in X$ and $w\in A^*$, $[\![x]\!]_B(w)=\lfloor\!\lfloor x\rfloor\!\rfloor(w)$ and $[\![x]\!]_T(w)=\lceil\!\lceil x\rceil\!\rceil(w)$.*

Theorem 23 and Corollary 24 provide a characterisation of $\equiv$, $\equiv_B$ and $\equiv_T$ in terms of resolutions. We next show that $\equiv_B$ coincides with the randomized $\sqcup$-trace equivalence investigated in [40] and inspired by [25], [26].

Coincidence with randomized $\sqcup$-trace equivalence. Let $t\colon X\to(\mathcal{P}\mathcal{D}X)^A$ be an NPLTS. A *fully probabilistic resolution* for t is a triple $\mathcal{R}=(Y,\mathrm{corr},r)$ such that Y is a set, $\mathrm{corr}\colon Y\to X$, and $r\colon Y\to(A\times\mathcal{D}Y)+1$ such that for every $y\in Y$ and $a\in A$ it holds:
if $r(y)=\langle a,\Delta\rangle$ then $\mathcal{D}(\mathrm{corr})(\Delta)\in\mathrm{conv}(t(\mathrm{corr}(y))(a))$.

While resolutions resolve only internal nondeterminism, fully probabilistic resolutions resolve both internal and external nondeterminism. Indeed, in a resolution a state can perform transitions with different labels, while in a fully probabilistic resolution a state can perform at most one transition. Moreover, a state y in a fully probabilistic resolutions might not perform any transition (i.e., $r(y)=\star$), even if the corresponding state $\mathrm{corr}(y)$ may perform a transition (i.e., $t(\mathrm{corr}(y))(a)\neq\star$ for some a).

Example 25. *As in Example 22, consider the NPLTS on the left of Figure 1. The resolution $\mathcal{R}_1$ in Figure 2 is a fully probabilistic resolution, while $\mathcal{R}_2$ is not, since x_2 is allowed to perform more than one transition, even if labelled by different actions. Other examples of fully probabilistic resolutions are given in Appendix B.2.*

As for resolutions, we can define $\mathrm{prob}_{\mathcal{R}}\colon Y\to[0,1]^{A^*}$ for $\mathcal{R}=(Y,\mathrm{corr},r)$ a fully probabilistic resolution inductively for all $y\in Y$ and all $w\in A^*$ as

$$\mathrm{prob}_{\mathcal{R}}(y)(\varepsilon)=1;$$
$$\mathrm{prob}_{\mathcal{R}}(y)(aw)=$$
$$\begin{cases}\sum_{y'\in\mathrm{supp}(\Delta)}\Delta(y')\cdot\mathrm{prob}_{\mathcal{R}}(y')(w) & \text{if } r(y)=\langle a,\Delta\rangle,\\ 0 & \text{otherwise.}\end{cases}$$

Given an NPLTS (X,t), we define for all $x\in X$ and $w\in A^*$ $\lfloor\!\lfloor x\rfloor\!\rfloor_{fp}(w)$ as

$$\bigsqcup\{\mathrm{prob}_{\mathcal{R}}(y)(w)\ \mid \mathcal{R}=(Y,\mathrm{corr},r) \text{ is a fully probabilistic resolution of } (X,t) \text{ and } \mathrm{corr}(y)=x\}.$$

In [40] (following [25], [26]), two states x and y are defined to be *randomized $\sqcup$-trace equivalent* whenever $\lfloor\!\lfloor x\rfloor\!\rfloor_{fp}(w)=\lfloor\!\lfloor y\rfloor\!\rfloor_{fp}(w)$ $w\in A^*$[2]. The following proposition guarantees that such equivalence coincides with $\equiv_B$.

Proposition 26. *Let (X,t) be an NPLTS. For all $x\in X$ and $w\in A^*$, it holds that $[\![x]\!]_B(w)=\lfloor\!\lfloor x\rfloor\!\rfloor(w)=\lfloor\!\lfloor x\rfloor\!\rfloor_{fp}(w)$.*

Remark 27. *The correspondence in Proposition 26 does not hold when infima are considered, instead of suprema. Indeed define $\lceil\!\lceil x\rceil\!\rceil_{fp}(w)$ as expected, namely, by replacing $\bigsqcup$ with $\bigsqcap$ in $\lfloor\!\lfloor x\rfloor\!\rfloor_{fp}(w)$. Then for any state x of an arbitrary NPLTS it holds that $\lceil\!\lceil x\rceil\!\rceil_{fp}(w)=0$ for all $w\neq\varepsilon$. To see this, observe that $\mathcal{R}'=(\{y\},\mathrm{corr}',r')$ with $\mathrm{corr}'(y)=x$ and $r'(y)=\star$ is always a fully probabilistic resolution, and that $\mathrm{prob}_{\mathcal{R}'}(y)(w)=0$.*

To avoid this problem, one typically modifies the definition of $\lceil\!\lceil\cdot\rceil\!\rceil_{fp}$ by restricting only to those fully probabilistic resolutions that can perform a certain trace (see e.g. [25], [26]). Instead, with our notion of resolution based on RPLTSs (Definition 21), this problems does not arise and the definition of $\lceil\!\lceil\cdot\rceil\!\rceil$ is totally analogous to the one of $\lfloor\!\lfloor\cdot\rfloor\!\rfloor$.

Why may, must, may-must? Trace equivalences as testing equivalences. The notion of resolution is at the basis not just of the definitions of trace equivalences for NPLTS investigated in the literature, but also of testing equivalences for nondeterministic and probabilistic processes [33], [34], [35], [36]. In testing equivalences, we say that x,y are may testing equivalent if, for every test, they have the same greatest probabilities of passing the test, with respect to any resolution $\mathcal{R}$ of the system resulting from the interaction between the test and the NPLTS. Analogously, x,y are must testing equivalent if the smallest probabilities coincide, and the may-must testing equivalence requires both the greatest and the smallest probabilities to coincide.

Now, take tests to be finite traces, and the probability of passing a given test in a resolution as the probability of performing the trace in the resolution. Then it becomes clear, by the correspondence between the local and the global view proven in Theorem 23, that each of our three trace equivalences indeed coincides with the corresponding testing equivalence, when tests are finite traces.

8. Conclusion

We developed an algebra-and-coalgebra-based trace theory for systems with nondeterminism and probability, that covers intricate trace semantics from the literature. The abstract approach sheds light on all choices and leaves no space for ad-hoc solutions.

The combination of nondeterminism and probability has been considered notorious for many years, and for good reasons. In our view, this new algebraic theory of traces for NPLTS shows that their bad reputation is not deserved.

2. Actually, [25], [26], [40] use a notion of resolution which is equal to our fully-probabilistic resolution modulo a tiny modification due to a mistake in [25], [26], as confirmed by the authors in a personal communication.

References

[1] C. Baier and J. Katoen, *Principles of model checking*. MIT Press, 2008.

[2] H. Hermanns, J. Krcál, and J. Kretínský, "Probabilistic bisimulation: Naturally on distributions," in *Proc. CONCUR'14*, ser. LNCS, vol. 8704, 2014, pp. 249–265.

[3] M. Z. Kwiatkowska, G. Norman, and D. Parker, "Prism: Probabilistic symbolic model checker," in *Computer Performance Evaluation / TOOLS*. LNCS 2324, 2002, pp. 200–204.

[4] C. Dehnert, S. Junges, J. Katoen, and M. Volk, "A storm is coming: A modern probabilistic model checker," in *Proc. CAV 2017*, ser. LNCS, vol. 10427, 2017, pp. 592–600.

[5] M. Y. Vardi, "Automatic verification of probabilistic concurrent finite state programs," in *Foundations of Computer Science, 1985., 26th Annual Symposium on*. IEEE, 1985, pp. 327–338.

[6] H. A. Hansson, "Time and probability in formal design of distributed systems," *PhD thesis, Uppsala University*, 1991.

[7] R. Segala and N. Lynch, "Probabilistic simulations for probabilistic processes," *Nordic Journal of Computing*, vol. 2, no. 2, pp. 250–273, 1995.

[8] P. S. Castro, P. Panangaden, and D. Precup, "Equivalence relations in fully and partially observable markov decision processes," in *IJCAI*, 2009, pp. 1653–1658.

[9] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra, "Planning and Acting in Partially Observable Stochastic Domains." *Artif. Intell.*, 1998.

[10] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*. Prentice Hall, 2009.

[11] C. Heunen, O. Kammar, S. Staton, and H. Yang, "A convenient category for higher-order probability theory," *CoRR*, vol. abs/1701.02547, 2017. [Online]. Available: http://arxiv.org/abs/1701.02547

[12] S. Staton, H. Yang, F. Wood, C. Heunen, and O. Kammar, "Semantics for probabilistic programming: higher-order functions, continuous distributions, and soft constraints," in *Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in Computer Science, LICS '16, New York, NY, USA, July 5-8, 2016*, 2016, pp. 525–534. [Online]. Available: http://doi.acm.org/10.1145/2933575.2935313

[13] H. Hermanns, A. Parma, R. Segala, B. Wachter, and L. Zhang, "Probabilistic logical characterization," *Information and Computation*, vol. 209, no. 2, pp. 154–172, 2011.

[14] D. Varacca and G. Winskel, "Distributing probabililty over nondeterminism," *MSCS*, vol. 16, no. 1, pp. 87–113, 2006.

[15] K. Keimel and G. D. Plotkin, "Mixed powerdomains for probability and nondeterminism," *Logical Methods in Computer Science*, vol. 13, no. 1, 2017. [Online]. Available: https://doi.org/10.23638/LMCS-13(1:2)2017

[16] M. Mio, "Upper-expectation bisimilarity and łukasiewicz μ-calculus," in *Proc. FOSSACS'14*, ser. LNCS, vol. 8412, 2014, pp. 335–350.

[17] B. Jacobs, "Coalgebraic trace semantics for combined possibilitistic and probabilistic systems," *Electr. Notes Theor. Comput. Sci.*, vol. 203, no. 5, pp. 131–152, 2008.

[18] D. Varacca, *Probability, Nondeterminism and Concurrency. Two Denotational Models for Probabilistic Computation*. PhD Thesis, 2003.

[19] M. W. Mislove, "Nondeterminism and probabilistic choice: Obeying the laws," in *CONCUR 2000*. LNCS 1877, 2000, pp. 350–364. [Online]. Available: https://doi.org/10.1007/3-540-44618-4_26

[20] J. Goubault-Larrecq, "Prevision domains and convex powercones," in *FOSSACS 2008*. LNCS 4962, 2008, pp. 318–333. [Online]. Available: https://doi.org/10.1007/978-3-540-78499-9_23

[21] R. Tix, K. Keimel, and G. D. Plotkin, "Semantic domains for combining probability and non-determinism," *ENTCS*, vol. 222, pp. 3–99, 2009. [Online]. Available: https://doi.org/10.1016/j.entcs.2009.01.002

[22] R. Glabbeek, "The linear time - branching time spectrum i. the semantics of concrete, sequential processes," in *Handbook of Process Algebra*, J. A. Bergstra, A. Ponse, and S. Smolka, Eds. Amsterdam: Elsevier, 2001, pp. 3–99.

[23] R. Segala, "Modeling and verification of randomized distributed real-time systems," Ph.D. dissertation, MIT, 1995.

[24] ——, "A compositional trace-based semantics for probabilistic automata," in *Proc. CONCUR'95*, ser. LNCS, vol. 962. Springer, 1995, pp. 234–248.

[25] M. Bernardo, R. De Nicola, and M. Loreti, "Revisiting trace and testing equivalences for nondeterministic and probabilistic processes," *Logical Methods in Computer Science*, vol. 10(1:16), pp. 1–42, 2014.

[26] ——, "Relating strong behavioral equivalences for processes with nondeterminism and probabilities," *Theoretical Computer Science*, 2014.

[27] J. Rutten, "Universal coalgebra: A theory of systems," *Theoretical Computer Science*, vol. 249, pp. 3–80, 2000.

[28] B. Jacobs, "Introduction to coalgebra. towards mathematics of states and observations," 2005, book in preparation, draft available via `http://www.cs.ru.nl/~bart`.

[29] I. Hasuo, B. Jacobs, and A. Sokolova, "Generic trace semantics via coinduction," *Logical Methods in Computer Science*, vol. 3, no. 4, 2007. [Online]. Available: https://doi.org/10.2168/LMCS-3(4:11)2007

[30] A. Silva, F. Bonchi, M. Bonsangue, and J. Rutten, "Generalizing the powerset construction, coalgebraically," in *Proc. FSTTCS 2010*, ser. Leibniz International Proceedings in Informatics (LIPIcs), vol. 8, 2010, pp. 272–283.

[31] M. Rabin, "Probabilistic automata," *Information and Control*, vol. 6, pp. 230–245, 1963.

[32] D. Varacca, "Probability, nondeterminism and concurrency: Two denotational models for probabilistic computation," Ph.D. dissertation, Univ. Aarhus, 2003, bRICS Dissertation Series, DS-03-14.

[33] W. Yi and K. Larsen, "Testing probabilistic and nondeterministic processes," in *Proc. PSTV'92*. North-Holland, 1992, pp. 47–61.

[34] B. Jonsson, C. Ho-Stuart, and W. Yi, "Testing and refinement for nondeterministic and probabilistic processes," in *Proc. FTRTFT'94*, ser. LNCS, vol. 863. Springer, 1994, pp. 418–430.

[35] Y. Deng, R. Glabbeek, M. Hennessy, C. Morgan, and C. Zhang, "Characterising testing preorders for finite probabilistic processes," in *Proc. LICS'07*. IEEE-CS Press, 2007, pp. 313–325.

[36] Y. Deng, R. Glabbeek, M. Hennessy, and C. Morgan, "Testing finitary probabilistic processes," in *Proc. CONCUR'09*, ser. LNCS, vol. 5710. Springer, 2009, pp. 274–288.

[37] R. Milner, *Communication and Concurrency*. Prentice Hall, 1989.

[38] D. Pous and D. Sangiorgi, "Enhancements of the bisimulation proof method," in *Advanced Topics in Bisimulation and Coinduction*, D. Sangiorgi and J. Rutten, Eds. Cambridge University Press, 2012.

[39] R. Segala, "Modeling and verification of randomized distributed real-time systems," Ph.D. dissertation, MIT, 1995.

[40] V. Castiglioni, "Trace and testing metrics on nondeterministic probabilistic processes," in *Proceedings Combined 25th International Workshop on Expressiveness in Concurrency and 15th Workshop on Structural Operational Semantics and 15th Workshop on Structural Operational Semantics, EXPRESS/SOS 2018, Beijing, China, September 3, 2018.*, 2018, pp. 19–36. [Online]. Available: https://doi.org/10.4204/EPTCS.276.4

[41] T. Świrszcz, "Monadic functors and convexity," *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.*, vol. 22, pp. 39–42, 1974.

[42] Z. Semadeni, *Monads and their Eilenberg-Moore algebras in functional analysis*. Queen's University, Kingston, Ont., 1973, queen's Papers in Pure and Applied Mathematics, No. 33.

[43] E.-E. Doberkat, "Eilenberg-Moore algebras for stochastic relations," *Inform. and Comput.*, vol. 204, no. 12, pp. 1756–1781, 2006. [Online]. Available: http://dx.doi.org/10.1016/j.ic.2006.09.001

[44] ——, "Erratum and addendum: Eilenberg-Moore algebras for stochastic relations [mr2277336]," *Inform. and Comput.*, vol. 206, no. 12, pp. 1476–1484, 2008. [Online]. Available: http://dx.doi.org/10.1016/j.ic.2008.08.002

[45] B. Jacobs, "Convexity, duality and effects," in *Theoretical computer science*, ser. IFIP Adv. Inf. Commun. Technol. Springer, Berlin, 2010, vol. 323, pp. 1–19. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-15240-5_1

[46] M. Hyland, G. Plotkin, and J. Power, "Combining computational effects: Commutativity and sum," in *Foundations of Information Technology in the Era of Network and Mobile Computing*. Springer, 2002, pp. 474–484.

[47] A. Sokolova and H. Woracek, "Termination in convex sets of distributions," *Logical Methods in Computer Science*, vol. 14, no. 4, 2018.

[48] F. Dahlqvist, L. Parlant, and A. Silva, "Layer by layer - combining monads," in *Theoretical Aspects of Computing - ICTAC 2018 - 15th International Colloquium, Stellenbosch, South Africa, October 16-19, 2018, Proceedings*, ser. Lecture Notes in Computer Science, B. Fischer and T. Uustalu, Eds., vol. 11187. Springer, 2018, pp. 153–172. [Online]. Available: https://doi.org/10.1007/978-3-030-02508-3_9

[49] G. van Heerdt, J. Hsu, J. Ouaknine, and A. Silva, "Convex Language Semantics for Nondeterministic Probabilistic Automata," in *Theoretical Aspects of Computing - ICTAC 2018 - 15th International Colloquium, Stellenbosch, South Africa, October 16-19, 2018, Proceedings*, ser. Lecture Notes in Computer Science, B. Fischer and T. Uustalu, Eds., vol. 11187. Springer, 2018, pp. 472–492. [Online]. Available: https://doi.org/10.1007/978-3-030-02508-3

[50] S. Staton, "Relating coalgebraic notions of bisimulation," *Logical Methods in Computer Science*, vol. 7, no. 1, 2011.

