Testing Finitary Probabilistic Processes*

Yuxin Deng[1 †], Rob van Glabbeek[2,4], Matthew Hennessy[3 ‡] & Carroll Morgan[4 §]

[1]Shanghai Jiao Tong University, China
[2]NICTA, Sydney, Australia
[3]Trinity College Dublin, Ireland
[4]University of New South Wales, Sydney, Australia

June 3, 2011

We provide both modal- and relational characterisations of may- and must-testing preorders for recursive CSP processes with divergence, featuring probabilistic as well as nondeterministic choice. May testing is characterised in terms of simulation, and must testing in terms of failure simulation. To this end we develop weak transitions between probabilistic processes, elaborate their topological properties, and express divergence in terms of partial distributions.

Contents

*An extended abstract of this paper, without Sections 4.1 and 4.3, appeared in Proc. CONCUR 2009, LNCS 5710, Springer, pp. 274–288. An extended abstract of Sections 4.1 and 4.3 was included in [3].
†Deng was supported by the National Natural Science Foundation of China (60703033).
‡Hennessy gratefully acknowledges the financial support of Science Foundation Ireland.
§Morgan acknowledges the support of ARC Discovery Grant DP0879529.

1 Introduction

It has long been a challenge for theoretical computer scientists to provide a firm mathematical foundation for process-description languages that incorporate both nondeterministic and probabilistic behaviour in such a way that processes are semantically distinguished just when they can be told apart by some notion of testing.

In our earlier work [4, 2] a semantic theory was developed for one particular language with these characteristics, a finite process calculus called pCSP: nondeterminism is present in the form of the standard choice operators inherited from CSP [10], that is $P \sqcap Q$ and $P \,\Box\, Q$, while probabilistic behaviour is added via a new choice operator $P \,{}_p{\oplus}\, Q$ in which P is chosen with probability p and Q with probability $1-p$. The intensional behaviour of a pCSP process is given in terms of a probabilistic labelled transition system [24, 4], or pLTS, a generalisation of labelled transition systems [20]. In a pLTS the result of performing an action in a given state results in a *probability distribution* over states, rather than a single state; thus the relations $s \xrightarrow{\alpha} t$ in an LTS are replaced by relations $s \xrightarrow{\alpha} \Delta$, with Δ a distribution. Closed pCSP expressions P are interpreted as probability distributions $[\![P]\!]$ in the associated pLTS. Our semantic theory [4, 2] naturally generalises the two preorders of standard testing theory [6] to pCSP:

- $P \sqsubseteq_{\text{pmay}} Q$ indicates that Q is at least as good as P from the point of view of *possibly* passing probabilistic tests; and
- $P \sqsubseteq_{\text{pmust}} Q$ indicates instead that Q is at least as good as P from the point of view of *guaranteeing* the passing of probabilistic tests.

The most significant result of [2] was an alternative characterisation of these preorders as particular forms of co-inductively defined *simulation* relations, $\sqsubseteq_S$ and $\sqsubseteq_{FS}$, over the underlying pLTS. We also provided a characterisation in terms of a modal logic.

The object of the current paper is to extend the above results to a version of pCSP with recursive process descriptions: we add a construct $\mathsf{rec}\, x.\, P$ for recursion, and extend the intensional semantics of [2] in a straightforward manner. We restrict ourselves to *finitary* pCSP processes, those having finitely many states and displaying finite branching.

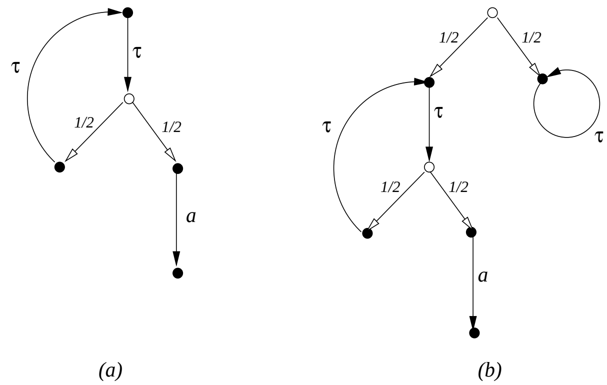


Figure 1: The pLTSs of processes Q_1 and Q_2

The simulation relations $\sqsubseteq_S$ and $\sqsubseteq_{FS}$ in [2] were defined in terms of weak transitions $\stackrel{\hat{\tau}}{\Longrightarrow}$ between distributions, obtained as the transitive closure of a relation $\stackrel{\hat{\tau}}{\longrightarrow}$ between distributions that allows one part of a distribution to make a τ-move with the other part remaining in place. This definition is however inadequate for processes that can do an unbounded number of τ-steps. The problem is highlighted by the process $Q_1 = \mathsf{rec}\, x.\, (\tau.x \;{}_{\frac{1}{2}}\!\oplus\; a)$ illustrated in Figure 1(a). Process Q_1 is indistinguishable, using tests, from the simple process a: we have $Q_1 \simeq_{\text{pmay}} a$ and $Q_1 \simeq_{\text{pmust}} a$. This is because the process Q_1 will eventually perform the action a with probability 1. However, the action $[\![a]\!] \xrightarrow{a} [\![\,\mathbf{0}\,]\!]$ can not be simulated by a corresponding move $[\![Q_1]\!] \stackrel{\hat{\tau}}{\Longrightarrow}\xrightarrow{a}$. No matter which distribution Δ we obtain from executing a finite sequence of internal moves $[\![Q_1]\!] \stackrel{\hat{\tau}}{\Longrightarrow} \Delta$, still part of it is unable to subsequently perform the action a.

To address this problem we propose a new relation $\Delta \Longrightarrow \Theta$, to indicate that Θ can be derived from Δ by performing an unbounded sequence of internal moves; we call Θ a *weak derivative* of Δ. For example $[\![a]\!]$ will turn out to be a weak derivative of $[\![Q_1]\!]$, i.e. $[\![Q_1]\!] \Longrightarrow [\![a]\!]$, via the infinite sequence of internal moves

$$[\![Q_1]\!] \xrightarrow{\hat{\tau}} [\![Q_1 \;{}_{\frac{1}{2}}\!\oplus a]\!] \xrightarrow{\hat{\tau}} [\![Q_1 \;{}_{\frac{1}{2^2}}\!\oplus a]\!] \xrightarrow{\hat{\tau}} \ldots \xrightarrow{\hat{\tau}} [\![Q_1 \;{}_{\frac{1}{2^n}}\!\oplus a]\!] \xrightarrow{\hat{\tau}} \ldots\,.$$

One of our contributions here is the significant use of "subdistributions" that sum to *no more than* one [11, 19]. For example, the empty subdistribution ε elegantly represents the chaotic behaviour of processes that in CSP and in must-testing semantics is tantamount to divergence, because we have $\varepsilon \xrightarrow{\alpha} \varepsilon$ for any action α, and a process like $\mathsf{rec}\, x.\, x$ that diverges via an infinite τ path gives rise to the weak transition $\mathsf{rec}\, x.\, x \Longrightarrow \varepsilon$. So the process $Q_2 = Q_1 \;{}_{\frac{1}{2}}\!\oplus \mathsf{rec}\, x.\, x$ illustrated in Figure 1(b) will enable the weak transition $[\![Q_2]\!] \Longrightarrow \frac{1}{2}[\![a]\!]$, where intuitively the latter is a proper subdistribution mapping the state a to the probability $\frac{1}{2}$. Our weak transition relation $\Longrightarrow$ can be regarded as an extension of the *weak hyper-transition* from [17] to partial distributions; the latter, although defined in a very different way, can be represented in terms of ours by requiring weak derivatives to be total distributions.

We end this introduction with a brief glimpse at our proof strategy. In [2] the characterisations for finite pCSP processes were obtained using a probabilistic extension of the Hennessy-Milner logic [20]. Moving to recursive processes, we know that process behaviour can be captured by a finite modal logic only if the underlying LTS is finitely branching, or at least image-finite [20]. Thus to take advantage of a finite probabilistic HML we need a property of pLTSs corresponding to finite branching in LTSs: this is topological compactness, whose relevance we now sketch.

Subdistributions over (derivatives of) finitary pCSP processes inherit the standard (complete) Euclidean metric. One of our key results is that

Theorem 1.1 For every finitary pCSP process P, the set $\{\, \Delta \mid [\![P]\!] \Longrightarrow \Delta \,\}$ is convex and compact.

Indeed, using techniques from Markov Decision Theory [22] we can show that the potentially uncountable set $\{\, \Delta \mid [\![P]\!] \Longrightarrow \Delta \,\}$ is nevertheless the convex closure of a *finite* set of subdistributions, from which Theorem 1.1 follows.

This key result allows an *inductive* characterisation of the simulation preorders $\sqsubseteq_S$ and $\sqsubseteq_{FS}$, here defined using our novel weak derivation relation $\Longrightarrow$. We first construct a sequence of approximations $\sqsubseteq_S^k$ for $k \geq 0$ and, using Theorem 1.1, we prove

Theorem 1.2 For every finitary pCSP process P, and for every $k \geq 0$, the set $\{ \Delta \mid [\![P]\!] \sqsubseteq_S^k \Delta \}$ is convex and compact.

This in turn enables us to use the *Finite Intersection Property* of compact sets to prove

Theorem 1.3 For finitary pCSP processes we have $P \sqsubseteq_S Q$ iff $P \sqsubseteq_S^k Q$ for all $k \geq 0$.

Our main characterisation results can then be obtained by extending the probabilistic modal logic used in [2], so that for example

- it characterises $\sqsubseteq_S^k$ for every $k \geq 0$, and therefore it also characterises $\sqsubseteq_S$
- every probabilistic modal formula can be captured by a may-test.

Similar results accrue for must testing and the new failure simulation preorder $\sqsubseteq_{FS}$: details are given in Section 8.

In the next section we introduce a probabilistic CSP with recursion. In Section 3 we elaborate on our approach to weak derivations and discuss some of their elementary properties. In Section 4 we present two methods of testing and show that they coincide for finitary processes. In Section 5 we introduce yet another method of testing. It appears simpler than the previous two methods because only extremal testing outcomes are considered. However, it turns out to coincide with them for finitary processes. In Section 6 we investigate the topological properties of weak derivations. In Section 7 we define a notion of failure simulation preorder. It is shown to be a precongruence relation and is sound for must testing. In Section 8 we show that failure simulation is also complete for must testing. Therefore, must testing can be characterised as failure simulation. In Section 9 we characterise may testing as simulation. Finally, related work is briefly discussed in Section 10.