[51] A. Kurz, "Logics for coalgebras and applications to computer science," Ph.D. dissertation, Ludwig-Maximilians-Universität München, 2000.

[52] U. Wolter, "On corelations, cokernels, and coequations," *Electronic Notes in Theoretical Computer Science*, vol. 33, 2000.

[53] F. Bartels, A. Sokolova, and E. d. Vink, "A hierarchy of probabilistic system types," *Theoretical Computer Science*, vol. 327, pp. 3–22, 2004.

[54] A. Sokolova, "Probabilistic systems coalgebraically: A survey," *Theor. Comput. Sci.*, vol. 412, no. 38, pp. 5095–5110, 2011.

[55] B. Jacobs, A. Silva, and A. Sokolova, "Trace semantics via determinization," *J. Comput. Syst. Sci.*, vol. 81, no. 5, pp. 859–879, 2015.

[56] D. Turi and G. D. Plotkin, "Towards a mathematical operational semantics," in *Proceedings, 12th Annual IEEE Symposium on Logic in Computer Science, Warsaw, Poland, June 29 - July 2, 1997*. IEEE Computer Society, 1997, pp. 280–291. [Online]. Available: https://doi.org/10.1109/LICS.1997.614955

[57] B. Klin, "Bialgebras for structural operational semantics: An introduction," *Theor. Comput. Sci.*, vol. 412, no. 38, pp. 5043–5069, 2011. [Online]. Available: https://doi.org/10.1016/j.tcs.2011.03.023

[58] F. Bonchi, D. Petrisan, D. Pous, and J. Rot, "A general account of coinduction up-to," *Acta Inf.*, vol. 54, no. 2, pp. 127–190, 2017. [Online]. Available: https://doi.org/10.1007/s00236-016-0271-4

[59] F. Bonchi, M. M. Bonsangue, G. Caltais, J. Rutten, and A. Silva, "A coalgebraic view on decorated traces," *Mathematical Structures in Computer Science*, vol. 26, no. 7, pp. 1234–1268, 2016. [Online]. Available: https://doi.org/10.1017/S0960129514000449

[60] Y. Feng and L. Zhang, "When equivalence and bisimulation join forces in probabilistic automata," in *FM 2014: Formal Methods*, C. Jones, P. Pihlajasaari, and J. Sun, Eds. Cham: Springer International Publishing, 2014, pp. 247–262.

[61] P. Yang, D. N. Jansen, and L. Zhang, "Distribution-based bisimulation for labelled markov processes," in *FORMATS 2017*, 2017, pp. 170–186. [Online]. Available: https://doi.org/10.1007/978-3-319-65765-3_10

[62] D. Sangiorgi, "On the bisimulation proof method," vol. 8, pp. 447–479, 1998. [Online]. Available: http://dx.doi.org/10.1017/S0960129598002527

[63] D. Pous, "Complete lattices and up-to techniques," in *APLAS*, vol. 4807, 2007, pp. 351–366. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-76637-7_24

[64] D. Pumplün and H. Röhrl, "Convexity theories. IV. Klein-Hilbert parts in convex modules," *Appl. Categ. Structures*, vol. 3, no. 2, pp. 173–200, 1995. [Online]. Available: http://dx.doi.org/10.1007/BF00877635

[65] M. Stone, "Postulates for the barycentric calculus," *Ann. Mat. Pura Appl. (4)*, vol. 29, pp. 25–30, 1949. [Online]. Available: http://dx.doi.org/10.1007/BF02413910

[66] S. MacLane, *Categories for the working mathematician*. Springer-Verlag, 1971.

[67] F. Bonchi, A. Silva, and A. Sokolova, "The Power of Convex Algebras," in *CONCUR 2017*, vol. 85. LIPIcs, 2017, pp. 23:1–23:18.

[68] B. Fischer and T. Uustalu, Eds., *Theoretical Aspects of Computing - ICTAC 2018 - 15th International Colloquium, Stellenbosch, South Africa, October 16-19, 2018, Proceedings*, ser. Lecture Notes in Computer Science, vol. 11187. Springer, 2018. [Online]. Available: https://doi.org/10.1007/978-3-030-02508-3

Appendix A. Coinduction Up-to

As anticipated in Theorem 19, $\equiv$, $\equiv_B$ and $\equiv_T$ can be proved coinductively by means of bisimulation up-to. In order to define uniformly the proof techniques for the three equivalences, we let $\equiv_i$ to range over $\equiv$, $\equiv_B$ and $\equiv_T$; T_i to range over $T_{\mathcal{P}\mathcal{C}\mathcal{S}}$, $T_{\mathcal{C}\mathcal{S}\mathcal{B}}$ and $T_{\mathcal{C}\mathcal{S}\mathcal{T}}$; $\bar{t}_i^\sharp$ over $\bar{t}^\sharp$, $\bar{t}_B^\sharp$ and $\bar{t}_T^\sharp$; $\bar{o}_i^\sharp$ over $\bar{o}^\sharp$, $\bar{o}_B^\sharp$ and $\bar{o}_T^\sharp$.

Definition 28. *Let (X,t) be a NPLTS and $(T_i(X), \langle \bar{o}_i^\sharp, \bar{t}_i^\sharp \rangle)$ the corresponding determinised system. A relation $\mathcal{R} \subseteq T_i(X) \times T_i(X)$ is a bisimulation iff for all $S_1, S_2 \in T_i(X)$ it holds that*

1) $\bar{o}_i^\sharp(S_1) = \bar{o}_i^\sharp(S_2)$ and
2) $\bar{t}_i^\sharp(S_1)(a) \; \mathcal{R} \; \bar{t}_i^\sharp(S_2)(a)$ for all $a \in A$.

Coinduction tells us (see e.g. [58]) that for all $x, y \in X$, $x \equiv_i y$ iff there exists a bisimulation $\mathcal{R}$ such that $x \,\mathcal{R}\, y$.

To make this proof principle more effective, one can use *up-to techniques* [37], [38]. Particularly relevant for us is up-to *contextual closure* which, for all relations $\mathcal{R} \subseteq T_i(X) \times T_i(X)$, is defined inductively by the following rules.

$$\frac{S \; \mathcal{R} \; S'}{S \; Ctx(\mathcal{R}) \; S'} \qquad \frac{-}{* \; Ctx(\mathcal{R}) \; *}$$

$$\frac{S_1 \; Ctx(\mathcal{R}) \; S_1' \qquad S_2 \; Ctx(\mathcal{R}) \; S_2'}{S_1 \oplus S_2 \; Ctx(\mathcal{R}) \; S_1' \oplus S_2'}$$

$$\frac{S_1 \ Ctx(\mathcal{R}) \ S_1' \qquad S_2 \ Ctx(\mathcal{R}) \ S_2'}{S_1 +_p S_2 \ Ctx(\mathcal{R}) \ S_1' +_p S_2'}$$

Definition 29. *Bisimulations up-to context are defined as in Definition 28, but with $Ctx(\mathcal{R})$ replaced to $\mathcal{R}$ in point 2).*

By virtue of the general theory in [58], one has that Ctx is a *sound* up-to technique, that is $x \equiv_i y$ iff there exists a bisimulation up-to context $\mathcal{R}$ such that $x \,\mathcal{R}\, y$. Actually, the theory in [58] guarantees a stronger property known as compatibility [38], [62], [63]. Intuitively, this mean that the technique is sound and it can be safely combined with other compatible up-to techniques. We refer the interested reader to [38] for a detailed introduction to compatible up-to techniques.

We conclude with an example illustrating a finite bisimulation up-to context witnessing that the states x and y from Figure 1 are in $\equiv_B$.

Example 30. *Consider the NPLTS depicted in Figure 1. One can prove that $x \equiv_B y$ by exhibiting a bisimulation on $(T_{\mathcal{CSB}}X, \langle \bar{o}_B^\sharp, \bar{t}_B^\sharp \rangle)$ relating them. However, due to the presence of cycles, the determinization of the NPLTS is infinite and the bisimulation relation contains infinitely many pairs. The interested reader may check such bisimulation in Appendix B.3.*

With bisimulations up-to, only few pairs are necessary. Indeed, we prove that the relation

$$\begin{aligned}\mathcal{R} = \{&(x,y),(x_1,y_1),(x_3,y_4),\\ &(x_3 +_{\frac{1}{2}} x_2, (y_4 +_{\frac{1}{2}} y_2) \oplus ((y_2 +_{\frac{1}{2}} y_4) +_{\frac{1}{2}} y_3))\}\end{aligned}$$

is a bisimulation up-to context. First, note that the observation is trivially the same for all pairs in the relation, since $\bar{o}_B^\sharp(S) = 1$ for all S in the relation. We can now check that the clauses of bisimulation up-to context on the transitions are satisfied. Consider the first pair. In $(T_{\mathcal{CSB}}X, \langle \bar{o}_B^\sharp, \bar{t}_B^\sharp \rangle)$, we have

$$\begin{aligned}&x \xrightarrow{a} x_1 \oplus (x_3 +_{\frac{1}{2}} x_2)\\ &y \xrightarrow{a} y_1 \oplus (y_4 +_{\frac{1}{2}} y_2) \oplus ((y_2 +_{\frac{1}{2}} y_4) +_{\frac{1}{2}} y_3)\end{aligned}$$

The reached states are in $Ctx(\mathcal{R})$ by the second and fourth pairs of $\mathcal{R}$. For any action $a' \neq a$, we have $x \xrightarrow{a'} \star$, $y \xrightarrow{a'} \star$ and $\star \ Ctx(\mathcal{R}) \ \star$.

The second and the third pairs can be checked in a similar way. For the fourth pair, we have

$$\begin{gathered}x_3 +_{\frac{1}{2}} x_2 \xrightarrow{b} \star +_{\frac{1}{2}} x_3\\ (y_4 +_{\frac{1}{2}} y_2) \oplus ((y_2 +_{\frac{1}{2}} y_4) +_{\frac{1}{2}} y_3) \xrightarrow{b} (\star +_{\frac{1}{2}} y_4) \oplus ((\star +_{\frac{1}{2}} y_4) +_{\frac{1}{2}} \star)\end{gathered}$$

We observe that

$$\begin{aligned}&(\star +_{\frac{1}{2}} y_4) \oplus ((\star +_{\frac{1}{2}} y_4) +_{\frac{1}{2}} \star)\\ &\overset{(B)}{=} (\star +_{\frac{1}{2}} y_4) \oplus ((\star +_{\frac{1}{2}} y_4) +_{\frac{1}{2}} \star) \oplus \star\\ &\overset{(C)}{=} (\star +_{\frac{1}{2}} y_4) \oplus \star\\ &\overset{(B)}{=} \star +_{\frac{1}{2}} y_4\end{aligned}$$

and we conclude by $\star +_{\frac{1}{2}} x_3 \ Ctx(\mathcal{R}) \ \star +_{\frac{1}{2}} y_4$. The cases for a and c are simpler.

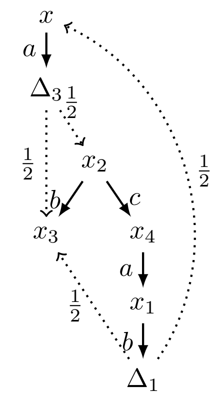


Figure 3. The resolution $\mathcal{R}_3$

Appendix B. Additional examples

B.1. Example: a resolution with non-injective correspondence function

In order to understand how a resolution allows to resolve differently nondeterministic choices at different times, when cycles occur in the original system, consider the NPLTS on the left of Figure 1 and its resolution in Figure 3. In the latter, the state space is enlarged with state x_4, which is mapped to x by the correspondence function. On the remaining states, the correspondence function is the identity over X. In this resolution, x first chooses the right-hand transition of the original NPLTS, and at the second cycle, represented by x_4, the left-hand transition is chosen. Observe that $\mathrm{prob}_{\mathcal{R}_3}(x)(abab) = 0$.

B.2. Example: fully probabilistic resolutions

In Figure 4, we show three examples of fully probabilistic resolutions of the process x in Figure 1. Note that neither of them is a resolution. In $\mathcal{R}_1$, state x_1 does not satisfy the first clause of the definition of resolution, since x_1 does not move while its corresponding state in the original NPLTS does. In $\mathcal{R}_2$ and $\mathcal{R}_3$, state x_2 respectively only performs a b-labelled transition and only performs a c-labelled transition. In a resolution, it should perform both.

B.3. Example: infinite determinization and bisimulation

Consider the NPLTS (X, t) depicted in Figure 1, and discussed in Example 18. Figure 1 shows the determinization of the system, where the terms are defined as follows

$$S_1 \overset{\text{def}}{=} x_1 \oplus (x_3 +_{\frac{1}{2}} x_2) \qquad S_2 \overset{\text{def}}{=} (x +_{\frac{1}{2}} x_3) \oplus (\star +_{\frac{1}{2}} x_3)$$

$$S_1' \overset{\text{def}}{=} y_1 \oplus (y_4 +_{\frac{1}{2}} y_2) \oplus ((y_2 +_{\frac{1}{2}} y_4) +_{\frac{1}{2}} y_3)$$

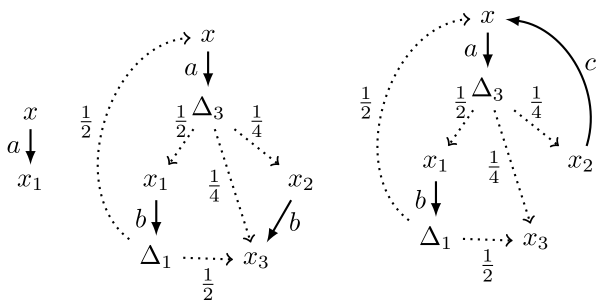


Figure 4. Fully probabilistic resolutions ($\mathcal{R}_1$, $\mathcal{R}_2$, $\mathcal{R}_3$, from left to right)

$$S'_2 \stackrel{\text{def}}{=} (y +_{\frac{1}{2}} y_4) \oplus (\star +_{\frac{1}{2}} y_4) \oplus ((\star +_{\frac{1}{2}} y_4) +_{\frac{1}{2}} \star)$$

$$S_3 \stackrel{\text{def}}{=} \star \oplus (\star +_{\frac{1}{2}} x) \qquad S'_3 \stackrel{\text{def}}{=} \star \oplus (\star +_{\frac{1}{2}} y)$$

and the depicted transitions are those given by $\bar{t}^\sharp$.

The determinized NPLTS is a system with infinitely many states, which are given by the presence of cycles in the original system. In the determinization of the automaton with algebra of observation $\mathbb{M}\mathrm{ax}_B$, each state S is in pair with the observation $\bar{o}^\sharp_B(S) \in [0,1]$. We prove that x and y are may trace equivalent by exhibiting the following bisimulation $\mathcal{R}$:

$$\begin{aligned}\mathcal{R} =&\{(x,y),(S_1,S'_1),(S_2,S'_2),(S_3,S'_3)\}\\ &\cup \{((S_i +_{\frac{1}{2^n}} \star)\oplus\star, (S'_i +_{\frac{1}{2^n}} \star)\oplus\star) \mid 1\le i\le 3, n\ge 1\}\end{aligned}$$

The relation satisfies the two clauses required by Definition 28 of bisimulation. As it emerges from Figure 5, the clause on transitions (clause 2) is satisfied by each pair in the relation. As to the clause on the observation (clause 1), we can derive as in Example 18 that for every pair $(S,S') \in \{(x,y),(S_1,S'_1),(S_2,S'_2),(S_3,S'_3)\}$ it holds $\bar{o}^\sharp_B(S) = \bar{o}^\sharp_B(S')$. Finally, clause 1 also holds for the remaining pairs, since for $1 \le i \le 3$ and $n \ge 1$ we have

$$\begin{aligned}\bar{o}^\sharp_B((S_i +_{\frac{1}{2^n}} \star)\oplus\star) &= (\bar{o}^\sharp_B(S_i) +_{\frac{1}{2^n}} 0) \max 0\\ &= (\bar{o}^\sharp_B(S'_i) +_{\frac{1}{2^n}} 0) \max 0\\ &= \bar{o}^\sharp_B((S'_i +_{\frac{1}{2^n}} \star)\oplus\star)\end{aligned}$$

Hence, $\mathcal{R}$ is a bisimulation.

As shown in Example 18, x,y are not bisimilar if the algebra of observation for the must equivalence, i.e, $\mathbb{M}\mathrm{in}_T$, is used instead of the one for the may equivalence, since $\bar{o}^\sharp_T(S_2) \neq \bar{o}^\sharp_B(S'_2)$. Analogously, they are not equivalent if we take the may-must algebra of observation $\mathbb{M}_{\mathcal{I},[0,0]}$.

Appendix C. Proofs for Section 2, Monads

For completeness, we recall that a *monad distributive law* of M over $\hat{M}$ is a natural transformation $\lambda\colon M\hat{M} \Rightarrow \hat{M}M$ that commutes appropriately with the units and the multiplications of the monads $\lambda \circ M\hat{\eta} = \eta\hat{M}$, $\lambda \circ M\hat{\mu} = \mu M \circ \hat{M}\lambda \circ \lambda\hat{M}$, and $\lambda \circ \eta\hat{M} = \hat{M}\eta$, $\lambda \circ \mu\hat{M} = \hat{M}\mu M \circ \lambda M \circ M\lambda$.

Lemma 1 follows directly from Lemma 31 and Lemma 32 below.

Lemma 31. *Given three monads M, $\hat{M}$, and T, two monad distributive laws $\lambda\colon TM \Rightarrow MT$ and $\hat{\lambda}\colon T\hat{M} \Rightarrow \hat{M}T$, ensuring that MT and $\hat{M}T$ are monads, and a monad map $\sigma\colon M \Rightarrow \hat{M}$. If the following diagram commutes, in which case we say that σ is a map of distributive laws,*

$$\begin{tikzcd} TM \arrow[r, "\lambda"] \arrow[d, "T\sigma"'] & MT \arrow[d, "\sigma T"] \\ T\hat{M} \arrow[r, "\hat{\lambda}"'] & \hat{M}T \end{tikzcd}$$

then $\sigma T\colon MT \Rightarrow \hat{M}T$ is a monad map. If σ is injective, then σT is as well.

Proof. We denote by η, μ the unit and multiplication of M, by $\hat{\eta}, \hat{\mu}$ those of $\hat{M}$ and by η^T, μ^T those of T. Note that $\sigma T_X = \sigma_{TX}$ and hence, using that σ is a monad map, we get immediately $\sigma T_X \circ \eta_{TX} \circ \eta^T_X = \hat{\eta}_{TX} \circ \eta^T_X$.

The following diagram commutes since σ is a monad map.

$$\begin{tikzcd} MMTX \arrow[r, "\sigma_{MTX}"] \arrow[d, "\mu"'] & \hat{M}MTX \arrow[r, "\hat{M}\sigma_{TX}"] & \hat{M}\hat{M}TX \arrow[d, "\hat{\mu}"] \\ MTX \arrow[rr, "\sigma_{TX}=\sigma T_X"] & & \hat{M}X \end{tikzcd} \tag{6}$$

From the naturality of σ, the following diagram also commutes.

$$\begin{tikzcd} MMTTX \arrow[r, "\sigma_{MTTX}"] \arrow[d, "MM\mu^T"'] & \hat{M}MTTX \arrow[r, "\hat{M}\sigma_{TTX}"] \arrow[d, "\hat{M}M\mu^T"] & \hat{M}\hat{M}TTX \arrow[d, "\hat{M}M\mu^T"] \\ MMTX \arrow[r, "\sigma_{MTX}"] & \hat{M}MTX \arrow[r, "\hat{M}\sigma_{TX}"] & \hat{M}\hat{M}TX \end{tikzcd} \tag{7}$$

Using once again the naturality of σ, for the left square, and the assumption that σ is a map of distributive laws, for the square on the right, we get the commutativity of the following diagram.

$$\begin{tikzcd} MTMTX \arrow[r, "\sigma_{TMTX}"] \arrow[d, "M\lambda_{TX}"'] & \hat{M}TMTX \arrow[r, "\hat{M}T\sigma_{TX}"] \arrow[d, "\hat{M}\lambda_{TX}"] & \hat{M}T\hat{M}TX \arrow[d, "\hat{M}\hat{\lambda}_{TX}"] \\ MMTTX \arrow[r, "\sigma_{MTTX}"] & \hat{M}MTTX \arrow[r, "\hat{M}\sigma T_{TX}"] & \hat{M}\hat{M}TTX \end{tikzcd} \tag{8}$$

Stacking diagram (8) on top of diagram (7) and further on top of diagram (6) gives the commutativity of

$$\begin{tikzcd} MTMTX \arrow[r, "\sigma T\sigma T"] \arrow[d, "\mu^{MT}"'] & \hat{M}T\hat{M}TX \arrow[d, "\mu^{\hat{M}T}"] \\ MTX \arrow[r, "\sigma T"'] & \hat{M}TX \end{tikzcd}$$

and completes the proof that σT is a monad map. Clearly, if all components of σ are injective, then all components of

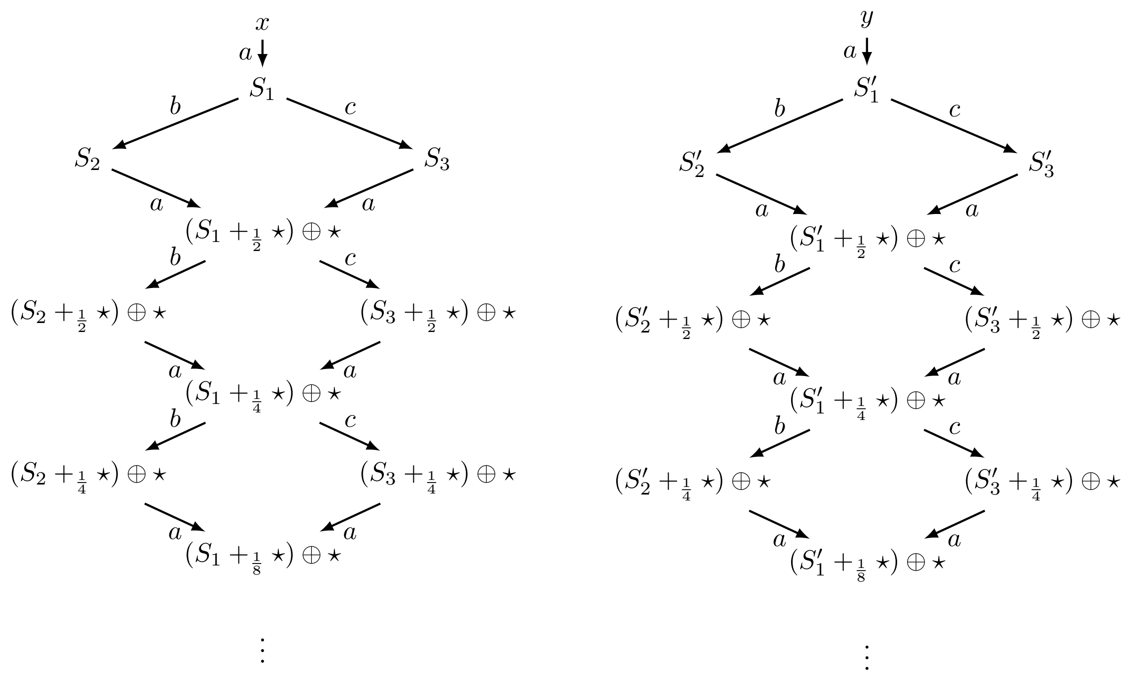


Figure 5. Determinization

$$
\begin{aligned}
x +_p y &\overset{(B_p)}{=} (x +_q \star) +_p y \\
&\overset{(A_p)}{=} x +_{pq} (\star +_{\frac{p(1-q)}{1-pq}} y) \\
&\overset{(B_p)}{=} x +_{pq} y \\
&\overset{(B_p)}{=} x +_{pq} (\star +_{\frac{q(1-p)}{1-pq}} y) \\
&\overset{(A_p)}{=} (x +_p \star) +_q y \\
&\overset{(B_p)}{=} x +_q y
\end{aligned}
$$

Figure 6. When adding (B_p) to the theory of pointed convex algebra $x +_p y = x +_q y$ holds for any $p, q \in (0, 1)$. At the monad level, adding the axioms (B_p) can be seen as the quotient of monads $\mathrm{supp}\colon \mathcal{D}(\cdot + 1) \Rightarrow \mathcal{P}$ mapping each sub-distribution into its support (e.g., $(x +_p y) +_q \star$ becomes $x + y$).

σT (which are the components of σ at TX) are injective as well. $\square$

Lemma 32. *Let M and $\hat{M}$ be two monads and $\sigma\colon M \Rightarrow \hat{M}$ be a monad map. Then the following commutes.*

$$
\begin{tikzcd}
MX + 1 \arrow[r, "\iota_X"] \arrow[d, "\sigma_X + id_1"'] & M(X+1) \arrow[d, "\sigma_{X+1}"] \\
\hat{M}X + 1 \arrow[r, "\iota_X"'] & \hat{M}(X+1)
\end{tikzcd}
$$

Proof. First observe that the following commutes: the left square commutes trivially; the right commutes since σ is a monad map.

$$
\begin{tikzcd}
1 \arrow[r, "i_r"] \arrow[d, "id_l"'] & X + 1 \arrow[r, "\eta_{X+1}"] \arrow[d, "id_{X+1}"] & M(X+1) \arrow[d, "\sigma_{X+1}"] \\
1 \arrow[r, "i_r"] & X + 1 \arrow[r, "\hat{\eta}_{X+1}"'] & \hat{M}(X+1)
\end{tikzcd}
$$

The following diagram commutes by naturality of σ.

$$
\begin{tikzcd}
MX \arrow[r, "Mi_l"] \arrow[d, "\sigma_X"'] & M(X+1) \arrow[d, "\sigma_{X+1}"] \\
\hat{M}X \arrow[r, "\hat{M}i_l"'] & \hat{M}(X+1)
\end{tikzcd}
$$

The statement of the lemma follows from the commutativity of the two above diagrams and the universal property of the coproduct. $\square$

Convex Algebras. Another presentation of convex algebras is given by the algebraic theory with infinitely many operations denoting arbitrary (and not only binary) convex combinations $(\Sigma_{\hat{P}}, E_{\hat{P}})$ where $\Sigma_{\hat{P}}$ consists of operations $\sum_{i=1}^{n} p_i(\cdot)_i$ for all $n \in \mathbb{N}$ and $(p_1, \ldots, p_n) \in [0,1]^n$ such that $\sum_{i=1}^{n} p_i = 1$ and $E_{\hat{P}}$ is the set of the following two axioms.

$$\sum_{i=0}^{n} p_i x_i \stackrel{(P)}{=} x_j \qquad \text{if } p_j = 1$$

$$\sum_{i=0}^{n} p_i \left(\sum_{j=0}^{m} q_{i,j} x_j \right) \stackrel{(BC)}{=} \sum_{j=0}^{m} \left(\sum_{i=0}^{n} p_i q_{i,j} \right) x_j .$$

Here, (P) stands for projection, and (BC) for barycentre.

Convex algebras are known under many names: "convex modules" in [64], "positive convex structures" in [43] (where X is taken to be endowed with the discrete topology), "sets with a convex structure" in [41], and barycentric algebras [65].

Remark 33. *Let $\mathbb{X}$ be a $(\Sigma_{\hat{P}}, E_{\hat{P}})$-algebra. Then (for $p_n \neq 1$ and $\overline{p_n} = 1 - p_n$)*

$$\sum_{i=1}^{n} p_i x_i = \overline{p_n} \left(\sum_{j=1}^{n-1} \frac{p_j}{\overline{p_n}} x_j \right) + p_n x_n . \tag{9}$$

Hence, an n-ary convex combination can be written as a binary convex combination using an $(n-1)$-ary convex combination.

One can also see Equation (9) as a definition – the classical definition of Stone [65, Definition 1]. The following property, whose proof follows by induction along the lines of [65, Lemma 1–Lemma 4], gives the connection:

Let X be the carrier of a (Σ_P, E_P)-algebra. Define n-ary convex operations inductively by the projection axiom and the formula (9). Then X becomes an algebra in $(\Sigma_{\hat{P}}, E_{\hat{P}})$.

Appendix D. Proofs for Section 3, the Presentation for C

of Lemma 3. For $n = 1$ the property amounts to idempotence. Assume $n > 1$ and the property holds for $n - 1$.

Below, we will write (D) also for generalised distributivity as in

$$\bigoplus_i a_i +_p \bigoplus_j b_j \stackrel{(D)}{=} \bigoplus_{i,j} (a_i +_p b_j).$$

First, we observe that

$$a_1 \oplus \ldots \oplus a_n = a_1 \oplus \ldots \oplus a_n \oplus \bigoplus_i (a_i +_{p_1} \bigoplus_{j \neq i} a_j) \tag{10}$$

which follows from

$$\begin{aligned} a_1 \oplus \ldots \oplus a_n &\stackrel{(I_p)}{=} (a_1 \oplus \ldots \oplus a_n) +_{p_1} (a_1 \oplus \ldots \oplus a_n) \\ &\stackrel{(D)}{=} \bigoplus_{i,j} (a_i +_{p_1} a_j) \\ &\stackrel{(I_p, D)}{=} a_1 \oplus \ldots \oplus a_n \oplus \bigoplus_i (a_i +_{p_1} \bigoplus_{j \neq i} a_j) \end{aligned}$$

Recall that we write $\overline{p}$ for $1 - p$ if $p \in [0, 1]$. Furthermore, having in mind that $\sum_{i=1}^{n} p_i a_i = a_1 +_{p_1} (\sum_{i=2}^{n} \frac{p_i}{\overline{p_1}} a_i)$ we have

$$\begin{aligned} a_1 +_{p_1} (\bigoplus_{j \neq 1} a_j) &\stackrel{IH}{=} a_1 +_{p_1} (\bigoplus_{j \neq 1} a_j \oplus \sum_{j \neq 1} \frac{p_j}{\overline{p_1}} a_j) \\ &\stackrel{(D)}{=} (a_1 +_{p_1} \bigoplus_{j \neq 1} a_j) + (a_1 +_{p_1} \sum_{j \neq 1} \frac{p_j}{\overline{p_1}} a_j) \\ &= (a_1 +_{p_1} \bigoplus_{j \neq 1} a_j) + \sum_i p_i a_i . \end{aligned}$$

Using this in the second equality below, we get

$$\begin{aligned} & a_1 \oplus \ldots \oplus a_n \oplus \sum_{i=1}^{n} p_i a_i \\ &\stackrel{Eq.\ (10)}{=} a_1 \oplus \ldots \oplus a_n \oplus \bigoplus_i (a_i +_{p_1} (\bigoplus_{j \neq i} a_j) \oplus \sum_{i=1}^{n} p_i a_i \\ &= a_1 \oplus \ldots \oplus a_n \oplus \bigoplus_i (a_i +_{p_1} (\bigoplus_{j \neq i} a_j) \\ &\stackrel{Eq.\ (10)}{=} a_1 \oplus \ldots \oplus a_n . \end{aligned}$$

□

We next formulate a property that provides a way to prove that an algebraic theory is a presentation for a monad, which we use in the proof that (Σ_{NP}, E_{NP}) is a presentation for C.

Proposition 34. *Let $\mathcal{V}$ be the variety of (Σ, E)-algebras with signature Σ and equations E. Let $U \colon \mathcal{V} \to \mathbf{Sets}$ be the forgetful functor. In order to prove that (Σ, E) is a presentation for a monad (M, η, μ), it suffices to:*

1. *For any set X, define Σ-operations Σ_X on MX and prove that with these operations (MX, Σ_X) is an algebra in $\mathcal{V}$. Moreover prove that for any map $f \colon X \to Y$, Mf is a $\mathcal{V}$-homomorphism from (MX, Σ_X) to (MY, Σ_Y).*
2. *Prove that (MX, Σ_X) is the free algebra in $\mathcal{V}$ with basis $\eta(X)$, i.e., for any algebra $\mathbb{A} = (A, \Sigma_A)$ in $\mathcal{V}$ and any map $f \colon X \to A$, there is a unique homomorphism $f^{\#} \colon (MX, \Sigma_X) \to \mathbb{A}$ that extends f, i.e., that satisfies $f = Uf^{\#} \circ \eta$.*
3. *Prove that $\mu_X = (\mathit{id}_{MX})^{\#}$.*

Proof. Assume that 1.-3. hold. Let $F \colon \mathbf{Sets} \to \mathcal{V}$ be the functor defined on objects as $FX = (MX, \Sigma_X)$. This shows that $UFX = MX$.

On arrows $f \colon X \to Y$, we set $Ff = (\eta \circ f)^{\#}$. Then F is a left adjoint of the forgetful functor U, and the adjunction is given by the bijective correspondence $(f \colon X \to U\mathbb{A}) \mapsto (f^{\#} \colon FX \to \mathbb{A})$.

Next, we see that $UFf = Mf$ as a consequence of naturality of η. Namely, we have that $Ff = (\eta \circ f)^{\#}$ is the unique homomorphism with the property $UFf \circ \eta = \eta \circ f$. Hence, using 1., since $Mf \circ \eta = \eta \circ f$ by naturality of η, we get $UFf = Mf$.