2 The language pCSP

Let Act be a set of visible actions which a process can perform, and let Var be an infinite set of variables. The language pCSP of probabilistic CSP processes is given by the following two-sorted syntax, in which $p \in [0,1]$, $a \in \mathsf{Act}$ and $A \subseteq \mathsf{Act}$:

$$\begin{array}{lcl} P & ::= & S \mid P \,{}_p\!\oplus P \\ S & ::= & \mathbf{0} \mid x \in \mathsf{Var} \mid a.P \mid P \sqcap P \mid S \,\Box\, S \mid S \,|_A\, S \mid \mathsf{rec}\, x.\, P\,. \end{array}$$

This is essentially the finite language of [2, 4] plus the recursive construct $\mathsf{rec}\, x.\, P$ in which x is a variable and P a term. Intuitively $\mathsf{rec}\, x.\, P$ represents the solution of the fixed-point equation $x = P$. The notions of free- and bound variables are standard; by $Q[x \mapsto P]$ we indicate substitution of term P for variable x in Q, with renaming if necessary. We write pCSP for the set of closed P-terms defined by this grammar, and sCSP for its *state-based* subset of closed S-terms.

The process $P \,{}_p\!\oplus Q$, for $0 \leq p \leq 1$, represents a *probabilistic choice* between P and Q: with probability p it will act like P and with probability $1-p$ it will act like Q.[1] Any process is a probabilistic combination of state-based processes built by repeated application of the operator ${}_p\!\oplus$. The state-based processes have a CSP-like syntax, involving the stopped process $\mathbf{0}$, action prefixing $a.\_$ for $a \in \mathsf{Act}$, *internal*- and *external choices* $\sqcap$ and $\Box$, and a *parallel composition* $|_A$ for $A \subseteq \mathsf{Act}$.

The process $P \sqcap Q$ will first do a so-called *internal action* $\tau \notin \mathsf{Act}$, choosing *nondeterministically* between P and Q. Therefore $\sqcap$, like $a.\_$, acts as a *guard*, in the sense that it converts any process arguments into a state-based process. The same applies to $\mathsf{rec}\, x.\, P$ as, following CSP [21], our recursion construct performs an internal action when unfolding. As our testing semantics will abstract from internal actions, these τ-steps are harmless and merely simplify the semantics.

[1] In our semantics we have $[\![P \,{}_0\!\oplus Q]\!] = [\![Q]\!]$ and $[\![P \,{}_1\!\oplus Q]\!] = [\![P]\!]$, so without limitation of generality we could have required that $0<p<1$. In papers involving axiomatisations this is customary, as the most natural formulation of the law of associativity involves dividing by p.

The process $s \mathbin{\Box} t$ on the other hand does not perform actions itself but rather allows its arguments to proceed, disabling one argument as soon as the other has done a visible action. In order for this process to start from a state rather than a probability distribution of states, we require its arguments to be state-based as well; the same requirement applies to $|_A$.

Finally, the expression $s \;|_A\; t$, where $A \subseteq \mathsf{Act}$, represents processes s and t running in parallel. They may synchronise by performing the same action from A simultaneously; such a synchronisation results in τ. In addition s and t may independently do any action from $(\mathsf{Act}\backslash A) \cup \{\tau\}$.

Although formally the operators $\Box$ and $|_A$ can only be applied to state-based processes, informally we use expressions of the form $P \mathbin{\Box} Q$ and $P \;|_A\; Q$, where P and Q are *not* state-based, as syntactic sugar for expressions in the above syntax obtained by distributing $\Box$ and $|_A$ over ${}_p\!\oplus$. Thus for example $s \mathbin{\Box} (t_1 \;{}_p\!\oplus\; t_2)$ abbreviates the term $(s \mathbin{\Box} t_1) \;{}_p\!\oplus\; (s \mathbin{\Box} t_2)$.

The full language of CSP [1, 10, 21] has many more operators; we have simply chosen a representative selection, and have added probabilistic choice. Our parallel operator is not a CSP primitive, but it can easily be expressed in terms of them — in particular $P \;|_A\; Q = (P\|_A Q)\backslash A$, where $\|_A$ and $\backslash A$ are the parallel composition and hiding operators of [21]. It can also be expressed in terms of the parallel composition, renaming and restriction operators of CCS. We have chosen this (non-associative) operator for convenience in defining the application of tests to processes.

As usual we may elide $\mathbf{0}$; the prefixing operator $a.\_$ binds stronger than any binary operator; and precedence between binary operators is indicated via brackets or spacing. We will also sometimes use indexed binary operators, such as $\bigoplus_{i\in I} p_i{\cdot}P_i$ with $\sum_{i\in I} p_i = 1$ and all $p_i > 0$, and $\Box_{i\in I} P_i$, for some finite index set I.

Our language is interpreted as a *probabilistic labelled transition system* [4, 2]. Essentially the same model has appeared in the literature under different names such as *NP-systems* [12], *probabilistic processes* [13], *simple probabilistic automata* [23], *probabilistic transition systems* [14] etc. Furthermore, there are strong structural similarities with *Markov Decision Processes* [22, 5].

We now fix some notation. A (discrete) probability *subdistribution* over a set S is a function $\Delta : S \to [0,1]$ with $\sum_{s\in S} \Delta(s) \leq 1$; the *support* of such a Δ is $\lceil \Delta \rceil := \{\, s \in S \mid \Delta(s) > 0 \,\}$, and its *mass* $|\Delta|$ is $\sum_{s\in\lceil\Delta\rceil} \Delta(s)$. A subdistribution is a (total, or full) *distribution* if $|\Delta| = 1$. The point distribution $\overline{s}$ assigns probability 1 to s and 0 to all other elements of S, so that $\lceil \overline{s} \rceil = \{s\}$. With $\mathscr{D}(S)$ we denote the set of subdistributions over S, and with $\mathscr{D}_1(S)$ its subset of full distributions. For $\Delta, \Theta \in \mathscr{D}(S)$ we write $\Delta \leq \Theta$ iff $\Delta(s) \leq \Theta(s)$ for all $s \in S$.

Let $\{\Delta_k \mid k \in K\}$ be a set of subdistributions, possibly infinite. Then $\sum_{k\in K} \Delta_k$ is the real-valued function in $S \to \mathbb{R}$ defined by $(\sum_{k\in K} \Delta_k)(s) := \sum_{k\in K} \Delta_k(s)$. This is a partial operation on subdistributions because for some state s the sum of $\Delta_k(s)$ might exceed 1. If the index set is finite, say $\{1..n\}$, we often write $\Delta_1 + \ldots + \Delta_n$. For p a real number from $[0,1]$ we use $p{\cdot}\Delta$ to denote the subdistribution given by $(p{\cdot}\Delta)(s) := p{\cdot}\Delta(s)$. Finally we use ε to denote the everywhere-zero subdistribution that thus has empty support. These operations on subdistributions do not readily adapt themselves to distributions; yet if $\sum_{k\in K} p_k = 1$ for some collection of $p_k \geq 0$, and the Δ_k are distributions, then so is $\sum_{k\in K} p_k{\cdot}\Delta_k$. In general when $0{\leq}p{\leq}1$ we write $x \;{}_p\!\oplus\; y$ for $p{\cdot}x + (1{-}p){\cdot}y$ where that makes sense, so that for example $\Delta_1 \;{}_p\!\oplus\; \Delta_2$ is always defined, and is full if Δ_1 and Δ_2 are.

The expected value $\sum_{s\in S} \Delta(s) \cdot f(s)$ over a (sub)distribution Δ of a bounded non-negative function f to the reals or tuples of them, or to $\mathscr{D}(S)$, is written $\mathrm{Exp}_\Delta(f)$, and the image of a (sub)distribution Δ through a function f is written $\mathrm{Img}_f(\Delta)$ — the latter is the (sub)distribution over the range of f given by $\mathrm{Img}_f(\Delta)(t) := \sum_{f(s)=t} \Delta(s)$.

Definition 2.1 A *probabilistic labelled transition system* (pLTS) is a triple $\langle S, L, \to \rangle$, where

(i) S is a set of states,

(ii) L is a set of transition labels,

(iii) relation $\to$ is a subset of $S \times L \times \mathscr{D}_1(S)$.

A (non-probabilistic) labelled transition system (LTS) may be viewed as a degenerate pLTS — one in which only point distributions are used. As with LTSs, we write $s \xrightarrow{\alpha} \Delta$ for $(s, \alpha, \Delta) \in {\to}$, as well as $s \xrightarrow{\alpha}$ for $\exists \Delta : s \xrightarrow{\alpha} \Delta$ and $s \to$ for $\exists \alpha : s \xrightarrow{\alpha}$. A pLTS is *finitely branching* if the set $\{\langle \alpha, \Delta \rangle \mid s \xrightarrow{\alpha} \Delta,\ \alpha \in L\}$ is finite for all states s; if moreover S is finite, then the pLTS is *finitary*. A pLTS is *deterministic* if for each state s and label α, there is at most one distribution Δ with $s \xrightarrow{\alpha} \Delta$.