Let $(T, \bar{\eta}, \bar{\mu})$ be the monad of this adjunction. Then, see e.g. [66], $\bar{\eta}^{\#} = \mathit{id}_{FX}$. We next show that $\eta^{\#} = \mathit{id}_{FX}$ which implies $\bar{\eta} = \eta$. All we need to observe is that $U\mathit{id}_{FX} \circ \eta = \mathit{id}_{UFX} \circ \eta = \eta$ and since $\eta^{\#}$ is the unique homomorphism with $U\eta^{\#} \circ \eta = \eta$ and id_{FX} is a homomorphism from FX to itself, we get $\eta^{\#} = \mathit{id}_{FX}$. Finally, see e.g. [66], $\bar{\mu}_X = (\mathit{id}_{MX})^{\#}$ so item 3. proves that $\bar{\mu} = \mu$. □

Before we proceed with the proof of the presentation, we recall several properties that are known or immediate to check, but very helpful in our further proofs.

Lemma 35. *Let X be a set and $S \in CCX$. Then $\bigcup S \in CX$.*

Proof. Let $S = \{S_i \mid i \in I\}$. Let $\Phi, \Psi \in \bigcup S$. Then there exist $i, j \in I$ with $\Phi \in S_i$ and $\Psi \in S_j$. We have $p\Phi + \overline{p}\Psi \in pS_i + \overline{p}S_j \in S$ as S is convex. □

Lemma 36. *Let $\mathbb{A}$ and $\mathbb{B}$ be two convex algebras, and $f\colon \mathbb{A} \to \mathbb{B}$ a convex homomorphism. Then the image map $\overline{f} = \mathcal{P}_u f\colon \mathcal{P}_u A \to \mathcal{P}_u B$, for $\mathcal{P}_u$ being the unrestricted (not necessarily finite) powerset, is a convex map, i.e. if $S = X +_p Y$ for $X \in \mathcal{P}_u A, Y \in \mathcal{P}_u B$, then $\overline{f}(S) = \overline{f}(X) +_p \overline{f}(Y)$.*

Proof. Let $S = X +_p Y$ for $X \in \mathcal{P}_u A, Y \in \mathcal{P}_u B$. Then

$$\begin{aligned} \overline{f}(S) &= \{f(s) \mid s \in S\} \\ &= \{f(px + \overline{p}y) \mid x \in X, y \in Y\} \\ &\overset{(*)}{=} \{pf(x) + \overline{p}f(y) \mid x \in X, y \in Y\} \\ &= p\overline{f}(X) + \overline{p}\overline{f}(Y). \end{aligned}$$

where the equality marked by $(*)$ holds by the assumption that f is a convex homomorphism. □

Lemma 37. *Let $\mathbb{A}$ and $\mathbb{B}$ be two convex algebras, and $f\colon \mathbb{A} \to \mathbb{B}$ a convex homomorphism. Then for all $X \in \mathcal{P}_u A$, for $\mathcal{P}_u$ being the unrestricted (not necessarily finite) powerset, $\mathrm{conv}_{\mathbb{B}}\, \overline{f}(X) = \overline{f}(\mathrm{conv}_{\mathbb{A}}\, X)$. In particular, if X is convex then also $\overline{f}(X)$ is convex.*

Proof. For $\subseteq$, for an arbitrary $pf(x) + \overline{p}f(y) \in \mathrm{conv}_{\mathbb{B}}\, \overline{f}(X)$ with $x, y \in X$, we have

$$pf(x) + \overline{p}f(y) \overset{(*)}{=} f(px + \overline{p}y) \in \overline{f}(\mathrm{conv}_{\mathbb{A}}\, X)$$

and here, again, $(*)$ holds since f is convex. For $\supseteq$, consider $f(a) \in \overline{f}(\mathrm{conv}_{\mathbb{A}}\, X)$. Then $a = px + \overline{p}y$ for some $x, y \in X$. Since f is convex, $f(a) = pf(x) + \overline{p}f(y)$ and $f(x), f(y) \in \overline{f}(X)$. Hence $f(a) \in \mathrm{conv}_{\mathbb{B}}\, \overline{f}(X)$. □

The proof of the presentation follows the structure of Proposition 34 via the following three lemmas.

Lemma 38. *With the above defined operations $(CX, \oplus, +_p)$ is a convex semilattice, for any set X. Moreover, for a map $f\colon X \to Y$, the map $Cf\colon CX \to CY$ is a convex semilattice homomorphism from $(CX, \oplus, +_p)$ to $(CY, \oplus, +_p)$.*

of Lemma 38. In any convex algebra $\mathbb{A}$ for $S, T \subseteq A$ we have

$$\mathrm{conv}(\mathrm{conv}(S) \cup T) = \mathrm{conv}(S \cup T).$$

As a consequence, using the associativity of union, we get that the axiom (A) holds. For $S_1, S_2, S_3 \in CX$:

$$\begin{aligned} S_1 \oplus (S_2 \oplus S_3) &= \mathrm{conv}(S_1 \cup \mathrm{conv}(S_2 \cup S_3)) \\ &= \mathrm{conv}(S_1 \cup (S_2 \cup S_3)) \\ &= \mathrm{conv}((S_1 \cup S_2) \cup S_3) \\ &= \mathrm{conv}(\mathrm{conv}(S_1 \cup S_2) \cup S_3) \\ &= (S_1 \oplus S_2) \oplus S_3. \end{aligned}$$

Commutativity and idempotence hold due to commutativity and idempotence of union.

Defining convex operations on CX using Minkowski sum, see [67], leads to a convex algebra, i.e., $(A_p), (C_p), (I_p)$ hold.

The axiom (D) holds as:

$$\begin{aligned} &(S_1 \oplus S_2) +_p S_3 \\ &\qquad = p\,\mathrm{conv}(S_1 \cup S_2) + \overline{p}S_3 = \\ &\qquad = \{pqd_1 + p\overline{q}d_2 + \overline{p}d_3 \mid q \in [0,1], d_i \in S_i\} \\ &\qquad = \mathrm{conv}((pS_1 + \overline{p}S_3) \cup (pS_2 + \overline{p}S_3)). \end{aligned}$$

Finally, Cf is a homomorphism from $(CX, \oplus, +_p)$ to $(CY, \oplus, +_p)$ as

$$\begin{aligned} Cf(S_1 \oplus S_2) &= \overline{\mathcal{D}f}(S_1 \oplus S_2) \\ &\overset{(a)}{=} \mathrm{conv}(\overline{\mathcal{D}f}(S_1 \cup S_2)) \\ &= \mathrm{conv}(\overline{\mathcal{D}f}(S_1) \cup \overline{\mathcal{D}f}(S_2))) \\ &= \overline{\mathcal{D}f}(S_1) \oplus \overline{\mathcal{D}f}(S_2) \\ &= Cf(S_1) \oplus Cf(S_2) \end{aligned}$$

where the equality marked by (a) holds by Lemma 37. Similarly

$$\begin{aligned} Cf(S_1 +_p S_2) &= \overline{\mathcal{D}f}(S_1 +_p S_2) \\ &\overset{(b)}{=} \overline{\mathcal{D}f}(S_1) +_p \overline{\mathcal{D}f}(S_2) \\ &= Cf(S_1) +_p Cf(S_2). \end{aligned}$$

where the equality marked by (b) holds by Lemma 36. □

Lemma 39. *The convex semilattice $(CX, \oplus, +_p)$ is the free convex semilattice generated by $\eta(X)$.*

Proof. We need to show that for any map $f\colon X \to A$ for a convex semilattice $\mathbb{A} = (A, \oplus, +_p)$, there is a unique convex semilattice homomorphism $f^{\#}\colon (CX, \oplus, +_p) \to \mathbb{A}$ such that $Uf^{\#} \circ \eta = f$. So, let $\mathbb{A} = (A, \oplus, +_p)$ be a convex semilattice, and let $f\colon X \to A$ be a map. We use the same notation for the operations in A and in CX for simplicity.

Note that, since any convex semilattice is a convex algebra, there is a unique convex homomorphism $f^{\#}_{\mathcal{D}}\colon \mathcal{D}X \to (A, +_p)$, as $\mathcal{D}X$ is the free convex algebra generated by $\eta_{\mathcal{D}}(X)$. Hence, $Uf^{\#}_{\mathcal{D}} \circ \eta_{\mathcal{D}} = f$.

Now, given a convex set $S = \mathrm{conv}\{d_1, \ldots, d_n\} \in CX$ we put

$$f^{\#}(S) = f^{\#}_{\mathcal{D}}(d_1) \oplus f^{\#}_{\mathcal{D}}(d_2) \oplus \ldots \oplus f^{\#}_{\mathcal{D}}(d_n).$$

We first prove that $f^{\#}$ is well defined, which is the most important step. We show that whenever

$$\operatorname{conv}\{d_1, \ldots, d_n\} = \operatorname{conv}\{e_1, \ldots, e_m\} \tag{11}$$

then

$$f^{\#}_{\mathcal{D}}(d_1) \oplus \ldots \oplus f^{\#}_{\mathcal{D}}(d_n) = f^{\#}_{\mathcal{D}}(e_1) \oplus \ldots \oplus f^{\#}_{\mathcal{D}}(e_m).$$

Clearly, if Equation (11) holds, then for all $i \in \{1, \ldots, n\}$, $d_i \in \operatorname{conv}\{e_1, \ldots, e_m\}$ and for all $j \in \{1, \ldots, m\}$, $e_j \in \operatorname{conv}\{d_1, \ldots, d_n\}$. Hence,

$$\begin{aligned}\operatorname{conv}\{d_1, \ldots, d_n, e_1, \ldots, e_m\} &= \operatorname{conv}\{d_1, \ldots, d_n\} \\ &= \operatorname{conv}\{e_1, \ldots, e_n\}.\end{aligned}$$

If we can prove that whenever $e \in \operatorname{conv}\{d_1, \ldots, d_n\}$ then

$$f^{\#}_{\mathcal{D}}(d_1) \oplus \ldots \oplus f^{\#}_{\mathcal{D}}(d_n) \oplus f^{\#}_{\mathcal{D}}(e) = f^{\#}_{\mathcal{D}}(d_1) \oplus \ldots \oplus f^{\#}_{\mathcal{D}}(d_n),$$

we would be done with well defined-ness as then

$$\begin{aligned}& f^{\#}_{\mathcal{D}}(d_1) \oplus \ldots \oplus f^{\#}_{\mathcal{D}}(d_n) \\ = \quad & f^{\#}_{\mathcal{D}}(d_1) \oplus \ldots \oplus f^{\#}_{\mathcal{D}}(d_n) \oplus f^{\#}_{\mathcal{D}}(e_1) \oplus \ldots \oplus f^{\#}_{\mathcal{D}}(e_m) \\ = \quad & f^{\#}_{\mathcal{D}}(e_1) \oplus \ldots \oplus f^{\#}_{\mathcal{D}}(e_m).\end{aligned}$$

So, let $e \in \operatorname{conv}\{d_1, \ldots, d_n\}$. Then $e = \sum_i p_i d_i$ and since $f^{\#}_{\mathcal{D}}$ is a convex algebra homomorphism, $f^{\#}_{\mathcal{D}}(e) = \sum_i p_i f^{\#}_{\mathcal{D}}(d_i)$. Now, by the convexity law, Lemma 3, we have that for any $a_1, \ldots, a_n \in A$ and any $p_1, \ldots, p_n \in [0, 1]$ with $\sum_i p_i = 1$

$$a_1 \oplus \ldots \oplus a_n \oplus \sum_i p_i a_i \quad = \quad a_1 \oplus \ldots \oplus a_n.$$

Hence, indeed

$$f^{\#}_{\mathcal{D}}(d_1) \oplus \ldots \oplus f^{\#}_{\mathcal{D}}(d_n) \oplus f^{\#}_{\mathcal{D}}(e) = f^{\#}_{\mathcal{D}}(d_1) \oplus \ldots \oplus f^{\#}_{\mathcal{D}}(d_n).$$

It remains to show that $f^{\#}$ is a homomorphism and that it is uniquely extending $\eta(X)$. Let $S, T \in CX$. Let $S = \operatorname{conv}\{d_1, \ldots, d_n\}$, $T = \operatorname{conv}\{e_1, \ldots, e_m\}$.

Then $S \oplus T = \operatorname{conv}(S \cup T) = \operatorname{conv}\{d_1, \ldots, d_n, e_1, \ldots, e_m\}$ and we get

$$\begin{aligned}& f^{\#}(S \oplus T) \\ &= f^{\#}_{\mathcal{D}}(d_1) \oplus \ldots \oplus f^{\#}_{\mathcal{D}}(d_n) \oplus f^{\#}_{\mathcal{D}}(e_1) \oplus \ldots \oplus f^{\#}_{\mathcal{D}}(e_m) \\ &= f^{\#}(S) \oplus f^{\#}(T).\end{aligned}$$

Next, we first notice that $S +_p T = \operatorname{conv}\{pd_i + \overline{p}e_j \mid i \in \{1, \ldots, n\}, j \in \{1, \ldots, m\}\}$. For $\supseteq$, we see that

$$\sum_{i,j} q_{i,j}(pd_i + \overline{p}e_j) = p \sum_{i,j} q_{i,j} d_i + \overline{p} \sum_{i,j} q_{i,j} e_j \in S +_p T.$$

For $\subseteq$, take $pd + \overline{p}e \in S +_p T$. So, $d = \sum_i q_i d_i$ and $e = \sum_j r_j e_j$ and we have

$$\begin{aligned}pd + \overline{p}e &= p \sum_i q_i d_i + \overline{p} \sum_j r_j e_j \\ &= p \sum_i q_i \left(\sum_j r_j\right) d_i + \overline{p} \sum_j r_j \left(\sum_i q_i\right) e_j \\ &= \sum_{i,j} q_i r_j (pd_i + \overline{p}e_j).\end{aligned}$$

Now

$$\begin{aligned}& f^{\#}(S +_p T) \\ &= \bigoplus_{i,j} f^{\#}_{\mathcal{D}}(pd_i + \overline{p}e_j) \\ &= \bigoplus_{i,j} p f^{\#}_{\mathcal{D}}(d_i) + \overline{p} f^{\#}_{\mathcal{D}}(e_j) \\ &= \bigoplus_{i,j} f^{\#}_{\mathcal{D}}(d_i) +_p f^{\#}_{\mathcal{D}}(e_j) \\ &\overset{(D)}{=} (f^{\#}_{\mathcal{D}}(d_1) \oplus \ldots \oplus f^{\#}_{\mathcal{D}}(d_n)) +_p (f^{\#}_{\mathcal{D}}(e_1) \oplus \ldots \oplus f^{\#}_{\mathcal{D}}(e_m)) \\ &= f^{\#}(S) +_p f^{\#}(T).\end{aligned}$$

Finally, assume $f^*\colon (CX, \oplus, +_p) \to \mathbb{A}$ is another homomorphism that extends f on $\eta(X)$, i.e., such that $Uf^* \circ \eta = f$. Then $f^*(\{\delta_x\}) = f^{\#}(\{\delta_x\}) = f(x)$. Since both $f^{\#}$ and f^* are convex homomorphisms, and $\{\sum_i p_i x_i\} = \sum_i p_i\{\delta_{x_i}\}$, we get

$$\begin{aligned}f^*(\{\sum_i p_i x_i\}) &= \sum_i p_i f^*(\{\delta_{x_i}\}) \\ &= \sum_i p_i f^{\#}(\{\delta_{x_i}\}) \quad = f^{\#}(\{\sum_i p_i x_i\}).\end{aligned}$$

Further on, for $S = \operatorname{conv}\{d_1, \ldots, d_n\}$ we have $S = \{d_1\} \oplus \ldots \oplus \{d_n\}$ and hence $f^*(S) = f^*(\{d_1\}) \oplus \ldots \oplus f^*(\{d_n\}) = f^{\#}(\{d_1\}) \oplus \ldots \oplus f^{\#}(\{d_n\}) = f^{\#}(S)$ shows that $f^* = f^{\#}$ and completes the proof. ☐

The final missing property for the presentation, Lemma 41, is an easy consequence of the next property that clarifies the definition of $f^{\#}$.

Lemma 40. *Let X be a set and $f\colon X \to CY$ a map. Then for all S in CX*

$$f^{\#}(S) = \bigcup \overline{f^{\#}_{\mathcal{D}}}(S) = \bigcup_{\Phi \in S} \sum_{u \in \operatorname{supp} \Phi} \Phi(u) \cdot f(u).$$

Proof. The first task is to prove that $f^{\#}(S) = \bigcup \overline{f^{\#}_{\mathcal{D}}}(S)$. Before we proceed, let's recall all the types. We have $f\colon X \to CY$ (and CY is the carrier of a convex semilattice), so $f^{\#}\colon CX \to CY$. Also, $f^{\#}_{\mathcal{D}}\colon \mathcal{D}X \to CY$ and hence $\overline{f^{\#}_{\mathcal{D}}}\colon \mathcal{P}_u\mathcal{D}X \to \mathcal{P}_u CY$ for $\mathcal{P}_u$ denoting the unrestricted (and not just finite) powerset. Finally, here $\bigcup\colon \mathcal{P}_u\mathcal{P}_u\mathcal{D}Y \to \mathcal{P}_u\mathcal{D}Y$. Clearly, $CZ \subseteq \mathcal{P}_u\mathcal{D}Z$ for any set Z.