The operational semantics of pCSP is defined by a particular pLTS $\langle \mathsf{sCSP}, \mathsf{Act}_\tau, \to \rangle$ in which sCSP is the set of states and $\mathsf{Act}_\tau := \mathsf{Act} \cup \{\tau\}$ is the set of transition labels; we let a range over Act and α over Act_τ. We interpret

(ACTION)
$$a.P \xrightarrow{a} [\![P]\!]$$

(RECURSION)
$$\mathsf{rec}\, x.\, P \xrightarrow{\tau} [\![P[x \mapsto \mathsf{rec}\, x.\, P]]\!]$$

(INT.L)
$$P \sqcap Q \xrightarrow{\tau} [\![P]\!]$$

(INT.R)
$$P \sqcap Q \xrightarrow{\tau} [\![Q]\!]$$

(EXT.L)
$$\frac{s_1 \xrightarrow{a} \Delta}{s_1 \,\Box\, s_2 \xrightarrow{a} \Delta}$$

(EXT.R)
$$\frac{s_2 \xrightarrow{a} \Delta}{s_1 \,\Box\, s_2 \xrightarrow{a} \Delta}$$

(EXT.I.L)
$$\frac{s_1 \xrightarrow{\tau} \Delta}{s_1 \,\Box\, s_2 \xrightarrow{\tau} \Delta \,\Box\, s_2}$$

(EXT.I.R)
$$\frac{s_2 \xrightarrow{\tau} \Delta}{s_1 \,\Box\, s_2 \xrightarrow{\tau} s_1 \,\Box\, \Delta}$$

(PAR.L)
$$\frac{s_1 \xrightarrow{\alpha} \Delta}{s_1 \mid_A s_2 \xrightarrow{\alpha} \Delta \mid_A s_2} \quad \alpha \notin A$$

(PAR.R)
$$\frac{s_2 \xrightarrow{\alpha} \Delta}{s_1 \mid_A s_2 \xrightarrow{\alpha} s_1 \mid_A \Delta} \quad \alpha \notin A$$

(PAR.I)
$$\frac{s_1 \xrightarrow{a} \Delta_1,\ s_2 \xrightarrow{a} \Delta_2}{s_1 \mid_A s_2 \xrightarrow{\tau} \Delta_1 \mid_A \Delta_2} \quad a \in A$$

In the above inferences A ranges over subsets of Act, and actions a, α are elements of Act, Act_τ respectively.

Figure 2: Operational semantics of pCSP

pCSP processes P as distributions $[\![P]\!] \in \mathscr{D}_1(\mathsf{sCSP})$ via the function $[\![\_]\!] : \mathsf{pCSP} \to \mathscr{D}_1(\mathsf{sCSP})$ defined by

$$[\![s]\!] \;:=\; \overline{s} \quad \text{for } s \in \mathsf{sCSP}, \quad \text{and} \quad [\![P\,{}_p\!\oplus Q]\!] \;:=\; [\![P]\!]\,{}_p\!\oplus [\![Q]\!] \;.$$

The transition relation $\rightarrow$ is defined in Figure 2. This is a slight extension of the rules we used earlier [4, 2] for finite processes: one new rule is required to interpret recursive processes. All rules are very similar to the standard ones used to interpret CSP as a labelled transition system [21], but are modified so that the result of an action is a distribution. The rules for external choice and parallel composition use an obvious notation for distributing an operator over a distribution; for example $\Delta \,\Box\, s$ represents the distribution given by

$$(\Delta \,\Box\, s)(t) = \begin{cases} \Delta(s') & \text{if } t = s' \,\Box\, s \\ 0 & \text{otherwise.} \end{cases}$$

We sometimes write $\tau.P$ for $P \sqcap P$, thus giving $\tau.P \xrightarrow{\tau} [\![P]\!]$.

The set of states *reachable* from a subdistribution Δ is the smallest set that contains $\lceil \Delta \rceil$ and is closed under transitions, meaning that if some state s is reachable and $s \xrightarrow{\alpha} \Theta$ then every state in $\lceil \Theta \rceil$ is reachable as well. We graphically depict the operational semantics of a pCSP expression P by drawing the part of the pLTS reachable from $[\![P]\!]$ as a directed graph with states represented by filled nodes $\bullet$ and distributions by open nodes $\circ$. For any state s and distribution Δ with $s \xrightarrow{\alpha} \Delta$ we draw an edge from s to Δ labelled with α; and for any distribution Δ and state s in $\lceil \Delta \rceil$, the support of Δ, we draw an edge from Δ to s labelled with $\Delta(s)$. We often leave out point-distributions—diverting an incoming edge to the unique state in its support. Sometimes we partially unfold this graph by drawing the same nodes multiple times; in doing so, all outgoing edges of a given instance of a node are always drawn, but not necessarily all incoming edges.

Note that for each $P \in \mathsf{pCSP}$ the distribution $[\![P]\!]$ has finite support. Moreover, our pLTS is *finitely branching* in the sense that for each state $s \in \mathsf{sCSP}$ there are only finitely many pairs $(\alpha, \Delta) \in \mathsf{Act}_\tau \times \mathscr{D}_1(\mathsf{sCSP}))$ with $s \xrightarrow{\alpha} \Delta$. In spite of $[\![P]\!]$'s finite support, and the finite branching of our pLTS, it is possible for there to be infinitely many states reachable from $[\![P]\!]$; when there are only finitely many, then P is said to be finitary [5].

Definition 2.2 A subdistribution $\Delta \in \mathscr{D}(S)$ in a pLTS $\langle S, L, \rightarrow \rangle$ is *finitary* if only finitely many states are reachable from Δ; a pCSP expression P is *finitary* if $[\![P]\!]$ is.

3 A novel approach to weak derivations

In this section we develop a new definition of what it means for a recursive process to evolve by silent activity into another process; it allows the simulation and failure-simulation preorders of [2] to be adapted to characterise the testing preorders for at least finitary probabilistic processes.

Recall for example the process Q_1 defined in the introduction. It turns out that in our testing framework this process is indistinguishable from a: both processes can do nothing else than an a-action, possibly after some internal moves, and in both cases the probability that the process will never do the a-action is 0. In [4, 2], where we did not deal with recursive processes like Q_1, we defined a weak transition relation $\stackrel{\hat{a}}{\Longrightarrow}$ in such a way that $P \stackrel{\hat{a}}{\Longrightarrow}$ iff there is a finite number of τ-moves after which the entire distribution $[\![P]\!]$ will have done an a-action. Lifting this definition verbatim to a setting with recursion would create a difference between a and Q_1, for only the former admits such a weak transition $\stackrel{\hat{a}}{\Longrightarrow}$. The purpose of this section is to propose a new definition of weak transitions, with which we can capture the intuition that the process Q_1 can perform the action a with probability 1, provided it is allowed to run for an unbounded amount of time.

We construct our generalised definition of weak move by revising what it means for a probabilistic process to execute an indefinite sequence of (internal) τ moves. The key technical innovation is to change the focus from distributions to *subdistributions* that enable us to express divergence very conveniently.[2]

First some relatively standard terminology. For any subset X of $\mathscr{D}(S)$, with S a set, let $\updownarrow X$, the *convex closure* of X, be the smallest convex set containing X. So it satisfies:

(i) $X \subseteq \updownarrow X$

(ii) $\Delta \in \updownarrow X$ if and only if $\Delta = \sum_{i\in I} p_i \cdot \Delta_i$, where $\Delta_i \in X$ and $p_i \in [0,1]$, for some index set I such that $\sum_{i\in I} p_i = 1$.

In case S is a finite set, it makes no difference whether we restrict I to being finite or not~~; in fact, index sets of size 2 will suffice~~. In fact, requiring I to be finite is equivalent to defining convexity of a set Y by $\Delta, \Theta \in Y \Rightarrow \Delta\,{}_p\!\oplus \Theta \in Y$ for any $p \in [0,1]$. However, in general there is a difference:

Example 3.1 Let $S = \{s_i \mid i \in \mathbb{N}\}$. Then $\updownarrow\{\overline{s_i} \mid i \in \mathbb{N}\}$ consists of all total distributions whose support is included in S. However, with a definition of convex closure that requires only binary interpolations of distributions to be included, $\updownarrow\{\overline{s_i} \mid i \in \mathbb{N}\}$ would merely consist of all such distributions with finite support. □

Convex closure is a closure operator in the standard sense, in that it satisfies

- $X \subseteq \updownarrow X$
- $X \subseteq Y$ implies $\updownarrow X \subseteq \updownarrow Y$
- $\updownarrow\updownarrow X = \updownarrow X$.

We say a set X is *convex* if $\updownarrow X = X$. Furthermore, we say that a relation $\mathscr{R} \subseteq Y \times \mathscr{D}(S)$ is convex whenever the set $\{\Delta \mid y \mathrel{\mathscr{R}} \Delta\}$ is convex for every y in Y, and $\updownarrow\mathscr{R}$ denotes the smallest convex relation containing $\mathscr{R}$.

3.1 Lifted relations

In a pLTS actions are only performed by states, in that actions are given by relations from states to distributions. But pCSP processes in general correspond to distributions over states, so in order to define what it means for a process to perform an action, we need to *lift* these relations so that they also apply to distributions. In fact we will find it convenient to lift them to subdistributions.

Definition 3.2 (Lifting) Let $\langle S, L, \rightarrow\rangle$ be a pLTS and $\mathscr{R} \subseteq S \times \mathscr{D}(S)$ be a relation from states to subdistributions. Then $\overline{\mathscr{R}} \subseteq \mathscr{D}(S) \times \mathscr{D}(S)$ is the smallest relation that satisfies:

(1) $s \mathrel{\mathscr{R}} \Theta$ implies $\overline{s} \mathrel{\overline{\mathscr{R}}} \Theta$, and

(2) (Linearity) $\Delta_i \mathrel{\overline{\mathscr{R}}} \Theta_i$ for $i \in I$ implies $(\sum_{i\in I} p_i \cdot \Delta_i) \mathrel{\overline{\mathscr{R}}} (\sum_{i\in I} p_i \cdot \Theta_i)$ for any $p_i \in [0,1]$ $(i \in I)$ with $\sum_{i\in I} p_i \le 1$.

[2]Subdistributions' nice properties with respect to divergence are due to their being equivalent to the discrete probabilistic powerdomain over a flat domain [11].

Remark 3.3 For $\mathscr{R}_1, \mathscr{R}_2 \subseteq S \times \mathscr{D}(S)$, if $\mathscr{R}_1 \subseteq \mathscr{R}_2$ then $\overline{\mathscr{R}_1} \subseteq \overline{\mathscr{R}_2}$.

Remark 3.4 By construction $\overline{\mathscr{R}}$ is convex. Moreover, because $s(\updownarrow\mathscr{R})\Theta$ implies $\overline{s}\ \overline{\mathscr{R}}\ \Theta$ we have $\overline{\mathscr{R}}=\overline{\updownarrow\mathscr{R}}$, which means that when considering a lifted relation we can without loss of generality assume the original relation to have been convex. In fact when $\mathscr{R}$ is indeed convex, we have that $\overline{s}\ \overline{\mathscr{R}}\ \Theta$ and $s\ \mathscr{R}\ \Theta$ are equivalent.

An application of this notion is when the relation is $\xrightarrow{\alpha}$ for $\alpha \in \mathsf{Act}_\tau$; in that case we also write $\xrightarrow{\alpha}$ for $\overline{\xrightarrow{\alpha}}$. Thus, as source of a relation $\xrightarrow{\alpha}$ we now also allow distributions, and even subdistributions. A subtlety of this approach is that for any action α, we have

$$\varepsilon \xrightarrow{\alpha} \varepsilon \tag{1}$$

simply by taking $I = \emptyset$ or $\sum_{i \in I} p_i = 0$ in Definition 3.2. That will turn out to make ε especially useful for modelling the "chaotic" aspects of divergence, in particular that in the must-case a divergent process can mimic any other.

Definition 3.2 is very similar to our previous definition in [2], although there it applied only to (full) distributions:

Lemma 3.5 $\Delta\ \overline{\mathscr{R}}\ \Theta$ if and only if

(i) $\Delta = \sum_{i \in I} p_i \cdot \overline{s_i}$, where I is an index set and $\sum_{i \in I} p_i \leq 1$,

(ii) For each $i \in I$ there is a subdistribution Θ_i such that $s_i\ \mathscr{R}\ \Theta_i$,

(iii) $\Theta = \sum_{i \in I} p_i \cdot \Theta_i$.

Proof: Straightforward. □

An important point here is that a single state can be split into several pieces: that is, the decomposition of Δ into $\sum_{i \in I} p_i \cdot \overline{s_i}$ is not unique. One important property of this lifting operation is the following:

Lemma 3.6 Suppose $\Delta\ \overline{\mathscr{R}}\ \Theta$, where $\mathscr{R}$ is any relation in $S \times \mathscr{D}(S)$. Then

(i) $|\Delta| \geq |\Theta|$.

(ii) If $\mathscr{R}$ is a relation in $S \times \mathscr{D}_1(S)$ then $|\Delta| = |\Theta|$.

Proof: This follows immediately from the characterisation in Lemma 3.5. □

So for example if $\varepsilon\ \overline{\mathscr{R}}\ \Theta$ then $0 = |\varepsilon| \geq |\Theta|$, whence Θ is also ε.

Remark 3.7 From Lemma 3.5 it also follows that lifting enjoys the following two properties:

(i) (Scaling) If $\Delta\ \overline{\mathscr{R}}\ \Theta$, $p \in \mathbb{R}$ and $|p \cdot \Delta| \leq 1$ then $p \cdot \Delta\ \overline{\mathscr{R}}\ p \cdot \Theta$.

(ii) (Additivity) If $\Delta_i\ \overline{\mathscr{R}}\ \Theta_i$ for $i \in I$ and $|\sum_{i \in I} \Delta_i| \leq 1$ then $(\sum_{i \in I} \Delta_i)\ \overline{\mathscr{R}}\ (\sum_{i \in I} \Theta_i)$.

In fact, we could have presented Definition 3.2 using scaling and additivity instead of linearity.

The lifting operation has yet another characterisation, this time in terms of *choice functions*.

Definition 3.8 Let $\mathscr{R} \subseteq S \times \mathscr{D}(S)$ be a relation from states to subdistributions. Then $f : dom(\mathscr{R}) \rightarrow \mathscr{D}(S)$ is a *choice function for* $\mathscr{R}$ if $s\ \mathscr{R}\ f(s)$ for every $s \in dom(\mathscr{R})$. We write $\mathbf{Ch}(\mathscr{R})$ for the set of all choice functions of $\mathscr{R}$.

Proposition 3.9 Suppose $\mathscr{R} \subseteq S \times \mathscr{D}(S)$ is a convex relation. Then for any $\Delta, \Theta \in \mathscr{D}(S)$, $\Delta\ \overline{\mathscr{R}}\ \Theta$ if and only if $\lceil\Delta\rceil \subseteq dom(R)$ and there is some choice function $f \in \mathbf{Ch}(\mathscr{R})$ such that $\Theta = \mathrm{Exp}_\Delta(f)$.

Proof: First suppose $\lceil\Delta\rceil \subseteq dom(R)$ and $\Theta = \mathrm{Exp}_\Delta(f)$ for some choice function $f \in \mathbf{Ch}(\mathscr{R})$, that is $\Theta = \sum_{s \in \lceil\Delta\rceil} \Delta(s) \cdot f(s)$. It now follows from Lemma 3.5 that $\Delta\ \overline{\mathscr{R}}\ \Theta$ since $s\ \mathscr{R}\ f(s)$ for each $s \in \lceil\Delta\rceil$.

Conversely suppose $\Delta\ \overline{\mathscr{R}}\ \Theta$. Applying Lemma 3.5 we know that

(i) $\Delta = \sum_{i \in I} p_i \cdot \overline{s_i}$, for some index set I, with $\sum_{i \in I} p_i \leq 1$

(ii) $\Theta = \sum_{i \in I} p_i \cdot \Theta_i$ for some Θ_i satisfying $s_i\ \mathscr{R}\ \Theta_i$.

First of all, this implies that $\lceil\Delta\rceil \subseteq dom(\mathscr{R})$. Now define the function $f : dom(\mathscr{R}) \rightarrow \mathscr{D}(S)$ as follows:

- if $s \in \lceil\Delta\rceil$ then $f(s) = \sum_{\{ i \in I \mid s_i = s \}} \dfrac{p_i}{\Delta(s)} \cdot \Theta_i$;
- otherwise, $f(s) = \Theta'$ for any Θ' with $s \mathrel{\mathscr{R}} \Theta'$;

Note that $\Delta(s) = \sum_{\{ i \in I \mid s_i = s \}} p_i$ and therefore by convexity $s \mathrel{\mathscr{R}} f(s)$; so f is a choice function for $\mathscr{R}$. Moreover, a simple calculation shows that $\mathrm{Exp}_\Delta(f) = \sum_{i \in I} p_i \cdot \Theta_i$, which by (ii) above is Θ. $\Box$

An important further property is the following:

Proposition 3.10 If $(\sum_{i \in I} p_i \cdot \Delta_i) \mathrel{\overline{\mathscr{R}}} \Theta$ then $\Theta = \sum_{i \in I} p_i \cdot \Theta_i$ for some subdistributions Θ_i such that $\Delta_i \mathrel{\overline{\mathscr{R}}} \Theta_i$ for $i \in I$.

Proof: Let $\Delta \mathrel{\overline{\mathscr{R}}} \Theta$ where $\Delta = \sum_{i \in I} p_i \cdot \Delta_i$. By Proposition 3.9, using that $\overline{\mathscr{R}} = \overline{\updownarrow\mathscr{R}}$, there is a choice function $f \in \mathbf{Ch}(\updownarrow\mathscr{R})$ such that $\Theta = \mathrm{Exp}_\Delta(f)$. Take $\Theta_i := \mathrm{Exp}_{\Delta_i}(f)$ for $i \in I$. Using that $\lceil\Delta_i\rceil \subseteq \lceil\Delta\rceil$, Proposition 3.9 yields $\Delta_i \mathrel{\overline{\mathscr{R}}} \Theta_i$ for $i \in I$. Finally,

$$\sum_{i \in I} p_i \cdot \Theta_i = \sum_{i \in I} p_i \cdot \sum_{s \in \lceil\Delta_i\rceil} \Delta_i(s) \cdot f(s) = \sum_{s \in \lceil\Delta\rceil} \sum_{i \in I} p_i \cdot \Delta_i(s) \cdot f(s) = \sum_{s \in \lceil\Delta\rceil} \Delta(s) \cdot f(s) = \mathrm{Exp}_\Delta(f) = \Theta. \qquad \Box$$

The converse to the above is not true in general: from $\Delta \mathrel{\overline{\mathscr{R}}} (\sum_{i \in I} p_i \cdot \Theta_i)$ it does not follow that Δ can correspondingly be decomposed. For example, we have $\overline{a.(b \,{}_{\frac{1}{2}}\!\oplus c)} \xrightarrow{a} \frac{1}{2} \cdot \overline{b} + \frac{1}{2} \cdot \overline{c}$, yet $\overline{a.(b \,{}_{\frac{1}{2}}\!\oplus c)}$ cannot be written as $\frac{1}{2} \cdot \Delta_1 + \frac{1}{2} \cdot \Delta_2$ such that $\Delta_1 \xrightarrow{a} b$ and $\Delta_2 \xrightarrow{a} c$.

A simplified form of Proposition 3.10 holds for unlifted relations, provided they are convex:

Corollary 3.11 If $(\sum_{i \in I} p_i \cdot \overline{s_i}) \mathrel{\overline{\mathscr{R}}} \Theta$ and $\mathscr{R}$ is convex, then $\Theta = \sum_{i \in I} p_i \cdot \Theta_i$ for subdistributions Θ_i with $s_i \mathrel{\mathscr{R}} \Theta_i$ for $i \in I$.

Proof: Take Δ_i to be $\overline{s_i}$ in Proposition 3.10, whence $\Theta = \sum_{i \in I} p_i \cdot \Theta_i$ for some subdistributions Θ_i such that $\overline{s_i} \mathrel{\overline{\mathscr{R}}} \Theta_i$ for $i \in I$. Because $\mathscr{R}$ is convex, we then have $s_i \mathrel{\mathscr{R}} \Theta_i$ from Remark 3.4. $\Box$

Lifting satisfies the following monadic property with respect to composition.

Lemma 3.12 Let $\mathscr{R}_1, \mathscr{R}_2 \subseteq S \times \mathscr{D}(S)$. Then the forward relational composition $\overline{\mathscr{R}_1};\overline{\mathscr{R}_2}$ is equal to the lifted composition $\overline{\mathscr{R}_1;\overline{\mathscr{R}_2}}$.