Now, since S is convex, by Lemma 37 also $\overline{f^{\#}_{\mathcal{D}}}(S)$ is convex. Each element of $\overline{f^{\#}_{\mathcal{D}}}(S)$ is of the form $f^{\#}_{\mathcal{D}}(\Phi)$ for $\Phi \in S$ and hence it is in CY, i.e., is convex. By Lemma 35, we get that $\bigcup \overline{f^{\#}_{\mathcal{D}}}(S)$ is convex.

Let $\Psi_1, \ldots, \Psi_n \in \mathcal{D}X$ be such that $S = \operatorname{conv}\{\Psi_1, \ldots, \Psi_n\}$. Clearly, $\Psi_1, \ldots, \Psi_n \in S$. Now, we have

$$\{f^{\#}_{\mathcal{D}}(\Psi_i) \mid i = 1, \ldots, n\} \subseteq \{f^{\#}_{\mathcal{D}}(\Phi) \mid \Phi \in S\}$$

and hence

$$\bigcup\{f^{\#}_{\mathcal{D}}(\Psi_i) \mid i=1,\ldots,n\} \subseteq \bigcup\{f^{\#}_{\mathcal{D}}(\Phi) \mid \Phi \in S\} \\ = \bigcup \overline{f^{\#}_{\mathcal{D}}}(S)$$

and since the set on the right hand side is convex, as we noted above,

$$f^{\#}(S) = \operatorname{conv}\bigcup\{f^{\#}_{\mathcal{D}}(\Psi_i) \mid i=1,\ldots,n\} \\ \subseteq \bigcup\{f^{\#}_{\mathcal{D}}(\Phi) \mid \Phi \in S\},$$

where the first equality is simply the definition of $f^{\#}$.

For the other inclusion, let $\Phi \in S$. Then $S = \operatorname{conv}\{\Psi_1,\ldots,\Psi_n,\Phi\}$ and

$$f^{\#}(S) = \operatorname{conv}\bigcup\{f^{\#}_{\mathcal{D}}(\Psi_1),\ldots,f^{\#}_{\mathcal{D}}(\Psi_n),f^{\#}_{\mathcal{D}}(\Phi)\}$$

by the definition of $f^{\#}$. Therefore, $f^{\#}_{\mathcal{D}}(\Phi) \subseteq f^{\#}(S)$ and since Φ was arbitrary,

$$\bigcup\{f^{\#}_{\mathcal{D}}(\Phi) \mid \Phi \in S\} \subseteq f^{\#}(S).$$

This proves the first equality of our statement. For the second equality, note that

$$\begin{aligned} f^{\#}(S) &= \bigcup\{f^{\#}_{\mathcal{D}}(\Phi) \mid \Phi \in S\} \\ &\overset{(*)}{=} \bigcup\{\sum_{u\in\operatorname{supp}\Phi} \Phi(u)\cdot f(u) \mid \Phi \in S\} \\ &= \bigcup_{\Phi\in S}\sum_{u\in\operatorname{supp}\Phi} \Phi(u)\cdot f(u) \end{aligned}$$

where the equality $(*)$ holds as $f^{\#}_{\mathcal{D}}$ is convex. ☐

Lemma 41. *The multiplication μ of the monad C satisfies $\mu = (id_{CX})^{\#}$.*

Proof. Using Lemma 40, we immediately get

$$(id_{CX})^{\#}(S) = \bigcup_{\Phi\in S}\sum_{A\in\operatorname{supp}\Phi} \Phi(A)\cdot A = \mu_X(S).$$

☐

Appendix E. Proofs of Section 4

Proof. (of Proposition 8) We denote by $\{\bullet\}$ the generating set. Let $2 = \{\bullet,\star\}$. Note that the carrier of the free pointed semilattice generated by $\{\bullet\}$ is $C(1+1) = C(2)$. Recall that $(C(2),\oplus,+_p)$, where $\oplus$ is the convex union and $+_p$ is the Minkowski sum, is the free convex semilattice generated by 2.

We first show that $(C(2),\oplus,+_p)$ is isomorphic to $\mathbb{M}_{\mathcal{I}}$. Indeed $\mathcal{D}(2)$ is isomorphic to $[0,1]$: the real number 0 corresponds to $\delta_\star$, 1 to $\delta_\bullet$ and $p\in(0,1)$ to $\bullet +_p \star$. Furthermore, the non-empty finitely-generated convex subsets of $[0,1]$ are the closed intervals. To conclude, it suffices to see that min-max is $\oplus$ on $\mathcal{I}$ and $+^{\mathcal{I}}_p$ is the Minkowski sum. ☐

Proof. (of Proposition 9) By Proposition 8, we know that $C(2)$ is isomorphic to $\mathcal{I}$. We show that $\mathbb{M}_{\mathcal{I},[0,0]}$ modulo the axiom (B) is isomorphic to $\mathbb{M}\mathrm{ax}_B$. We have

$$\text{min-max}([x,y],[0,0]) \overset{(B)}{=} [x,y],$$

for $[x,y]\in\mathcal{I}$. From

$$[0,y] = \text{min-max}([x,y],[0,0]) = [x,y],$$

we derive that $[x_1,y] = [x_2,y]$ for any x_1,x_2,y. Hence, we define the isomorphism $[x,y]\mapsto y$ mapping any interval $[x,y]$ to its upper bound y.

The interval $[0,0]$ is mapped to the bottom element 0, and the operations are such that:

$$\text{min-max}([x_1,y_1],[x_2,y_2]) = \text{min-max}([0,y_1],[0,y_2]) \\ = [0,\max(y_1,y_2)]$$

hence $\text{min-max}([x_1,y_1],[x_2,y_2]) \mapsto \max(y_1,y_2)$ and

$$[x_1,y_1]+_p[x_2,y_2] = [0,y_1]+_p[0,y_2] = [0,y_1+_py_2] \mapsto y_1+_py_2.$$

☐

Proof. (of Proposition 10) We show that $\mathbb{M}_{\mathcal{I},[0,0]}$ modulo the (T) axiom

$$\text{min-max}([x,y],[0,0]) \overset{(T)}{=} [0,0]$$

is isomorphic to $\mathbb{M}\mathrm{in}_T$. First, we derive $[x,y_1]=[x,y_2]$ for any x,y_1,y_2 as follows. For $x=1$ the property trivially holds. For $x=0$ we have

$$[0,y_1] = \text{min-max}([x,y_1],[0,0]) \\ \overset{(T)}{=} [0,0] \overset{(T)}{=} \text{min-max}([x,y_2],[0,0]) = [0,y_2] \quad (*)$$

Finally, for $x\in(0,1)$ and $y_1,y_2\geq x$ we derive

$$[x,y_1] = [1,1]+_x[0,\frac{y_1-x}{1-x}] \overset{(*)}{=} [1,1]+_x[0,\frac{y_2-x}{1-x}] = [x,y_2].$$

Hence, we can now map every interval $[x,y]$ to its lower bound x. Then $[0,0]$ is mapped to the top element 0, and

$$\text{min-max}([x_1,y_1],[x_2,y_2]) = [\min(x_1,x_2),\max(y_1,y_2)] \\ \mapsto \min(x_1,x_2)$$

$$[x_1,y_1]+_p[x_2,y_2] = [x_1+_px_2, y_1+_py_2] \mapsto x_1+_px_2.$$

☐

Appendix F. Proofs of Section 5

Proof. (of Corollary 17) We first transform the automaton $c = \langle o,t\rangle$ to $\hat{c} = \langle\hat{o},t\rangle$ and apply Theorem 16.2 and then transform $\hat{c} = \langle\hat{o},t\rangle$ to $\hat{\hat{c}} = \langle\hat{o},\hat{t}\rangle$ and apply Theorem 16.1.

For the determinisation, take for a in Theorem 16.2 the free algebra $\mu_1\colon MM1 \to M1$ and as $\hat{a}$ the M-algebra $(\hat{\mu}\circ\sigma_{\hat{M}1})\colon M\hat{M}1 \to \hat{M}1$. It is easy to see that $\hat{a}$ is indeed an M-algebra using that σ is a monad map, its naturality, and the associativity of $\hat{\mu}$. Since σ is a monad map, the following diagram commutes showing that σ_1 is an M-algebra homomorphism.

$$\begin{tikzcd}
MM1 \arrow[r, "M\sigma_1"] \arrow[dd, "\mu"'] & M\hat{M}1 \arrow[d, "\sigma_{\hat{M}1}"] \\
& \hat{M}\hat{M}1 \arrow[d, "\hat{\mu}"] \\
M1 \arrow[r, "\sigma_1"] & \hat{M}1
\end{tikzcd}$$

Observe that $\hat{o} = \sigma_1 \circ o$, again since σ is a monad map. Then, by Theorem 16.2 $\hat{[\![\cdot]\!]} = \sigma_1^{A^*} \circ [\![\cdot]\!]$ where $\hat{[\![\cdot]\!]}$ is the semantics obtained by determinisation of $\hat{c}$ and $[\![\cdot]\!]$ the one after determinisation of c. Since σ_1 is injective, also $\sigma_1^{A^*}$ is injective and we have that for all $x, y \in X$, $[\![\widehat{\eta(x)}]\!] = [\![\widehat{\eta(y)}]\!]$ iff $[\![\eta(x)]\!] = [\![\eta(y)]\!]$, i.e., the semantics remains the same.

For the second step, take $\hat{\hat{a}} = \hat{\mu}\colon \hat{M}\hat{M}1 \to \hat{M}1$ for the determinisation of $\hat{\hat{c}}$. By definition $\hat{a} = \hat{\hat{a}} \circ \sigma_{\hat{O}}$. Therefore Theorem 16.1 guarantees that the semantics after determinisation of $\hat{\hat{c}}$ again remains the same as the semantics after determinisation of $\hat{c}$. □

Proof. (of Theorem 16.1) The proof proceeds in two steps. First, we show that the following diagram commutes

$$\begin{tikzcd}
MFX \arrow[r, "\sigma_{FX}"] \arrow[d, "\lambda_X"'] & \hat{M}FX \arrow[d, "\hat{\lambda}_X"] \\
FMX \arrow[r, "F\sigma_X"'] & F\hat{M}X
\end{tikzcd} \tag{12}$$

where σ is the monad map from the hypothesis, and λ and $\hat{\lambda}$ are the distributive laws from Proposition 11 used for the determinisation of FM- and $F\hat{M}$-coalgebras using the algebras a and $\hat{a}$, and the strengths *st* and $\hat{st}$, respectively.

The following diagram commutes by naturality of σ:

$$\begin{tikzcd}
M(X^A) \arrow[r, "\sigma_{X^A}"] \arrow[d, "M\mathit{ev}_a"'] & \hat{M}(X^A) \arrow[d, "\hat{M}\mathit{ev}_a"] \\
MX \arrow[r, "\sigma_X"'] & \hat{M}X
\end{tikzcd}$$

Using this, by definition of the strengths, we have

$$\begin{aligned}
(\sigma_X^A \circ \mathit{st}(\varphi))(a) &= \sigma_X^A(\mathit{st}(\varphi)(a)) \\
&= \sigma_X^A(M\mathit{ev}_a(\varphi)) \\
&= (\sigma_X^A \circ M\mathit{ev}_a)(\varphi) \\
&\overset{(*)}{=} \hat{M}\mathit{ev}_a \circ \sigma_{X^A}(\varphi) \\
&= \hat{M}\mathit{ev}_a(\sigma_{X^A}(\varphi)) \\
&= \hat{\mathit{st}}(\sigma_{X^A}(\varphi))(a) \\
&= (\hat{\mathit{st}} \circ \sigma_{X^A}(\varphi))(a).
\end{aligned}$$

where the equality marked by $(*)$ holds by the commutativity of the diagram above. Hence, the following diagram commutes.

$$\begin{tikzcd}
M(X^A) \arrow[r, "\sigma_{X^A}"] \arrow[d, "\mathit{st}"'] & \hat{M}(X^A) \arrow[d, "\hat{\mathit{st}}"] \\
(MX)^A \arrow[r, "\sigma_X^A"'] & (\hat{M}X)^A
\end{tikzcd}$$

Recall now that by hypothesis $a = \hat{a} \circ \sigma_O$. Therefore, the following commutes.

$$\begin{tikzcd}
MO \times M(X^A) \arrow[r, "\sigma_O \times \sigma_{X^A}"] \arrow[d, "a \times \mathit{st}"'] & \hat{M}O \times \hat{M}(X^A) \arrow[d, "\hat{a} \times \hat{\mathit{st}}"] \\
O \times (MX)^A \arrow[r, "id_O \times \sigma_X^A"'] & O \times (\hat{M}X)^A
\end{tikzcd}$$

Finally, the following two squares commute by naturality of σ.

$$\begin{tikzcd}
M(O \times X^A) \arrow[r, "\sigma_{O \times X^A}"] \arrow[d, "M\pi_1"'] & \hat{M}(O \times X^A) \arrow[d, "\hat{M}\pi_1"] \\
MO \arrow[r, "\sigma_O"'] & \hat{M}O
\end{tikzcd}$$

$$\begin{tikzcd}
M(O \times X^A) \arrow[r, "\sigma_{O \times X^A}"] \arrow[d, "M\pi_2"'] & \hat{M}(O \times X^A) \arrow[d, "\hat{M}\pi_2"] \\
M(X^A) \arrow[r, "\sigma_{X^A}"'] & \hat{M}(X^A)
\end{tikzcd}$$

By pasting together the last three diagrams, we obtain that the following commutes.

$$\begin{tikzcd}
M(O \times X^A) \arrow[r, "\sigma_{O \times X^A}"] \arrow[d, "\langle M\pi_1, M\pi_2 \rangle"'] & \hat{M}(O \times X^A) \arrow[d, "\langle \hat{M}\pi_1, \hat{M}\pi_2 \rangle"] \\
MO \times M(X^A) \arrow[r, "\sigma_O \times \sigma_{X^A}"] \arrow[d, "a \times \mathit{st}"'] & \hat{M}O \times \hat{M}(X^A) \arrow[d, "\hat{a} \times \hat{\mathit{st}}"] \\
O \times M(X)^A \arrow[r, "id_O \times \sigma_X^A"'] & O \times \hat{M}(X)^A
\end{tikzcd}$$

Observe that, by the definition of the distributive law (Proposition 11), this diagram is exactly (12). Using (12), we can now easily show that the following commutes.

$$\begin{tikzcd}
MX \arrow[r, "\sigma_X"] \arrow[d, "M\langle o,t\rangle"'] & \hat{M}X \arrow[d, "\hat{M}\langle o,t\rangle"] & \\
MFMX \arrow[r, "\sigma_{FMX}"] \arrow[d, "\lambda_{MX}"'] & \hat{M}FMX \arrow[d, "\hat{\lambda}_{MX}"'] \arrow[dr, "\hat{M}F\sigma_X"] & \\
FMMX \arrow[r, "F\sigma_{MX}"] \arrow[dd, "F\mu"'] & F\hat{M}MX \arrow[d, "F\hat{M}\sigma_X"'] & \hat{M}F\hat{M}X \arrow[dl, "\hat{\lambda}_{\hat{M}X}"] \\
 & F\hat{M}\hat{M}X \arrow[d, "F\hat{\mu}_X"] & \\
FMX \arrow[r, "F\sigma_X"'] & F\hat{M}X &
\end{tikzcd}$$

Indeed, commutativity of the topmost square is given by naturality of σ. The fact that σ is a monad morphism entails commutativity of the bottom square. The rightmost square commutes by naturality of $\hat{\lambda}$. The missing square, the one in the centre, is exactly (12).