Proof: Suppose $\Delta \mathrel{\overline{\mathscr{R}_1};\overline{\mathscr{R}_2}} \Phi$. Then there is some Θ such that $\Delta \mathrel{\overline{\mathscr{R}_1}} \Theta \mathrel{\overline{\mathscr{R}_2}} \Phi$. By Lemma 3.5 we have the decomposition $\Delta = \sum_{i \in I} p_i \cdot \overline{s_i}$ and $\Theta = \sum_{i \in I} p_i \cdot \Theta_i$ with $s_i \mathrel{\mathscr{R}_1} \Theta_i$ for each $i \in I$. By Proposition 3.10 we obtain $\Phi = \sum_{i \in I} p_i \cdot \Phi_i$ with $\Theta_i \mathrel{\overline{\mathscr{R}_2}} \Phi_i$. It follows that $s_i \mathrel{\mathscr{R}_1;\overline{\mathscr{R}_2}} \Phi_i$, and thus $\Delta \mathrel{\overline{\mathscr{R}_1;\overline{\mathscr{R}_2}}} \Phi$. So we have shown that $\overline{\mathscr{R}_1};\overline{\mathscr{R}_2} \subseteq \overline{\mathscr{R}_1;\overline{\mathscr{R}_2}}$. The other direction can be proved similarly. $\Box$

3.2 Weak derivations

We now formally define a notion of weak derivatives.

Definition 3.13 (Weak τ moves to derivatives) Suppose we have subdistributions $\Delta,\ \Delta_k,\ \Delta_k^{\rightarrow}, \Delta_k^{\times}$, for $k \geq 0$, with the following properties:

$$\begin{array}{lclcl}
\Delta & = & \Delta_0 = \Delta_0^{\rightarrow} & + \Delta_0^{\times} & \text{— The } \times \text{ component stops “here” (even if it could have moved),} \\
\Delta_0^{\rightarrow} & \xrightarrow{\tau} & \Delta_1 = \Delta_1^{\rightarrow} & + \Delta_1^{\times} & \text{— but the } \rightarrow \text{ component moves on.} \\
\vdots & & & \vdots & \\
\Delta_k^{\rightarrow} & \xrightarrow{\tau} & \Delta_{k+1} = \Delta_{k+1}^{\rightarrow} & + \Delta_{k+1}^{\times} & \\
 & & & \vdots & \\
\hline
\text{In total:} & & \Delta' & = \sum_{k=0}^{\infty} \Delta_k^{\times} & \text{— Finally, all the stopped-somewhere components are summed.}
\end{array}$$

The $\xrightarrow{\tau}$ moves above with subdistribution sources are lifted in the sense of the previous section.

Then we call $\Delta' := \sum_{k=0}^{\infty} \Delta_k^{\times}$ a *weak derivative* of Δ, and write $\Delta \Longrightarrow \Delta'$ to mean that Δ can make a *weak τ move* to its derivative Δ'.

There is always at least one derivative of any distribution (the distribution itself) and there can be many. Using Lemma 3.6 it is easily checked that Definition 3.13 is well-defined in that derivatives do not sum to more than one.

Example 3.14 Let $\xrightarrow{\tau}{}^{\star}$ denote the reflexive transitive closure of the relation $\xrightarrow{\tau}$ over subdistributions. By the judicious use of the empty distribution ε in the definition of $\Longrightarrow$, and property (1) above, it is easy to see that

$$\Delta \xrightarrow{\tau}{}^{\star} \Theta \quad \text{implies} \quad \Delta \Longrightarrow \Theta$$

because $\Delta \xrightarrow{\tau}{}^{\star} \Theta$ means the existence of a finite sequence of subdistributions $\Delta = \Delta_0,\ \Delta_1, \ldots, \Delta_k = \Theta,\ k \geq 0$ for which we can write

$$\begin{array}{lcl}
\Delta & = & \Delta_0 + \varepsilon \\
\Delta_0 & \xrightarrow{\tau} & \Delta_1 + \varepsilon \\
\vdots & & \quad\ \vdots \\
\Delta_{k-1} & \xrightarrow{\tau} & \quad \varepsilon + \Delta_k \\
\varepsilon & \xrightarrow{\tau} & \quad \varepsilon + \varepsilon \\
 & & \qquad \vdots \\ \hline
 & \text{In total:} & \qquad \Theta\ .
\end{array}$$

This implies that $\Longrightarrow$ is indeed a generalisation of the standard notion for non-probabilistic transition systems of performing an indefinite sequence of internal τ moves. □

In [4, 2] we wrote $s \xrightarrow{\hat{\tau}} \Delta$ if either $s \xrightarrow{\tau} \Delta$ or $\Delta = \overline{s}$. Hence the lifted relation $\xrightarrow{\hat{\tau}}$ satisfies $\Delta \xrightarrow{\hat{\tau}} \Delta'$ iff there are $\Delta^{\rightarrow}$, $\Delta^{\times}$ and Δ_1 such that $\Delta = \Delta_{\rightarrow} + \Delta^{\times}$, $\Delta^{\rightarrow} \xrightarrow{\tau} \Delta_1$ and $\Delta' = \Delta_1 + \Delta^{\times}$. Clearly, $\Delta \xrightarrow{\hat{\tau}} \Delta'$ implies $\Delta \Longrightarrow \Delta'$. With a little effort, one can also show that $\Delta \xrightarrow{\hat{\tau}}{}^{\star} \Delta'$ implies $\Delta \Longrightarrow \Delta'$. In fact, this follows directly from the reflexivity and transitivity of $\Longrightarrow$; the latter will be established in Theorem 3.22.

Conversely, in [4, 2] we dealt with recursion-free pCSP processes P, and these have the property that in a sequence as in Definition 3.13 with $\Delta = [\![P]\!]$ we necessarily have that $\Delta_k = \varepsilon$ for some $k \geq 0$. On such processes we have that the relations $\xrightarrow{\hat{\tau}}{}^{\star}$ and $\Longrightarrow$ coincide.

In Definition 3.13 we can see that $\Delta' = \varepsilon$ iff $\Delta_k^{\times} = \varepsilon$ for all k. Thus $\Delta \Longrightarrow \varepsilon$ iff there is an infinite sequence of subdistributions Δ_k such that $\Delta = \Delta_0$ and $\Delta_k \xrightarrow{\tau} \Delta_{k+1}$, that is Δ can give rise to a divergent computation.

Example 3.15 Consider the process $\mathsf{rec}\, x.\, x$, which recall is a state, and for which we have $\mathsf{rec}\, x.\, x \xrightarrow{\tau} [\![\mathsf{rec}\, x.\, x]\!]$ and thus $[\![\mathsf{rec}\, x.\, x]\!] \xrightarrow{\tau} [\![\mathsf{rec}\, x.\, x]\!]$. Then $[\![\mathsf{rec}\, x.\, x]\!] \Longrightarrow \varepsilon$. □

Example 3.16 Recall the process $Q_1 = \mathsf{rec}\, x.\, (\tau.x \,{}_{\frac{1}{2}}{\oplus}\, a)$ from the introduction. We have $[\![Q_1]\!] \Longrightarrow [\![a]\!]$ because

$$\begin{array}{rcl}
[\![Q_1]\!] & = & [\![Q_1]\!] + \varepsilon \\
[\![Q_1]\!] & \xrightarrow{\tau} & \frac{1}{2} \cdot [\![\tau.Q_1]\!] + \frac{1}{2} \cdot [\![a]\!] \\
\frac{1}{2} \cdot [\![\tau.Q_1]\!] & \xrightarrow{\tau} & \frac{1}{2} \cdot [\![Q_1]\!] + \varepsilon \\
\frac{1}{2} \cdot [\![Q_1]\!] & \xrightarrow{\tau} & \frac{1}{2^2} \cdot [\![\tau.Q_1]\!] + \frac{1}{2^2} \cdot [\![a]\!] \\
\ldots & & \\
\frac{1}{2^k} \cdot [\![Q_1]\!] & \xrightarrow{\tau} & \frac{1}{2^{k+1}} \cdot [\![\tau.Q_1]\!] + \frac{1}{2^{k+1}} \cdot [\![a]\!] \\
\ldots & &
\end{array}$$

which means that by definition we have

$$[\![Q_1]\!] \Longrightarrow \varepsilon + \sum_{k \geq 1} \frac{1}{2^k} \cdot [\![a]\!]$$

thus generating the weak derivative $[\![a]\!]$ as claimed. □

Example 3.17 Consider the (infinite) collection of states s_k and probabilities p_k for $k \geq 2$ such that

$$s_k \xrightarrow{\tau} [\![a]\!]\,_{p_k}\!\oplus \overline{s_{k+1}}\ ,$$

where we choose p_k so that starting from any s_k the probability of eventually taking a left-hand branch, and so reaching $[\![a]\!]$ ultimately, is just $\frac{1}{k}$ in total. Thus p_k must satisfy $\frac{1}{k} = p_k + (1{-}p_k)\frac{1}{k+1}$, whence by arithmetic we have that $p_k := \frac{1}{k^2}$ will do. Therefore in particular $s_2 \Longrightarrow \frac{1}{2}[\![a]\!]$, with the remaining $\frac{1}{2}$ lost in divergence. □

Our final example demonstrates that derivatives of (interpretations of) pCSP processes may have infinite support, and hence that we can have $[\![P]\!] \Longrightarrow \Delta'$ such that there is no $P' \in \mathsf{pCSP}$ with $[\![P']\!] = \Delta'$.

Example 3.18 Let P denote the process $\mathsf{rec}\, x.\, b\,_{\frac{1}{2}}\!\oplus (x \mid_\emptyset \mathbf{0})$. Then we have the derivation:

$$\begin{array}{rcl}
[\![P]\!] & = & [\![P]\!] + \varepsilon \\
[\![P]\!] & \xrightarrow{\tau} & \dfrac{1}{2}\cdot[\![P \mid_\emptyset \mathbf{0}^1]\!] + \dfrac{1}{2}\cdot[\![b]\!] \\
\dfrac{1}{2}\cdot[\![P \mid_\emptyset \mathbf{0}^1]\!] & \xrightarrow{\tau} & \dfrac{1}{2^2}\cdot[\![P \mid_\emptyset \mathbf{0}^2]\!] + \dfrac{1}{2^2}\cdot[\![b \mid_\emptyset \mathbf{0}^1]\!] \\
\dots & & \dots \\
\dfrac{1}{2^k}\cdot[\![P \mid_\emptyset \mathbf{0}^k]\!] & \xrightarrow{\tau} & \dfrac{1}{2^{k+1}}\cdot[\![P \mid_\emptyset \mathbf{0}^{k+1}]\!] + \dfrac{1}{2^{k+1}}\cdot[\![b \mid_\emptyset \mathbf{0}^k]\!] \\
\dots & &
\end{array}$$

where $\mathbf{0}^k$ represents k instances of $\mathbf{0}$ running in parallel. This implies that

$$[\![P]\!] \Longrightarrow \Theta$$

where

$$\Theta = \sum_{k \geq 0} \frac{1}{2^{k+1}}\cdot[\![b \mid_\emptyset \mathbf{0}^k]\!]$$

a distribution with infinite support. □

3.3 Properties of weak derivations

Here we develop some properties of the weak move relation $\Longrightarrow$ which will be important later on in the paper. We wish to use weak derivation as much as possible in the same way as the lifted action relations $\xrightarrow{\alpha}$, and therefore we start with showing that $\Longrightarrow$ enjoys two of the most crucial properties of $\xrightarrow{\alpha}$: linearity of Definition 3.2 and the decomposition property of Proposition 3.10. To this end, we first establish that weak derivations do not increase the mass of distributions, and are preserved under scaling.

Lemma 3.19 For any subdistributions $\Delta, \Theta, \Gamma, \Lambda, \Pi$ we have

(i) If $\Delta \Longrightarrow \Theta$ then $|\Delta| \geq |\Theta|$.

(ii) If $\Delta \Longrightarrow \Theta$ and $p \in \mathbb{R}$ such that $|p\cdot\Delta| \leq 1$, then $p\cdot\Delta \Longrightarrow p\cdot\Theta$.

(iii) If $\Gamma + \Lambda \Longrightarrow \Pi$ then $\Pi = \Pi^\Gamma + \Pi^\Lambda$ with $\Gamma \Longrightarrow \Pi^\Gamma$ and $\Lambda \Longrightarrow \Pi^\Lambda$.

Proof: By definition $\Delta \Longrightarrow \Theta$ means that some $\Delta_k, \Delta_k^\times, \Delta_k^\rightarrow$ exist for all $k \geq 0$ such that

$$\Delta = \Delta_0, \qquad \Delta_k = \Delta_k^\times + \Delta_k^\rightarrow, \qquad \Delta_k^\rightarrow \xrightarrow{\tau} \Delta_{k+1}, \qquad \Theta = \sum_{k=0}^{\infty} \Delta_k^\times .$$

A simple inductive proof shows that

$$|\Delta| \quad = \quad |\Delta_i^\rightarrow| + \sum_{k \leq i} |\Delta_k^\times| \text{ for any } i \geq 0. \tag{2}$$

The sequence $\{\sum_{k\leq i}|\Delta_k^\times|\}_{i=0}^\infty$ is nondecreasing and by (2) each element of the sequence is not greater than $|\Delta|$. Therefore, the limit of this sequence is bounded by $|\Delta|$. That is,

$$|\Delta| \quad \geq \quad \lim_{i\to\infty}\sum_{k\leq i}|\Delta_k^\times| \quad = \quad |\Theta|.$$

Now suppose $p\in\mathbb{R}$ such that $|p\cdot\Delta|\leq 1$. From Remark 3.7(i) it follows that

$$p\cdot\Delta = p\cdot\Delta_0, \qquad p\cdot\Delta_k = p\cdot\Delta_k^\rightarrow + p\cdot\Delta_k^\times, \qquad p\cdot\Delta_k^\rightarrow \xrightarrow{\tau} p\cdot\Delta_{k+1}, \qquad p\cdot\Theta = \sum_k p\cdot\Delta_k^\times.$$

Hence Definition 3.13 yields $p\cdot\Delta \Longrightarrow p\cdot\Theta$.

Next suppose $\Gamma+\Lambda\Longrightarrow\Pi$. By Definition 3.13 there are subdistributions $\Pi_k,\Pi_k^\rightarrow,\Pi_k^\times$ for $k\in\mathbb{N}$ such that

$$\Gamma+\Lambda=\Pi_0, \qquad \Pi_k=\Pi_k^\rightarrow+\Pi_k^\times, \qquad \Pi_k^\rightarrow\xrightarrow{\tau}\Pi_{k+1}, \qquad \Pi=\sum_k\Pi_k^\times.$$

For any $s\in S$, define

$$\begin{array}{rcl}\Gamma_0^\rightarrow(s) & := & \min(\Gamma(s),\Pi_0^\rightarrow(s))\\ \Gamma_0^\times(s) & := & \Gamma(s)-\Gamma_0^\rightarrow(s)\\ \Lambda_0^\times(s) & := & \min(\Lambda(s),\Pi_0^\times(s))\\ \Lambda_0^\rightarrow(s) & := & \Lambda(s)-\Lambda_0^\times(s)\ ,\end{array} \tag{3}$$

and check that $\Gamma_0^\rightarrow+\Gamma_0^\times=\Gamma$ and $\Lambda_0^\rightarrow+\Lambda_0^\times=\Lambda$. To show that $\Lambda_0^\rightarrow+\Gamma_0^\rightarrow=\Pi_0^\rightarrow$ and $\Lambda_0^\times+\Gamma_0^\times=\Pi_0^\times$ we fix a state s and distinguish two cases: either (a) $\Pi_0^\rightarrow(s)\geq\Gamma(s)$ or (b) $\Pi_0^\rightarrow(s)<\Gamma(s)$. In Case (a) we have $\Pi_0^\times(s)\leq\Lambda(s)$ and the definitions (3) simplify to $\Gamma_0^\rightarrow(s)=\Gamma(s)$, $\Gamma_0^\times(s)=0$, $\Lambda_0^\times(s)=\Pi_0^\times(s)$ and $\Lambda_0^\rightarrow(s)=\Lambda(s)-\Pi_0^\times(s)$, whence immediately $\Gamma_0^\rightarrow(s)+\Lambda_0^\rightarrow(s)=\Pi_0^\rightarrow(s)$ and $\Gamma_0^\times(s)+\Lambda_0^\times(s)=\Pi_0^\times(s)$. Case (b) is similar.

Since $\Lambda_0^\rightarrow+\Gamma_0^\rightarrow\xrightarrow{\tau}\Pi_1$, by Proposition 3.10 we find Γ_1,Λ_1 with $\Gamma_0^\rightarrow\xrightarrow{\tau}\Gamma_1$ and $\Lambda_0^\rightarrow\xrightarrow{\tau}\Lambda_1$ and $\Pi_1=\Gamma_1+\Lambda_1$. Being now in the same position with Π_1 as we were with Π_0, we can continue this procedure to find Λ_k, Γ_k, $\Lambda_k^\rightarrow$, $\Gamma_k^\rightarrow$, $\Lambda_k^\times$ and $\Gamma_k^\times$ with

$$\begin{array}{ccc}\Gamma=\Gamma_0, & \Gamma_k=\Gamma_k^\rightarrow+\Gamma_k^\times, & \Gamma_k^\rightarrow\xrightarrow{\tau}\Gamma_{k+1},\\ \Lambda=\Lambda_0, & \Lambda_k=\Lambda_k^\rightarrow+\Lambda_k^\times, & \Lambda_k^\rightarrow\xrightarrow{\tau}\Lambda_{k+1},\\ \Gamma_k+\Lambda_k=\Pi_k, & \Gamma_k^\rightarrow+\Lambda_k^\rightarrow=\Pi_k^\rightarrow, & \Gamma_k^\times+\Lambda_k^\times=\Pi_k^\times.\end{array}$$

Let $\Pi^\Gamma:=\sum_k\Gamma_k^\times$ and $\Pi^\Lambda:=\sum_k\Lambda_k^\times$. Then $\Pi=\Pi^\Gamma+\Pi^\Lambda$ and Definition 3.13 yields $\Gamma\Longrightarrow\Pi^\Gamma$ and $\Lambda\Longrightarrow\Pi^\Lambda$. □

Together, Lemma 3.19(ii) and (iii) imply the binary counterpart of the decomposition property of Proposition 3.10. We now generalise this result to infinite (but still countable) decomposition, and also establish linearity.

Theorem 3.20 (Linearity and decomposition property) Let $p_i\in[0,1]$ for $i\in I$ with $\sum_{i\in I}p_i\leq 1$. Then

(i) If $\Delta_i\Longrightarrow\Theta_i$ for all $i\in I$ then $\sum_{i\in I}p_i\cdot\Delta_i\Longrightarrow\sum_{i\in I}p_i\cdot\Theta_i$.

(ii) If $\sum_{i\in I}p_i\cdot\Delta_i\Longrightarrow\Theta$ then $\Theta=\sum_{i\in I}p_i\cdot\Theta_i$ for subdistributions Θ_i such that $\Delta_i\Longrightarrow\Theta_i$ for all $i\in I$.

Proof: (i) Suppose $\Delta_i\Longrightarrow\Theta_i$ for all $i\in I$. By Definition 3.13 there are subdistributions $\Delta_{ik},\Delta_{ik}^\rightarrow,\Delta_{ik}^\times$ such that

$$\Delta_i=\Delta_{i0}, \qquad \Delta_{ik}=\Delta_{ik}^\rightarrow+\Delta_{ik}^\times, \qquad \Delta_{ik}^\rightarrow\xrightarrow{\tau}\Delta_{i(k+1)}, \qquad \Theta_i=\sum_k\Delta_{ik}^\times.$$

Therefore, we have that $\sum_{i\in I}p_i\cdot\Delta_i=\sum_{i\in I}p_i\cdot\Delta_{i0}$, $\sum_{i\in I}p_i\cdot\Delta_{ik}=\sum_{i\in I}p_i\cdot\Delta_{ik}^\rightarrow+\sum_{i\in I}p_i\cdot\Delta_{ik}^\times$, $\sum_{i\in I}p_i\cdot\Delta_{ik}^\rightarrow\xrightarrow{\tau}\sum_{i\in I}p_i\cdot\Delta_{i(k+1)}$ by Clause (2) of Definition 3.2, and $\sum_{i\in I}p_i\cdot\Theta_i=\sum_{i\in I}p_i\cdot\sum_k\Delta_{ik}^\times=\sum_k(\sum_{i\in I}p_i\cdot\Delta_{ik}^\times)$. By Definition 3.13 we obtain $\sum_{i\in I}p_i\cdot\Delta_i\Longrightarrow\sum_{i\in I}p_i\cdot\Theta_i$.

(ii) In the light of Lemma 3.19(ii) it suffices to show that

if $\sum_{i=0}^\infty\Delta_i\Longrightarrow\Theta$ then $\Theta=\sum_{i=0}^\infty\Theta_i$ for subdistributions Θ_i such that $\Delta_i\Longrightarrow\Theta_i$ for all $i\geq 0$.