Now observe that the leftmost border in the above diagram, the morphism $MX \to FMX$, equals $c^\sharp = \langle o^\sharp, t^\sharp \rangle$ (see (2)). The determinisation $\hat{c}^\sharp$ of $\hat{c} = \langle o, \hat{t} \rangle = \langle o, (\sigma_X)^A \circ t \rangle$ obtained using $\hat{a}$ and $\hat{\lambda}$ coincides with the rightmost border of the above diagram, the morphism $\hat{M}X \to F\hat{M}X$. The commuting of the above diagram means that σ_X is a homomorphism of F-coalgebras. By postcomposing this homomorphism with the unique F-coalgebra morphism $\hat{[\![\cdot]\!]} \colon \hat{M}X \to O^{A^*}$, one obtains an F-coalgebra morphism of type $TX \to O^{A^*}$. Since $[\![\cdot]\!]$ is the unique such, $[\![\cdot]\!] = \hat{[\![\cdot]\!]} \circ \sigma_X$ follows.

$$\begin{tikzcd}
MX \arrow[r, "\sigma_X"] \arrow[d, "c^\sharp"'] \arrow[rr, bend left, "{[\![\cdot]\!]}"] & \hat{M}X \arrow[r, "\hat{[\![\cdot]\!]}"] \arrow[d, "\hat{c}^\sharp"] & O^{A^*} \arrow[d, "\cong"', "\zeta"] \\
FMX \arrow[r, "F\sigma_X"] \arrow[rr, bend right, "F{[\![\cdot]\!]}"'] & F\hat{M}X \arrow[r, "F\hat{[\![\cdot]\!]}"] & F(O^{A^*})
\end{tikzcd}$$

Now, since σ is a monad map, $\hat{\eta} = \sigma \circ \eta$. Therefore $[\![\cdot]\!] \circ \eta_X = \hat{[\![\cdot]\!]} \circ \sigma_X \circ \eta_X = \hat{[\![\cdot]\!]} \circ \hat{\eta}_X$. $\square$

Proof. (of Theorem 16.2) Consider the following diagram in $\mathbf{Sets}$. Both squares on the left trivially commute by definition. To prove that also the square on the right commutes, it is enough to show that $h^{A^*} \colon O^{A^*} \to \hat{O}^{A^*}$ coincides with the unique coalgebra morphism $[\![\cdot]\!]_d$ from the coalgebra $d = \langle h \circ \epsilon, (\cdot)_a \rangle$ where $\zeta = \langle \epsilon, (\cdot)_a \rangle$ to the final $\hat{O} \times (\cdot)^A$-coalgebra $\hat{\zeta}$. Here $\epsilon \colon O^{A^*} \to O$ is given by $\epsilon(\varphi) = \varphi(\varepsilon)$ for the empty word $\varepsilon \in A^*$, and $(\cdot)_a \colon O^{A^*} \to (O^{A^*})^A$ is defined by $(\varphi)_a(a) = \varphi_a = \lambda w \in A^*. \varphi(aw)$. The definition of $\hat{\zeta}$ is the same as the definition of ζ, with $\hat{O}$ instead of O.

From the inductive definition of $[\![\cdot]\!]_d$, see e.g. [28], we get $[\![\varphi]\!]_d = \lambda w \in A^*. h \circ \epsilon((\varphi)_w)$ where $(\varphi)_w(u) = \varphi(wu)$ which easily leads to $[\![\varphi]\!]_d = h \circ \varphi = h^{A^*}(\varphi)$.

Now observe that $h \circ o^\sharp$ is equal to $(h \circ o)^\sharp$, since h is an algebra morphism.

From this observation and the commuting of the above diagram, it follows that $h^{A^*} \circ [\![\cdot]\!]$ is the unique coalgebra morphism from $\hat{c}^\sharp = \langle (h \circ o)^\sharp, t^\sharp \rangle$ to the final $\hat{O} \times (\cdot)^A$-coalgebra, and hence it equals $\hat{[\![\cdot]\!]}$. $\square$

Appendix G. Bialgebras, Invariance and Proofs of Back-Compatibility, Section 6

The bialgebras of probabilistic traces. At this point we would like to explicitly mention each of the three pointed convex semilattices, with bottom, or with top, carried by the carrier of the final coalgebra. The algebraic operations of these bialgebras of probabilistic traces are defined pointwise, we illustrate here the explicit definition of the bialgebra of may probabilistic traces: The carrier of the final coalgebra is $[0,1]^{A^*}$. The coalgebra map is $\zeta = \langle \epsilon, \cdot_a \rangle$ where $\epsilon(\varphi) = \varphi(\varepsilon)$ for $\varphi \colon A^* \to [0,1]$ and ε the empty word in A^*. The algebraic structure is defined pointwise from $\mathbb{M}\mathrm{ax}_B = ([0,1], \max, +_p, 0)$ resulting in the pointed convex semilattice with bottom $([0,1]^{A*}, \oplus, +_p, \varphi_0)$ where $\varphi_0(w) = 0$ for all $w \in A^*$; $\varphi_1 \oplus \varphi_2 = \varphi$ for $\varphi(w) = \max\{\varphi_1(w), \varphi_2(w)\}$, for all $w \in A^*$; and $\varphi_1 +_p \varphi_2 = \varphi$ for $\varphi(w) = \varphi_1(w) +_p \varphi_2(w)$, for all $w \in A^*$. In the same way one can explicitly write the pointed convex semilattice (with top) operations of the may-must (and the must) probabilistic traces.

Consequences of the invariance theorem. We might have performed the generalised determinisation in a number of different ways. We now show that all these ways lead however to the above three semantics.

First consider the coalgebra $\langle \bar{o}_B, \bar{t} \rangle \colon X \to T_{\mathcal{CS}\mathcal{B}}1 \times (T_{\mathcal{PCS}}X)^A$ and observe that the algebra $\mathbb{M}\mathrm{ax}_B = ([0,1], \max, +_p, 0)$, namely $\mu_1 \colon T_{\mathcal{CS}\mathcal{B}}T_{\mathcal{CS}\mathcal{B}}1 \to T_{\mathcal{CS}\mathcal{B}}1$, is also a pointed convex semilattice—formally this is $\mu_1 \circ q_B \colon T_{\mathcal{PCS}}T_{\mathcal{CS}\mathcal{B}}1 \to T_{\mathcal{CS}\mathcal{B}}1$. One could thus perform the generalised determinisation w.r.t. this algebra and the monad $T_{\mathcal{PCS}}$ and obtain an equivalence that we denote by $\equiv'_B$. Theorem 16.1 guarantees however that $\equiv'_B = \equiv_B$. Similarly, one could start with the coalgebra $\langle \bar{o}_T, \bar{t} \rangle$, apply the same construction and end up with an equivalence which, by Theorem 16.1, coincides with $\equiv_T$.

More interestingly, the semantics does not change also when eliminating conv from the definition of $\bar{t}$. Indeed one can consider the monad $C' = T_{\Sigma_{NP}, E_N \cup E_P}$, namely the monad C without the axiom (D) and consider the injective natural tranformation $\kappa \colon \mathcal{P}_{ne}\mathcal{D} \Rightarrow C'$ defined by $\kappa(\{\varphi_1, \ldots, \varphi_n\}) = \bar{\varphi}_1 \oplus \ldots \oplus \bar{\varphi}_n$ with $\bar{\varphi}$ being the term in signature $+_p$ representing the distribution φ, e.g., for

$$\begin{tikzcd}
MX \arrow[r, "\llbracket\cdot\rrbracket"] \arrow[d, "c^\sharp"'] & O^{A^*} \arrow[r, "h^{A^*}"] \arrow[d, "\cong"', "\zeta"] & \hat{O}^{A^*} \arrow[dd, "\cong"', "\hat{\zeta}"] \\
O \times (MX)^A \arrow[r, "id_O \times \llbracket\cdot\rrbracket^A"] \arrow[d, "h \times id_{(MX)^A}"'] & O \times (O^{A^*})^A \arrow[d, "h \times id_{(O^{A^*})^A}"] & \\
\hat{O} \times (MX)^A \arrow[r, "id_{\hat{O}} \times \llbracket\cdot\rrbracket^A"'] & \hat{O} \times (O^{A^*})^A \arrow[r, "id_{\hat{O}} \times (h^{A^*})^A"'] & \hat{O} \times (\hat{O}^{A^*})^A
\end{tikzcd}$$

$\varphi = (x \mapsto \frac{1}{2}, y \mapsto \frac{1}{2})$, $\bar{\varphi} = x +_{\frac{1}{2}} y$. Let $q_D \colon C' \Rightarrow C$ be the monad morphism quotienting C' by the axiom (D). One can check that $\mathrm{conv} = q_D \circ \kappa$ and thus define

$$\bar{t}' = \big(X \xrightarrow{t} (\mathcal{P}_{ne}\mathcal{D}X + 1)^A \xrightarrow{(\kappa+1)^A} (C'X + 1)^A \xrightarrow{\iota^A} (C'(X+1))^A\big)$$

and $\bar{o}' = \bar{o}$. Since the algebra $\mathbb{M}_{\mathcal{J},[0,0]}$, the free $C(\cdot + 1)$-algebra generated by 1, is also an algebra for $C'(\cdot + 1)$ — formally this is $\mu_1 \circ q_D$ one can perform the generalised determinisation of $\langle \bar{o}', \bar{t}' \rangle$ and obtains an equivalence $\equiv'$. By observing that $\bar{t} = q_D \circ \bar{t}'$, using Lemma 32, Theorem 16.1 guarantees that $\equiv' = \equiv$.

Actually, one can even take the purely syntactic monad $T_{\Sigma_{NP}}$ and consider the monad map $q \colon T_{\Sigma_{NP}} \Rightarrow C$ with epi components. We pick a right inverse r_X of q_X and consider

$$\bar{t}'' = \big(X \xrightarrow{t} (\mathcal{P}_{ne}\mathcal{D}X + 1)^A \xrightarrow{(*)} (T_\Sigma X + 1)^A \xrightarrow{\iota^A} (T_\Sigma(X+1))^A\big)$$

where $(*) = (r_X \circ \mathrm{conv}_X + 1)^A$, and $\bar{o}'' = \bar{o}$. Then again using Lemma 32 it is easy to see that $\bar{t} = q \circ \bar{t}''$. Moreover, $\mathbb{M}_{\mathcal{J},[0,0]}$ is also an algebra for $T_\Sigma(\cdot + 1)$, hence determinising $\langle \bar{o}'', \bar{t}'' \rangle$ w.r.t. $T_\Sigma(\cdot + 1)$ gives the semantics $\equiv'' = \equiv$.

Backcompatibility. For back-compatibility, we consider an LTS:

$$t \colon X \longrightarrow (\mathcal{P}X)^A \xrightarrow{\cong} (\mathcal{P}_{ne}X + 1)^A \xrightarrow{\iota} (\mathcal{P}_{ne}(X+1))^A,$$

take as observation set $O = \mathcal{P}_{ne}(1+1)$, add observations $o \colon X \to (\mathcal{P}_{ne}(1+1))$ in our standard way, and use the free pointed semilattice generated by 1 as the algebra on observations.

We denote by $\equiv_*^{LTS}$ the semantics obtained via the determinisation $\langle o^\sharp, t^\sharp \rangle$ and call it the LTS may-must trace equivalence[3].

Similarly, quotienting $\mathcal{P}_{ne}(\cdot + 1)$ by (B) and (T), respectively, leads to two more trace semantics: $\equiv_B^{LTS}$ and $\equiv_T^{LTS}$ which we call may trace equivalence and must trace equivalence, respectively. Note that $\equiv_B^{LTS} = \equiv^{LTS}$ — the standard LTS trace semantics, due to the isomorphism $T_{\mathcal{SB}} \cong \mathcal{P}$.

3. A more appropriate but longer name is may testing semantics when tests are finite traces

We start with the following simple observations that are easy to check by unfolding the definitions.

Lemma 42. *We have two injective monad maps $\chi_{\mathcal{P}_{ne}} \colon \mathcal{P}_{ne} \Rightarrow C$ and $\chi_{\mathcal{D}} \colon \mathcal{D} \Rightarrow C$ given by $\chi_{\mathcal{P}_{ne}} = \mathrm{conv} \circ \mathcal{P}_{ne}\eta^{\mathcal{D}}$ and $\chi_{\mathcal{D}} = \mathrm{conv} \circ \eta^{\mathcal{P}_{ne}}$.* □

Note that $\chi_{\mathcal{D}}(\varphi) = \{\varphi\}$, for any distribution φ, as singleton sets are convex. Using Lemma 1, we immediately get that $\chi_{\mathcal{P}_{ne}}(\cdot + 1) \colon \mathcal{P}_{ne}(\cdot + 1) \Rightarrow C(\cdot + 1)$ and $\chi_{\mathcal{D}}(\cdot + 1) \colon \mathcal{D}(\cdot + 1) \Rightarrow C(\cdot + 1)$ are injective monad maps. From this fact, and Corollary 17, we immediately get backward compatibility for may-must trace semantics of LTS and RPLTS.

Corollary 43. *May-must trace semantics after determinisation $\equiv_*^{LTS}$ for LTS coincides with may-must trace semantics after determinisation $\equiv$ of the LTS seen as NPLTS. The same holds for trace semantics after determinisation of RPLTS, i.e., $\equiv^{RP} = \equiv$ for the may-must trace semantics after determinisation $\equiv$ of the RPLTS seen as NPLTS.* □

Proving back-compatibility of may trace semantics and must trace semantics is a little bit more involved.

Lemma 44. *There are injective monad maps $T_{\mathcal{SB}} \Rightarrow T_{\mathcal{CSB}}$ and $T_{\mathcal{ST}} \Rightarrow T_{\mathcal{CST}}$, and hence $\equiv_B^{LTS} = \equiv_B$ and $\equiv_T^{LTS} = \equiv_T$.*

Proof. Note that, again by Corollary 17, for $\equiv_B^{LTS} = \equiv_B$ and $\equiv_T^{LTS} = \equiv_T$ it is enough to find injective monad maps $T_{\mathcal{SB}} \Rightarrow T_{\mathcal{CSB}}$ and $T_{\mathcal{ST}} \Rightarrow T_{\mathcal{CST}}$. We present the proof for $T_{\mathcal{SB}} \Rightarrow T_{\mathcal{CSB}}$, the proof for $T_{\mathcal{ST}} \Rightarrow T_{\mathcal{CST}}$ is analogous.

We define $e \colon T_{\mathcal{SB}} \Rightarrow T_{\mathcal{CSB}}$ by $e_X([t]_{\mathcal{SB}}) = [t]_{\mathcal{CSB}}$ for any term t with variables in X in signature $\Sigma_N \cup \Sigma_T$, where $[t]_{\mathcal{SB}}$ on the left denotes the equivalence class of t modulo $E_N \cup \{(B)\}$ and $[t]_{\mathcal{CSB}}$ on the right the equivalence class of t modulo $E_{NP} \cup \{(B)\}$. This is justified as every $T_{\mathcal{SB}}$-term is a $T_{\mathcal{CSB}}$-term as well. This is easily seen to be a monad map, we need to check well-definedness and injectivity: $t =_{\mathcal{SB}} t' \Leftrightarrow t =_{\mathcal{CSB}} t'$. Well-definedness, the implication left-to-right, is immediate as the equations of a semilattice with bottom are included in the equations of a convex semilattice with bottom. Assume $t =_{\mathcal{CSB}} t'$. Let $\bar{s}$ denote the term obtained from a term s in $T_{\mathcal{CSB}}$ by replacing every occurrence of $+_p$ by $\oplus$. Then we have

$$s_1 =_{\mathcal{CSB}} s_2 \Rightarrow \bar{s}_1 =_{\mathcal{SB}} \bar{s}_2$$

which is easy to show by checking that it holds for each of the equations.

Now, let $t = t_1 =_{\mathcal{ESB}} t_2 \cdots =_{\mathcal{ESB}} t_n = t'$. Then $t = \bar{t}_1 =_{\mathcal{SB}} \bar{t}_2 \cdots =_{\mathcal{SB}} \bar{t}_n = \bar{t}'$ where the first and last equality hold since t and t' are terms in $\Sigma_N \cup \{(B)\}$ showing injectivity. $\square$

Using similar arguments about the underlaying algebraic theories, one can prove back-compatibility of may and must trace semantics after determinisation for RPLTS seen as NPLTS.

Appendix H. Proofs of Section 7

H.1. Proof of Theorem 23

Given a resolution $\mathcal{R} = (Y, \mathrm{corr}, r)$, we define the function $\mathrm{reach}_{\mathcal{R}} \colon Y \to \mathcal{D}(Y+1)^{A^*}$ inductively as

$$\begin{aligned}&\mathrm{reach}_{\mathcal{R}}(y)(\varepsilon) = \delta_y;\\ &\mathrm{reach}_{\mathcal{R}}(y)(aw) =\\ &\begin{cases}\delta_\star & \text{if } r(y)(a) = \star;\\ \sum_{y' \in \mathrm{supp}(\Delta)} \Delta(y') \cdot \mathrm{reach}_{\mathcal{R}}(y')(w) & \text{if } r(y)(a) = \Delta.\end{cases}\end{aligned}$$

Intuitively, this assigns to each state $y \in Y$ and word $w \in A^*$ a subdistribution over Y, which is the state of the determinised system that y reaches via w.

Let $o'^\sharp \colon \mathcal{D}(Y+1) \to [0,1]$ be the function assigning to a subdistribution Δ its total mass, namely $1 - \Delta(\star)$. More formally, this is defined inductively as

$$o'^\sharp(\Delta) = \begin{cases}0 & \text{if } \Delta = \delta_\star;\\ 1 & \text{if } \Delta = \delta_y \text{ for } y \in Y;\\ o'^\sharp(\Delta_1) +_p o'^\sharp(\Delta_2) & \text{if } \Delta = \Delta_1 +_p \Delta_2.\end{cases}$$

Lemma 45. $o'^\sharp \circ \mathrm{reach}_{\mathcal{R}} = \mathrm{prob}_{\mathcal{R}}$.