Since $\sum_{i=0}^{\infty} \Delta_i = \Delta_0 + \sum_{i\geq 1} \Delta_i$ and $\sum_{i=0}^{\infty} \Delta_i \Longrightarrow \Theta$, by Lemma 3.19(iii) there are $\Theta_0, \Theta_1^{\geq}$ such that

$$\Delta_0 \Longrightarrow \Theta_0, \qquad \sum_{i\geq 1} \Delta_i \Longrightarrow \Theta_1^{\geq}, \qquad \Theta = \Theta_0 + \Theta_1^{\geq} \ .$$

Using Lemma 3.19(iii) once more, we have $\Theta_1, \Theta_2^{\geq}$ such that

$$\Delta_1 \Longrightarrow \Theta_1, \qquad \sum_{i\geq 2} \Delta_i \Longrightarrow \Theta_2^{\geq}, \qquad \Theta_1^{\geq} = \Theta_1 + \Theta_2^{\geq},$$

thus in combination $\Theta = \Theta_0 + \Theta_1 + \Theta_2^{\geq}$. Continuing this process we have that

$$\Delta_i \Longrightarrow \Theta_i, \qquad \sum_{j\geq i+1} \Delta_j \Longrightarrow \Theta_{i+1}^{\geq}, \qquad \Theta = \sum_{j=0}^{i} \Theta_j + \Theta_{i+1}^{\geq}$$

for all $i \geq 0$. Lemma 3.19(i) ensures that $|\sum_{j\geq i+1} \Delta_j| \geq |\Theta_{i+1}^{\geq}|$ for all $i \geq 0$. But since $\sum_{i=0}^{\infty} \Delta_i$ is a subdistribution, we know that the tail sum $\sum_{j\geq i+1} \Delta_j$ converges to ε when i approaches ∞, and therefore that $\lim_{i\to\infty} \Theta_i^{\geq} = \varepsilon$. Thus by taking that limit we conclude that $\Theta = \sum_{i=0}^{\infty} \Theta_i$. □

With Theorem 3.20, the relation $\Longrightarrow \subseteq \mathscr{D}(S) \times \mathscr{D}(S)$ can be obtained as the lifting of a relation $\Longrightarrow_S$ from S to $\mathscr{D}(S)$, which is defined by writing $s \Longrightarrow_S \Theta$ just when $\overline{s} \Longrightarrow \Theta$.

Proposition 3.21 $\overline{(\Longrightarrow_S)} = (\Longrightarrow)$.

Proof: That $\Delta \overline{\Longrightarrow_S} \Theta$ implies $\Delta \Longrightarrow \Theta$ is a simple application of Part (i) of Theorem 3.20. For the other direction, suppose $\Delta \Longrightarrow \Theta$: given that $\Delta = \sum_{s\in\lceil\Delta\rceil} \Delta(s)\cdot\overline{s}$, Part (ii) of the same theorem enables us to decompose Θ into $\sum_{s\in\lceil\Delta\rceil} \Delta(s)\cdot\Theta_s$ where $\overline{s} \Longrightarrow \Theta_s$ for each s in $\lceil\Delta\rceil$. But the latter actually means that $s \Longrightarrow_S \Theta_s$, and so by definition this implies $\Delta \overline{\Longrightarrow_S} \Theta$. □

It is immediate that the relation $\Longrightarrow$ is convex because of its being a lifting.

We proceed with the important properties of reflexivity and transitivity of weak derivations. First note that reflexivity is straightforward; in Definition 3.13 it suffices to take $\Delta_0^{\rightarrow}$ to be the empty distribution ε.

Theorem 3.22 (Transitivity of $\Longrightarrow$) If $\Delta \Longrightarrow \Theta$ and $\Theta \Longrightarrow \Lambda$ then $\Delta \Longrightarrow \Lambda$.

Proof: By definition $\Delta \Longrightarrow \Theta$ means that some $\Delta_k, \Delta_k^{\times}, \Delta_k^{\rightarrow}$ exist for all $k \geq 0$ such that

$$\Delta = \Delta_0, \qquad \Delta_k = \Delta_k^{\times} + \Delta_k^{\rightarrow}, \qquad \Delta_k^{\rightarrow} \xrightarrow{\tau} \Delta_{k+1}, \qquad \Theta = \sum_{k=0}^{\infty} \Delta_k^{\times}. \tag{4}$$

Since $\Theta = \sum_{k=0}^{\infty} \Delta_k^{\times}$ and $\Theta \Longrightarrow \Lambda$, by Theorem 3.20(ii) there are Λ_k for $k \geq 0$ such that $\Lambda = \sum_{k=0}^{\infty} \Lambda_k$ and $\Delta_k^{\times} \Longrightarrow \Lambda_k$ for all $k \geq 0$. For each $k \geq 0$, we know that $\Delta_k^{\times} \Longrightarrow \Lambda_k$ gives us some Δ_{kl}, $\Delta_{kl}^{\times}, \Delta_{kl}^{\rightarrow}$ for $l \geq 0$ such that

$$\Delta_k^{\times} = \Delta_{k0}, \qquad \Delta_{kl} = \Delta_{kl}^{\times} + \Delta_{kl}^{\rightarrow}, \qquad \Delta_{kl}^{\rightarrow} \xrightarrow{\tau} \Delta_{k,l+1} \qquad \Lambda_k = \sum_{l\geq 0} \Delta_{kl}^{\times}. \tag{5}$$

Therefore we can put all this together with

$$\Lambda \quad = \quad \sum_{k=0}^{\infty} \Lambda_k \quad = \quad \sum_{k,l\geq 0} \Delta_{kl}^{\times} \quad = \quad \sum_{i\geq 0} \left(\sum_{k,l|k+l=i} \Delta_{kl}^{\times} \right) \ , \tag{6}$$

where the last step is a straightforward diagonalisation.

Now from the decompositions above we re-compose an alternative trajectory of ${\Delta'}_i$'s to take Δ via $\Longrightarrow$ to Λ directly. Define

$$ {\Delta'}_i = {\Delta'}^{\times}_i + {\Delta'}^{\rightarrow}_i , \qquad {\Delta'}^{\times}_i = \sum_{k,l|k+l=i} \Delta^{\times}_{kl}, \qquad {\Delta'}^{\rightarrow}_i = (\sum_{k,l|k+l=i} \Delta^{\rightarrow}_{kl}) + \Delta^{\rightarrow}_i \ , \tag{7} $$

so that from (6) we have immediately that

$$ \Lambda \quad = \quad \sum_{i\geq 0} {\Delta'}^{\times}_i \ . \tag{8} $$

We now show that

(i) $\Delta = {\Delta'}_0$

(ii) ${\Delta'}^{\rightarrow}_i \xrightarrow{\tau} {\Delta'}_{i+1}$

from which, with (8), we will have $\Delta \Longrightarrow \Lambda$ as required. For (i) we observe that

$$
\begin{array}{lll}
 & \Delta & \\
= & \Delta_0 & (4) \\
= & \Delta^{\times}_0 + \Delta^{\rightarrow}_0 & (4) \\
= & \Delta_{00} + \Delta^{\rightarrow}_0 & (5) \\
= & \Delta^{\times}_{00} + \Delta^{\rightarrow}_{00} + \Delta^{\rightarrow}_0 & (5) \\
= & (\sum_{k,l|k+l=0} \Delta^{\times}_{kl}) + (\sum_{k,l|k+l=0} \Delta^{\rightarrow}_{kl}) + \Delta^{\rightarrow}_0 & \text{index arithmetic} \\
= & {\Delta'}^{\times}_0 + {\Delta'}^{\rightarrow}_0 & (7) \\
= & {\Delta'}_0 \ . & (7)
\end{array}
$$

For (ii) we observe that

$$
\begin{array}{lll}
 & {\Delta'}^{\rightarrow}_i & \\
= & (\sum_{k,l|k+l=i} \Delta^{\rightarrow}_{kl}) + \Delta^{\rightarrow}_i & (7) \\
\xrightarrow{\tau} & (\sum_{k,l|k+l=i} \Delta_{k,l+1}) + \Delta_{i+1} & (4), (5), \text{Remark 3.7(ii)} \\
= & (\sum_{k,l|k+l=i} (\Delta^{\times}_{k,l+1} + \Delta^{\rightarrow}_{k,l+1})) + \Delta^{\times}_{i+1} + \Delta^{\rightarrow}_{i+1} & (4), (5) \\
= & (\sum_{k,l|k+l=i} \Delta^{\times}_{k,l+1}) + \Delta^{\times}_{i+1} + (\sum_{k,l|k+l=i} \Delta^{\rightarrow}_{k,l+1}) + \Delta^{\rightarrow}_{i+1} & \text{rearrange} \\
= & (\sum_{k,l|k+l=i} \Delta^{\times}_{k,l+1}) + \Delta_{i+1,0} + (\sum_{k,l|k+l=i} \Delta^{\rightarrow}_{k,l+1}) + \Delta^{\rightarrow}_{i+1} & (5) \\
= & (\sum_{k,l|k+l=i} \Delta^{\times}_{k,l+1}) + \Delta^{\times}_{i+1,0} + \Delta^{\rightarrow}_{i+1,0} + (\sum_{k,l|k+l=i} \Delta^{\rightarrow}_{k,l+1}) + \Delta^{\rightarrow}_{i+1} & (5) \\
= & (\sum_{k,l|k+l=i+1} \Delta^{\times}_{kl}) + (\sum_{k,l|k+l=i+1} \Delta^{\rightarrow}_{kl}) + \Delta^{\rightarrow}_{i+1} & \text{index arithmetic} \\
= & {\Delta'}^{\times}_{i+1} + {\Delta'}^{\rightarrow}_{i+1} & (7) \\
= & {\Delta'}_{i+1} \ , & (7)
\end{array}
$$

which concludes the proof. □

Finally, we need a property that is the converse of transitivity: if one executes a given weak derivation partly, by stopping more often and moving on less often, one makes another weak transition that can be regarded as an initial segment of the given one. We need the property that after executing such an initial segment, it is still possible to complete the given derivation.