Proof. We prove that $o'^\sharp(\mathrm{reach}_{\mathcal{R}}(y)(w)) = \mathrm{prob}_{\mathcal{R}}(y)(w)$ for all $y \in Y$ and $w \in A^*$. The proof proceeds by induction on w.

Base case: $w = \varepsilon$.

$$\mathrm{prob}_{\mathcal{R}}(y)(\varepsilon) = 1 = o'^\sharp(\delta_y) = o'^\sharp(\mathrm{reach}_{\mathcal{R}}(y)(\varepsilon))$$

Inductive case: $w = aw'$. If $r(y)(a) = \star$, then

$$\mathrm{prob}_{\mathcal{R}}(y)(aw') = 0 = o'^\sharp(\delta_\star) = o'^\sharp(\mathrm{reach}_{\mathcal{R}}(y)(aw')).$$

If $r(y)(a) = \Delta$, then

$$\begin{aligned}&\mathrm{prob}_{\mathcal{R}}(y)(aw')\\ &= \sum_{y' \in \mathrm{supp}(\Delta)} \Delta(y') \cdot \mathrm{prob}_{\mathcal{R}}(y')(w') && \text{(definition)}\\ &= \sum_{y' \in \mathrm{supp}(\Delta)} \Delta(y') \cdot o'^\sharp(\mathrm{reach}_{\mathcal{R}}(y')(w')) && \text{(IH)}\\ &= o'^\sharp(\sum_{y' \in \mathrm{supp}(\Delta)} \Delta(y') \cdot (\mathrm{reach}_{\mathcal{R}}(y')(w'))) && (o'^\sharp \text{ hom.})\\ &= o'^\sharp(\mathrm{prob}_{\mathcal{R}}(y)(aw')) && \text{(definition).}\end{aligned}$$

$\square$

Given a NPLTS (X, t), we define the function $\mathrm{reach} \colon X \to C(X+1)^{A^*}$ inductively as

$$\begin{aligned}&\mathrm{reach}(x)(\varepsilon)\{\delta_x\};\\ &\mathrm{reach}(x)(aw) =\\ &\begin{cases}\{\delta_\star\} & \text{if } t(x)(a) = \star;\\ \bigoplus_{\Delta \in \mathrm{conv}(S)} \sum_{x' \in \mathrm{supp}(\Delta)} \Delta(x') \cdot \mathrm{reach}(x')(w) & \text{if } t(x)(a) = S.\end{cases}\end{aligned}$$

For each NPLTS (X, t), we have a function $\langle\!\langle \cdot \rangle\!\rangle \colon C(X+1) \to C(X+1)^{A^*}$ defined for all $S \in C(X+1)$ and $w \in A^*$ as

$$\begin{aligned}\langle\!\langle S \rangle\!\rangle(\varepsilon) &= S;\\ \langle\!\langle S \rangle\!\rangle(aw) &= \langle\!\langle \bar{t}^\sharp(S)(a) \rangle\!\rangle(w).\end{aligned}$$

Lemma 46. $[\![\cdot]\!] = \bar{o}^\sharp \circ \langle\!\langle \cdot \rangle\!\rangle$

Proof. Trivial by the inductive definitions of $[\![\cdot]\!]$ and $\langle\!\langle \cdot \rangle\!\rangle$. $\square$

Lemma 47. $\langle\!\langle \cdot \rangle\!\rangle \circ \eta = \mathrm{reach}$

Proof. The proof goes by induction on $w \in A^*$.

Base case: if $w = \varepsilon$, then $\mathrm{reach}(x)(\varepsilon) = \{\delta_x\} = \eta(x) = \langle\!\langle \eta(x) \rangle\!\rangle(\varepsilon)$.

Inductive case: $w = aw'$. If $t(x)(a) = \star$, then $\langle\!\langle \eta(x) \rangle\!\rangle(aw') = \langle\!\langle \bar{t}^\sharp(\{\delta_x\})(a) \rangle\!\rangle(w') = \langle\!\langle \{\delta_\star\} \rangle\!\rangle(w') = \{\delta_\star\} = \mathrm{reach}(x)(aw')$.

If $t(x)(a) = S$, then $\mathrm{reach}(x)(aw) = \bigoplus_{\Delta \in \mathrm{conv}(S)} \sum_{x' \in \mathrm{supp}(\Delta)} \Delta(x') \cdot \mathrm{reach}(x')(w)$. By induction hypothesis, the latter is equal to $\bigoplus_{\Delta \in \mathrm{conv}(S)} \sum_{x' \in \mathrm{supp}(\Delta)} \Delta(x') \cdot \langle\!\langle \eta(x') \rangle\!\rangle(w')$. Since $\langle\!\langle \cdot \rangle\!\rangle$ is a homomorphism of convex semilattices, the latter is equal to $\langle\!\langle \bigoplus_{\Delta \in \mathrm{conv}(S)} \sum_{x' \in \mathrm{supp}(\Delta)} \Delta(x') \cdot \eta(x') \rangle\!\rangle(w)$ that is $\langle\!\langle \mathrm{conv}(S) \rangle\!\rangle(w') = \langle\!\langle \bar{t}^\sharp(\{\delta_x\})(a) \rangle\!\rangle(w') = \langle\!\langle \eta(x) \rangle\!\rangle(aw')$. $\square$

Proposition 48. $[\![\cdot]\!] \circ \eta = \bar{o}^\sharp \circ \mathrm{reach}$

Proof. By Lemma 46, $[\![\cdot]\!] \circ \eta = \bar{o}^\sharp \circ \langle\!\langle \cdot \rangle\!\rangle \circ \eta$. By Lemma 47, $\bar{o}^\sharp \circ \langle\!\langle \cdot \rangle\!\rangle \circ \eta = \bar{o}^\sharp \circ \mathrm{reach}$. $\square$

Proposition 49. *Let (X, t) be a NPLTS and let $\mathcal{R} = (Y, \mathrm{corr}, r)$ be one of its resolutions. Let $x \in X$ and $y \in Y$ such that $\mathrm{corr}(y) = x$. For all $w \in A^*$,*

$$\mathcal{D}(\mathrm{corr}+1)(\mathrm{reach}_{\mathcal{R}}(y)(w)) \in \mathrm{reach}(x)(w).$$

Proof. By induction on the structure of w. If $w = \epsilon$ then

$$\begin{aligned}\mathcal{D}(\mathrm{corr}+1)(\mathrm{reach}_{\mathcal{R}}(y)(\epsilon)) &= \mathcal{D}(\mathrm{corr}+1)(\delta_y)\\ &= \delta_x\\ &\subseteq \{\delta_x\}\\ &\in \mathrm{reach}(x)(\epsilon)\end{aligned}$$

If $w = aw'$ and $t(x)(a) = \star$ then we have $r(y)(a) = \star$, and

$$\begin{aligned}\mathcal{D}(\mathrm{corr}+1)(\mathrm{reach}_{\mathcal{R}}(y)(aw')) &= \mathcal{D}(\mathrm{corr}+1)(\delta_\star)\\ &= \delta_\star\\ &\in \{\delta_\star\}\\ &= \mathrm{reach}(x)(aw')\end{aligned}$$

If $t(x)(a) \neq \star$ then we have $r(y)(a) \neq \star$. Let $r(y)(a) = \Delta \in \mathcal{D}(Y)$. We have:

$$\begin{aligned}&\mathcal{D}(\mathrm{corr}+1)(\mathrm{reach}_{\mathcal{R}}(y)(aw'))\\ &= \mathcal{D}(\mathrm{corr}+1)(\sum_{y'\in\mathrm{supp}(\Delta)} \Delta(y')\cdot \mathrm{reach}_{\mathcal{R}}(y')(w'))\\ &= \sum_{y'\in\mathrm{supp}(\Delta)} \Delta(y')\cdot \mathcal{D}(\mathrm{corr}+1)(\mathrm{reach}_{\mathcal{R}}(y')(w'))\end{aligned}$$

By the inductive hypothesis, for each y' we have $\mathcal{D}(\mathrm{corr}+1)(\mathrm{reach}_{\mathcal{R}}(y')(w')) \in \mathrm{reach}(\mathrm{corr}(y'))(w')$. Hence, by the definition of Minkowski sum,

$$\begin{aligned}&\sum_{y'\in\mathrm{supp}(\Delta)} \Delta(y')\cdot \mathcal{D}(\mathrm{corr}+1)(\mathrm{reach}_{\mathcal{R}}(y')(w'))\\ &\in \sum_{y'\in\mathrm{supp}(\Delta)} \Delta(y')\cdot \mathrm{reach}(\mathrm{corr}(y'))(w'))\end{aligned}$$

Since $\mathcal{R}$ is a resolution, there is a $\Delta' \in \mathrm{conv}(t(x)(a))$ such that $\mathcal{D}(\mathrm{corr})(\Delta) = \Delta'$. This means that $\Delta'(x) = \sum_{\{y'\in\mathrm{supp}(\Delta) \mid \mathrm{corr}(y')=x'\}} \Delta(y')$, and thus:

$$\begin{aligned}&\sum_{y'\in\mathrm{supp}(\Delta)} \Delta(y')\cdot \mathrm{reach}(\mathrm{corr}(y'))(w'))\\ &= \sum_{x'\in\mathrm{supp}(\Delta')} \Delta'(x')\cdot \mathrm{reach}(x')(w')\end{aligned}$$

as easily follows from the axioms of convex algebras. We can then conclude by the definition of $\mathrm{reach}(x)(aw')$

$$\begin{aligned}&\sum_{x'\in\mathrm{supp}(\Delta')} \Delta'(x')\cdot \mathrm{reach}(x')(w')\\ &\subseteq \bigoplus_{\Delta'\in\mathrm{conv}(t(a)(x))} \sum_{x'\in\mathrm{supp}(\Delta')} \Delta'(x')\cdot \mathrm{reach}(x')(w')\\ &= \mathrm{reach}(x)(aw').\end{aligned}$$

□

Proposition 50. *Let (X,t) be a NPLTS. For all $x \in X$ and $w \in A^*$, if $\Delta \in \mathrm{reach}(x)(w)$ then there exists a resolution $\mathcal{R} = (Y, \mathrm{corr}, r)$ and a state $y \in Y$ such that*

(1) $\mathrm{corr}(y) = x$ *and*
(2) $\mathcal{D}(\mathrm{corr}+1)(\mathrm{reach}_{\mathcal{R}}(y)(w)) = \Delta$.

Proof. The proof proceeds by induction on $w \in A^*$.

In the base case $w = \varepsilon$. For all $x \in X$ and $a \in A$ such that $t(x)(a) \neq \star$, we can choose one distribution $\Delta_{x,a} \in t(x)(a)$. Then, we take $\mathcal{R} = (X, id_X, r)$ where $r\colon X \to (\mathcal{D}(X)+1)^A$ is defined for all $x \in X$ and $a \in A$ as

$$r(x)(a) = \begin{cases}\star & \text{if } t(x)(a) = \star;\\ \Delta_{x,a} & \text{otherwise.}\end{cases}$$

By construction $\mathcal{R}$ is a resolution. Then we take x as the selected state y of the resolution $\mathcal{R}$. Since the correspondence function is id_X, (1) is immediately satisfied. Now, by definition, $\mathrm{reach}_{\mathcal{R}}(x)(\varepsilon) = \delta_x$ and $\mathrm{reach}(x)(\varepsilon) = \{\delta_x\}$. We conclude by observing that $\mathcal{D}(id_X+1)(\delta_x) = \delta_x \in \{\delta_x\} = \mathrm{reach}(x)(\varepsilon)$.

In the inductive case $w = aw'$. Now we have two cases to consider: either $t(x)(a) = \star$ or $t(x)(a) = S$ for $S \in \mathcal{P}_{ne}\mathcal{D}(X)$.

Assume $t(x)(a) = \star$. Then $\mathrm{reach}(x)(aw') = \{\delta_\star\}$. Let $\mathcal{R} = (X, id_X, r)$ be the resolution defined as in the base case, and take x as the selected state y of the resolution $\mathcal{R}$. Since the correspondence function is id_X, (1) is immediately satisfied. Since $\mathcal{R}$ is a resolution, $t(x)(a) = \star$ implies $r(x)(a) = \star$. Hence, $\mathrm{reach}_{\mathcal{R}}(x)(a) = \delta_\star$ and $\mathrm{reach}(x)(a) = \{\delta_\star\}$. We conclude by $\mathcal{D}(id_X+1)(\delta_\star) = \delta_\star \in \{\delta_\star\} = \mathrm{reach}(x)(a)$.

Assume $t(x)(a) = S$. Then

$$\mathrm{reach}(x)(aw') = \bigoplus_{\Delta'\in\mathrm{conv}(S)} \sum_{x'\in\mathrm{supp}(\Delta')} \Delta'(x')\cdot\mathrm{reach}(x')(w').$$

By Proposition 40, it holds that

$$\mathrm{reach}(x)(aw') = \bigcup_{\Delta'\in\mathrm{conv}(S)} \sum_{x'\in\mathrm{supp}(\Delta')} \Delta'(x')\cdot\mathrm{reach}(x')(w').$$

This is equivalent to saying that there exists a $\Delta' \in \mathrm{conv}(S)$ such that

$$\Delta \in \sum_{x'\in\mathrm{supp}(\Delta')} \Delta'(x')\cdot \mathrm{reach}(x')(w').$$

This is in turn equivalent to saying (by the definition of Minkowski sum) that for every $x' \in \mathrm{supp}(\Delta')$ there exists a $\Delta'_{x'} \in \mathrm{reach}(x')(w')$ such that

$$\Delta = \sum_{x'\in\mathrm{supp}(\Delta')} \Delta'(x')\cdot \Delta'_{x'}.$$

We can now use the induction hypothesis on $\Delta'_{x'} \in \mathrm{reach}(x')(w')$: for each $\Delta'_{x'} \in \mathrm{reach}(x')(w')$ there exists a resolution $\mathcal{R}_{x'} = (Y_{x'}, \mathrm{corr}_{x'}, r_{x'})$ and a $y_{x'} \in Y_{x'}$ such that

(c) $\mathrm{corr}_{x'}(y_{x'}) = x'$ and
(d) $\mathcal{D}(\mathrm{corr}_{x'}+1)(\mathrm{reach}_{\mathcal{R}_{x'}}(y_{x'})(w)) = \Delta'_{x'}$.

Now we construct the coproduct of all the resolutions $\mathcal{R}_{x'}$. Take Z to be the disjoint union of all the $Y_{x'}$ and define $\mathrm{corr}_Z\colon Z \to X$ as $\mathrm{corr}_Z(z) = \mathrm{corr}_{x'}(z)$ if $z \in Y_{x'}$. Similarly, we define $r_Z\colon Z \to (\mathcal{D}(Z+1))^A$ as $r_Z(z) = r_{x'}(z)$ if $z \in Y_{x'}$. By construction, $\mathcal{R}_Z = (Z, \mathrm{corr}_Z, r_Z)$ is a resolution of (X,t).

Let $\mathcal{R}' = (X, id_X, r')$ be a resolution defined as in the base case, i.e., by arbitrarily choosing a distribution

$\Delta_{x,a} \in t(x)(a)$, for any x and a, as value of $r'(x)(a)$, whenever $t(x)(a) \neq \star$. We define the resolution $\mathcal{R} = (Y, \mathrm{corr}, r)$ needed to conclude the proof as follows: the state space is $Y = Z + X + \{y\}$, namely the disjoint union of Z, of X, and of the singleton containing a fresh state y; the correspondence function $\mathrm{corr}\colon Y \to X$ and the transition function $r\colon Y \to (\mathcal{D}(Y)+1)^A$ are defined for all $u \in Y$ as

$$\mathrm{corr}(u) = \begin{cases} \mathrm{corr}_Z(u) & \text{if } u \in Z, \\ id_X(u) & \text{if } u \in X, \\ x & \text{if } u = y, \end{cases}$$

$$r(u)(b) = \begin{cases} r_Z(u)(b) & \text{if } u \in Z, \\ r'(u)(b) & \text{if } u \in X, \\ \Delta_{x,b} & \text{if } u = y, a \neq b, t(x)(b) \neq \star, \\ \star & \text{if } u = y, a \neq b, t(x)(b) = \star, \\ \Delta'' & \text{if } u = y,\ a = b \end{cases}$$

where Δ'' is the distribution having as support the set of states $\{y_{x'} | x' \in \mathrm{supp}(\Delta)\} \subseteq Z$, and such that $\Delta''(y_{x'}) = \Delta'(x')$. Note that Δ'' is a distribution, since Δ' is and since we are taking exactly one $y_{x'}$ for each $x' \in \mathrm{supp}(\Delta')$.

The fact that $\mathcal{R}$ is a resolution follows from $\mathcal{R}_Z$ and $\mathcal{R}'$ being resolution and y respecting – by construction– the conditions of resolution: indeed $\mathrm{corr}(y) = x$, and

- if $a \neq b$ and $t(x)(b) \neq \star$, then $r(y)(b) = \Delta_{x,b}$ and $\mathcal{D}(\mathrm{corr})(\Delta_{x,b}) = \mathcal{D}(id_X)(\Delta_{x,b}) = \Delta_{x,b} \in t(x)(b)$
- if $a \neq b$ and $t(x)(b) = \star$, then $r(y)(b) = \star$;
- if $a = b$, then $r(y)(b) = \Delta''$, and $\mathcal{D}(\mathrm{corr})(\Delta'') = \Delta'$, with $\Delta' \in \mathrm{conv}(t(x)(a))$.

To conclude the proof we only need to show that points (1) and (2) hold. The former is trivially satisfied by definition of corr. For (2), we display the following derivation.

$$\begin{aligned}
&\mathcal{D}(\mathrm{corr}+1)(\mathrm{reach}_{\mathcal{R}}(y)(aw')) \\
&= \mathcal{D}(\mathrm{corr}+1)\Big(\sum_{y_{x'} \in \mathrm{supp}(\Delta'')} (\Delta''(y_{x'}) \cdot \mathrm{reach}_{\mathcal{R}}(y_{x'})(w'))\Big) \\
&= \mathcal{D}(\mathrm{corr}+1)\Big(\sum_{x' \in \mathrm{supp}(\Delta')} (\Delta'(x') \cdot \mathrm{reach}_{\mathcal{R}}(y_{x'})(w'))\Big) \\
&= \sum_{x' \in \mathrm{supp}(\Delta')} (\Delta'(x') \cdot (\mathcal{D}(\mathrm{corr}+1)(\mathrm{reach}_{\mathcal{R}}(y_{x'})(w')))) \\
&= \sum_{x' \in \mathrm{supp}(\Delta')} (\Delta'(x') \cdot \Delta'_{x'}) \\
&= \Delta
\end{aligned}$$

□

of Theorem 23. Before starting with the actual proof, we need the following elementary observation: for all $f\colon X \to Y$ and $\Delta \in \mathcal{D}(X+1)$, it holds that

$$o'^{\sharp}(\mathcal{D}(f+1)(\Delta)) = o'^{\sharp}(\Delta), \tag{13}$$

namely, the total mass is preserved by applying $\mathcal{D}(f+1)$.

Now, suppose that $[\![\eta(x)]\!](w) = [p,q]$ for some $p, q \in [0,1]$ with $p \leq q$. By Proposition 48, it holds that $\bar{o}^{\sharp}(\mathrm{reach}(x)(w)) = [p,q]$. By definition of $\bar{o}^{\sharp}$ there exists $\Delta_{min}, \Delta_{max} \in \mathrm{reach}(x)$ such that the total mass of $\Delta_{min} = p$, the total mass of $\Delta_{max} = q$ and for an arbitrary $\Delta \in \mathrm{reach}(x)$, its total mass is in between p and q. In other words,

(a) $o'^{\sharp}(\Delta_{min}) = p$,
(b) $o'^{\sharp}(\Delta_{max}) = q$ and
(c) $p \leq o'^{\sharp}(\Delta) \leq q$ for all $\Delta \in \mathrm{reach}(x)$.

By Proposition 49, for all resolutions $\mathcal{R}$, states y such that $\mathrm{corr}(y) = x$ and distributions Δ' such that $\mathrm{reach}_{\mathcal{R}}(y)(w) = \Delta'$, one has that $\mathcal{D}(\mathrm{corr}+1)(\Delta') \in \mathrm{reach}(x)(w)$. By (c), $p \leq o'^{\sharp}(\mathcal{D}(\mathrm{corr}+1)(\Delta')) \leq q$ and, by (13), $p \leq o'^{\sharp}(\Delta') \leq q$. This means $p \leq o'^{\sharp}(\mathrm{reach}_{\mathcal{R}}(y)(w)) \leq q$ that, by Lemma 45, coincides with $p \leq \mathrm{prob}_{\mathcal{R}}(y)(w) \leq q$. This proves that $\lceil\!\lceil x \rceil\!\rceil(w) \geq p$ and $\lfloor\!\lfloor x \rfloor\!\rfloor(w) \leq q$.

We now prove that $\lceil\!\lceil x \rceil\!\rceil(w) \leq p$; the proof for $\lfloor\!\lfloor x \rfloor\!\rfloor(w) \geq q$ is analogous.

By Proposition 50, there exist resolutions $\mathcal{R}$, a state y and distributions Δ'' such that

(d) $\mathrm{corr}(y) = x$,
(e) $\mathcal{D}(\mathrm{corr}+1)(\Delta'') = \Delta_{min}$,
(f) $\mathrm{reach}_{\mathcal{R}}(y)(w) = \Delta''$.

By (e) and (13), one immediately has that $o'^{\sharp}(\Delta'') = o'^{\sharp}(\Delta_{min}) = p$. By (f), the latter means that $o'^{\sharp}(\mathrm{reach}_{\mathcal{R}}(y)(w)) = p$ that, by Lemma 45, allows to conclude that $\mathrm{prob}_{\mathcal{R}}(y)(w) = p$. This proves that $\lceil\!\lceil x \rceil\!\rceil(w) \leq p$.

□

H.2. Proof of Corollary 24

Proof. Consider the monad morphisms $q_B\colon T_{\mathcal{PCS}} \Rightarrow T_{\mathcal{CSB}}$ and $q_T\colon T_{\mathcal{PCS}} \Rightarrow T_{\mathcal{CST}}$ quotienting $T_{\mathcal{PCS}}$ by (B) and (T), respectively (see Section 6). By Theorem 16 item 2 we have

$$[\![\eta(x)]\!]_B(w) = ({q_B}^{A*} \circ [\![\eta(x)]\!])(w)$$

$$[\![\eta(x)]\!]_T(w) = ({q_T}^{A*} \circ [\![\eta(x)]\!])(w)$$

For an interval $[p,q]$, $q_B([p,q]) = q$ and $q_T([p,q]) = p$. Then by Theorem 23 we derive

$$\begin{aligned}
({q_B}^{A*} \circ [\![\eta(x)]\!])(w) &= q_B([\![\eta(x)]\!](w)) \\
&= q_B([\, \lceil\!\lceil x \rceil\!\rceil(w), \lfloor\!\lfloor x \rfloor\!\rfloor(w) \,]) \\
&= \lfloor\!\lfloor x \rfloor\!\rfloor(w)
\end{aligned}$$

and, analogously,

$$\begin{aligned}
({q_T}^{A*} \circ [\![\eta(x)]\!])(w) &= q_T([\![\eta(x)]\!](w)) \\
&= q_T([\, \lceil\!\lceil x \rceil\!\rceil(w), \lfloor\!\lfloor x \rfloor\!\rfloor(w) \,]) \\
&= \lceil\!\lceil x \rceil\!\rceil(w)
\end{aligned}$$

□

H.3. Proof of Proposition 26

Proof. We first prove that $\lfloor\lfloor x\rfloor\rfloor(w) \leq \lfloor\lfloor x\rfloor\rfloor_{fp}(w)$.

Let $\mathcal{R} = (Y, \mathrm{corr}, r)$ be a resolution of (X, t), $x \in X$, and $w \in A^*$. Let $y \in Y$ such that $\mathrm{corr}(y) = x$. We show that there exists a fully probabilistic resolution $\mathcal{R}'$ of (X, t) with a state z such that z is mapped by the correspondence function of $\mathcal{R}'$ to x and such that $\mathrm{prob}_{\mathcal{R}}(y)(w) = \mathrm{prob}_{\mathcal{R}'}(z)(w)$.

We define $\mathcal{R}' = (Y \times A^*, \mathrm{corr}', r')$ as follows. The correspondece function $\mathrm{corr}' \colon Y \times A^* \to Y$ is $\mathrm{corr} \circ \pi_1$, namely $\mathrm{corr}'(y, w') = \mathrm{corr}(y)$ for all $w' \in A^*$. To define r', we use the notation $\Delta_{w'} \in \mathcal{D}(Y \times A^*)$ to denote, for all $\Delta \in \mathcal{D}(Y)$ and $w' \in A^*$, the distribution over $Y \times A^*$ given as

$$\Delta_{w'}(y, w'') \quad = \quad \begin{cases} \Delta(y) & \text{if } w' = w'', \\ 0 & \text{otherwise.} \end{cases}$$

Now $r' \colon Y \times A^* \to (A \times \mathcal{D}(Y \times A^*)) + 1$ is defined as:

$$\begin{aligned} r'(y, \epsilon) \quad &= \quad \star \\ r'(y, aw') \quad &= \quad \begin{cases} \langle a, \Delta_{w'} \rangle & \text{if } r(y)(a) = \Delta \neq \star \\ \star & \text{otherwise.} \end{cases} \end{aligned}$$

We proceed by proving that $\mathcal{R}'$ is a fully probabilistic resolution of (X, t). First, it is necessary to observe that r' is well defined: if $r(y)(a) = \Delta$ and $r(y)(b) = \Delta'$ for some $b \neq a$, then $r'(y, aw')$ is by definition $\langle a, \Delta_{w'} \rangle$: this explains why we needed to take as set of states $Y \times A^*$.

Now suppose that $r'(y, w'') \neq \star$. Then $w' = aw''$, and $r'(y, w'') = \langle a, \Delta_{w''} \rangle$ with $\Delta = r(y)(a)$. Hence, $\mathcal{D}(\mathrm{corr}')(\Delta_{w''}) = \mathcal{D}(\mathrm{corr})(\Delta)$. Since $\mathcal{R}$ is a resolution, $\mathcal{D}(\mathrm{corr})(\Delta) \in \mathrm{conv}(t(x)(a))$ and therefore $\mathcal{R}'$ is a fully probabilistic resolution.

We now prove that for all $w' \in A^*$ and for all $y \in Y$, it holds that $\mathrm{prob}_{\mathcal{R}}(y)(w') = \mathrm{prob}_{\mathcal{R}'}(y, w')(w')$. The proof goes by induction on w'.

If $w' = \epsilon$ then, $\mathrm{prob}_{\mathcal{R}}(y)(\epsilon) = 1 = \mathrm{prob}_{\mathcal{R}'}(y, \epsilon)(\epsilon)$.

If $w' = aw''$ and $r(y)(a) = \star$, then $r'(y, w') = \star$ and $\mathrm{prob}_{\mathcal{R}'}(y, w')(w') = 0 = \mathrm{prob}_{\mathcal{R}}(y)(w')$.

If $w' = aw''$ and $r(y)(a) = \Delta \neq \star$, then $r'(y, aw'') = \langle a, \Delta_{w''} \rangle$:

$$\begin{aligned} &\mathrm{prob}_{\mathcal{R}}(y)(w') \\ &= \sum_{y' \in \mathrm{supp}(\Delta)} \Delta(y') \cdot \mathrm{prob}_{\mathcal{R}}(y')(w'') \\ &= \sum_{y' \in \mathrm{supp}(\Delta)} \Delta(y') \cdot \mathrm{prob}_{\mathcal{R}'}(y', w'')(w'') && \text{(by IH)} \\ &= \sum_{(y', w'') \in \mathrm{supp}(\Delta_{w''})} \Delta_{w''}(y', w'') \cdot \mathrm{prob}_{\mathcal{R}'}(y', w'')(w'') \\ &= \mathrm{prob}_{\mathcal{R}'}(y, aw'')(aw'') \end{aligned}$$

Hence, $\mathrm{prob}_{\mathcal{R}}(y)(w) = \mathrm{prob}_{\mathcal{R}'}(y, w)(w)$, with $\mathrm{corr}(y) = \mathrm{corr}'(y, w) = x$.

We now prove that $\lfloor\lfloor x\rfloor\rfloor(w) \geq \lfloor\lfloor x\rfloor\rfloor_{fp}(w)$.

Let $\mathcal{R} = (Y, \mathrm{corr}, r)$ be a fully probabilistic resolution of (X, t), $x \in X$, and $w \in A^*$. Let $\mathrm{corr}(y) = x$. We show that there exists a resolution $\mathcal{R}' = (Y', \mathrm{corr}', r')$ of (X, t) with a state z such that $\mathrm{corr}'(z) = x$ and $\mathrm{prob}_{\mathcal{R}}(y)(w) \leq \mathrm{prob}_{\mathcal{R}'}(z)(w)$. We define $\mathcal{R}' = (Y', \mathrm{corr}', r')$ as follows:

- $Y' = Y + X$ is the (disjoint) union of Y and X
- $\mathrm{corr}' = \mathrm{corr} + id_X$ that is for all $y' \in Y'$

$$\mathrm{corr}'(y') \quad = \quad \begin{cases} \mathrm{corr}(y') & \text{if } y' \in Y; \\ id_X(y') & \text{if } y' \in X. \end{cases}$$

- $r' : Y' \to (\mathcal{D}(Y') + 1)^A$ is defined as:

$$r'(y')(a) \quad = \quad \begin{cases} \star & \text{if } t(\mathrm{corr}'(y'))(a) = \star \\ \Delta & \text{if } y' \in Y \text{ and } r(y') = \langle a, \Delta \rangle \\ \Delta_{\mathrm{corr}'(y'),a} & \text{otherwise} \end{cases}$$

where $\Delta_{x,a}$ are defined like in the base case of the proof of Proposition 50 (namely, an arbitrary choice amongst the distributions in $t(x)(a)$).

We prove that $\mathcal{R}'$ is a resolution. For elements $y' \in X$ the conditions of Definition 21 are trivially satisfied (see the analogous construction used in the base case in the proof of Proposition 50). Suppose $y' \in Y$.

1) By definition, $r'(y')(a) = \star$ iff $t(\mathrm{corr}'(y'))(a) = \star$.
2) If $r'(y')(a) \neq \star$, then we are either in the second or in the third case of the definition of r'. If we are in the second case, $r'(y')(a) = \Delta$ with $r(y') = \langle a, \Delta \rangle$. Since $\Delta \in \mathcal{D}Y$ we have $\mathcal{D}(\mathrm{corr}')(\Delta) = \mathcal{D}(\mathrm{corr})(\Delta)$, and by the definition of fully probabilistic resolution, it holds that $\mathcal{D}(\mathrm{corr})(\Delta) \in \mathrm{conv}(t(\mathrm{corr}(y'))(a))$. Therefore $\mathcal{D}(\mathrm{corr}')(\Delta) \in \mathrm{conv}(t(\mathrm{corr}'(y'))(a))$. If we are in the third case, we have $r'(y')(a) = \Delta_{\mathrm{corr}'(y'),a}$, with $\Delta_{\mathrm{corr}'(y'),a} \in t(\mathrm{corr}'(y'))(a)$. By definition of corr', $\mathcal{D}(\mathrm{corr}')(\Delta_{\mathrm{corr}'(y'),a}) = \Delta_{\mathrm{corr}'(y'),a}$. Therefore $\mathcal{D}(\mathrm{corr}')(\Delta_{\mathrm{corr}'(y'),a}) \in \mathrm{conv}(t(\mathrm{corr}'(y'))(a))$.

We conclude by showing that for all $y \in Y$ and for all w', $\mathrm{prob}_{\mathcal{R}}(y, w') \leq \mathrm{prob}_{\mathcal{R}'}(y, w')$. The proof goes by induction on w'. The case $w' = \epsilon$ is trivial, since $\mathrm{prob}_{\mathcal{R}}(y, \epsilon) = 1 = \mathrm{prob}_{\mathcal{R}'}(y, \epsilon)$. For the inductive case, take $w' = aw''$. Suppose $r(y) = \langle a, \Delta \rangle$. Then, by definition of r', $r'(y)(a) = \Delta$, and

$$\begin{aligned} \mathrm{prob}_{\mathcal{R}}(y)(aw'') &= \sum_{y' \in \mathrm{supp}(\Delta)} \Delta(y') \cdot \mathrm{prob}_{\mathcal{R}}(y')(w'') \\ &\leq \sum_{y' \in \mathrm{supp}(\Delta)} \Delta(y') \cdot \mathrm{prob}_{\mathcal{R}'}(y')(w'') \\ & \qquad\qquad\qquad\qquad\qquad\qquad \text{(by IH)} \\ &= \mathrm{prob}_{\mathcal{R}'}(y)(aw'') \end{aligned}$$

Now suppose that $r(y) \neq \langle a, \Delta \rangle$. $r(y) = \star$ then, by definition of $\mathrm{prob}_{\mathcal{R}}$, $\mathrm{prob}_{\mathcal{R}}(y)(aw'') = 0$ and there is nothing to prove.

Hence, $\mathrm{prob}_{\mathcal{R}}(y)(w) \leq \mathrm{prob}_{\mathcal{R}'}(y)(w)$, with $\mathrm{corr}(y) = \mathrm{corr}'(y) = x$. $\square$