Definition 3.23 A weak derivation $\Phi \Longrightarrow \Gamma$ is called an *initial segment* of a weak derivation $\Phi \Longrightarrow \Psi$ if for $k \geq 0$ there are $\Gamma_k, \Gamma^{\rightarrow}_k, \Gamma^{\times}_k, \Psi_k, \Psi^{\rightarrow}_k, \Psi^{\times}_k \in \mathscr{D}(S)$ such that $\Gamma_0 = \Psi_0 = \Phi$ and

$$
\begin{array}{lll}
\Gamma_k = \Gamma^{\rightarrow}_k + \Gamma^{\times}_k & \Psi_k = \Psi^{\rightarrow}_k + \Psi^{\times}_k & \Gamma^{\rightarrow}_k \leq \Psi^{\rightarrow}_k \\
\Gamma^{\rightarrow}_k \xrightarrow{\tau} \Gamma_{k+1} & \Psi^{\rightarrow}_k \xrightarrow{\tau} \Psi_{k+1} & \Gamma_k \leq \Psi_k \\
\Gamma = \sum_{i=0}^{\infty} \Gamma^{\times}_k & \Psi = \sum_{i=0}^{\infty} \Psi^{\times}_k & (\Psi^{\rightarrow}_k - \Gamma^{\rightarrow}_k) \xrightarrow{\tau} (\Psi_{k+1} - \Gamma_{k+1}).
\end{array}
$$

Proposition 3.24 If $\Phi \Longrightarrow \Gamma$ is an initial segment of $\Phi \Longrightarrow \Psi$, then $\Gamma \Longrightarrow \Psi$. $\square$

Proof: For subdistribution $\Delta, \Theta \in \mathscr{D}(S)$ define $\Delta \cap \Theta \in \mathscr{D}(S)$ by $\Delta \cap \Theta(s) := \min(\Delta(s), \Theta(s))$ and $\Delta - \Theta \in \mathscr{D}(S)$ by $\Delta - \Theta(s) := \min(\Delta(s) - \Theta(s), 0)$. So $\Delta - \Theta = \Delta - (\Delta \cap \Theta)$. Observe that in case $\Theta \leq \Delta$, and only then, we have that $(\Delta - \Theta) + \Theta = \Delta$.

Let $\Gamma_k, \Gamma_k^{\rightarrow}, \Gamma_k^{\times}, \Psi_k, \Psi_k^{\rightarrow}, \Psi_k^{\times} \in \mathscr{D}(S)$ be as in Definition 3.23. By induction on $k \geq 0$ we define Δ_{ki}, $\Delta_{ki}^{\rightarrow}$ and $\Delta_{ki}^{\times}$, for $0 \leq i \leq k$, such that

$$\Delta_{k0} = \Gamma_k^{\times} \qquad \Psi_k = \textstyle\sum_{i=0}^{k} \Delta_{ki} + \Gamma_k^{\rightarrow} \qquad \Psi_k^{\times} = \textstyle\sum_{i=0}^{k} \Delta_{ki}^{\times} \qquad \Delta_{ki} = \Delta_{ki}^{\rightarrow} + \Delta_{ki}^{\times} \qquad \Delta_{ki}^{\rightarrow} \xrightarrow{\tau} \Delta_{(k+1)(i+1)} \ .$$

Induction base: Let $\Delta_{00} := \Gamma_0^{\times} = \Gamma_0 - \Gamma_0^{\rightarrow} = \Psi_0 - \Gamma_0^{\rightarrow}$. This way the first two equations are satisfied for $k = 0$. All other statements will be dealt with fully by the induction step.

Induction step: Suppose Δ_{ki} for $0 \leq i \leq k$ are already known, and $\Psi_k = \sum_{i=0}^{k} \Delta_{ki} + \Gamma_k^{\rightarrow}$. With induction on i we define $\Delta_{ki}^{\times} := \Delta_{ki} \cap (\Psi_k^{\times} - \sum_{j=0}^{i-1} \Delta_{kj}^{\times})$ and establish that $\sum_{j=0}^{i} \Delta_{kj}^{\times} \leq \Psi_k^{\times}$. Namely, writing Θ_{ki} for $\sum_{j=0}^{i-1} \Delta_{kj}^{\times}$, surely $\Theta_{k0} = \varepsilon \leq \Psi_k^{\times}$, and when assuming that $\Theta_{ki} \leq \Psi_k^{\times}$, for some $0 \leq i \leq k$, and defining $\Delta_{ki}^{\times} := \Delta_{ki} \cap (\Psi_k^{\times} - \Theta_{ki})$ we obtain $\Theta_{k(i+1)} = \Delta_{ki}^{\times} + \Theta_{ki} \leq (\Psi_k^{\times} - \Theta_{ki}) + \Theta_{ki} = \Psi_k^{\times}$. So in particular $\sum_{i=0}^{k} \Delta_{ki}^{\times} \leq \Psi_k^{\times}$. Using that $\Gamma_k^{\rightarrow} \leq \Psi_k^{\rightarrow}$ we find

$$\Delta_{kk} = (\Psi_k - \Gamma_k^{\rightarrow}) - \sum_{i=0}^{k-1} \Delta_{ki} = (\Psi_k^{\times} + (\Psi_k^{\rightarrow} - \Gamma_k^{\rightarrow})) - \sum_{i=0}^{k-1} \Delta_{ki} \geq \Psi_k^{\times} - \sum_{i=0}^{k-1} \Delta_{ki},$$

hence $\Delta_{kk}^{\times} = \Delta_{kk} \cap (\Psi_k^{\times} - \sum_{i=0}^{k-1} \Delta_{ki}^{\times}) = \Psi_k^{\times} - \sum_{i=0}^{k-1} \Delta_{ki}$ and thus $\Psi_k^{\times} = \sum_{i=0}^{k} \Delta_{ki}$.

Now define $\Delta_{ki}^{\rightarrow} := \Delta_{ki} - \Delta_{ki}^{\times}$. This yields $\Delta_{ki} = \Delta_{ki}^{\rightarrow} + \Delta_{ki}^{\times}$ and thereby

$$\Psi_k^{\rightarrow} = \Psi_k - \Psi_k^{\times} = (\sum_{i=0}^{k} \Delta_{ki} + \Gamma_k^{\rightarrow}) - \sum_{i=0}^{k} \Delta_{ki}^{\times} = \sum_{i=0}^{k} \Delta_{ki}^{\rightarrow} + \Gamma_k^{\rightarrow}.$$

Since $\sum_{i=0}^{k} \Delta_{ki}^{\rightarrow} = (\Psi_k^{\rightarrow} - \Gamma_k^{\rightarrow}) \xrightarrow{\tau} (\Psi_{k+1} - \Gamma_{k+1})$, by Proposition 3.10 we have $\Psi_{k+1} - \Gamma_{k+1} = \sum_{i=0}^{k} \Delta_{(k+1)(i+1)}$ for some subdistributions $\Delta_{(k+1)(i+1)}$ such that $\Delta_{ki}^{\rightarrow} \xrightarrow{\tau} \Delta_{(k+1)(i+1)}$ for $i = 0, \ldots, k$. Furthermore, define $\Delta_{(k+1)0} := \Gamma_{k+1}^{\times} = \Gamma_{k+1} - \Gamma_{k+1}^{\rightarrow}$. It follows that

$$\Psi_{k+1} = \sum_{i=0}^{k} \Delta_{(k+1)(i+1)} + \Gamma_{k+1} = \sum_{i=1}^{k+1} \Delta_{(k+1)i} + (\Delta_{(k+1)0} + \Gamma_{k+1}^{\rightarrow}) = \sum_{i=0}^{k+1} \Delta_{(k+1)i} + \Gamma_{k+1}^{\rightarrow}.$$

This ends the inductive definition and proof. Now let $\Theta_i := \sum_{k=i}^{\infty} \Delta_{ki}$, $\Theta_i^{\rightarrow} := \sum_{k=i}^{\infty} \Delta_{ki}^{\rightarrow}$ and $\Theta_i^{\times} := \sum_{k=i}^{\infty} \Delta_{ki}^{\times}$. It follows that $\Theta_0 = \sum_{k=0}^{\infty} \Delta_{k0} = \sum_{k=0}^{\infty} \Gamma_k^{\times} = \Gamma$, $\Theta_i = \Theta_i^{\rightarrow} + \Theta_i^{+}$, and, using Remark 3.7(ii), $\Theta_i^{\rightarrow} \xrightarrow{\tau} \Theta_{i+1}$. Moreover, $\sum_{i=0}^{\infty} \Theta_i^{\times} = \sum_{i=0}^{\infty} \sum_{k=i}^{\infty} \Delta_{ki}^{\times} = \sum_{k=0}^{\infty} \sum_{i=0}^{k} \Delta_{ki}^{\times} = \sum_{k=0}^{\infty} \Psi_k^{\times} = \Psi$. Definition 3.13 yields $\Gamma \Longrightarrow \Psi$. $\square$

3.4 Derivations through policies

In Markov Decision Theory [22] *policies* are used to determine a run of a process. These are is essentially the same as the *schedulers* of [17]. Here we will show that this method agrees with our (weak) derivations.

In Markov Decision Processes (MDPs) [22] transitions are usually unlabelled. To lift the definition of a policy from MDPs to pLTSs, we need to map pLTSs to MDPs. Here we do this by considering an MDP to be a pLTS in which all transitions are labelled τ, and mapping a pLTS to an MDP by leaving out all non-τ transitions. This method yields the required match between policies and weak derivations. An alternative map from pLTSs to MDPs would be to simply forget the transition labels. In that case we would need to use a notion of derivation obtained from the one in Definition 3.13 by dropping the requirement that the transitions $\Delta_k^{\rightarrow} \longrightarrow \Delta_{k+1}$ are labelled τ.

A policy specifies for each state s is a pLTS $\langle S, \mathsf{Act}_\tau, \rightarrow \rangle$ a “way to proceed”. This “way” is a probabilistic combination of the outgoing τ-transitions of s, with as a special component in this probabilistic combination the possibility not to proceed further at all. A *history-dependent* policy makes the way to proceed from s depending on the way one arrives at s, called a *history* of s. Here, for the sake of generality, we postulate a set H of histories, equipped with a function *last* $: H \rightarrow S$ telling from a given history $h \in H$ of which state *last*(h) a history it is.

Definition 3.25 A *policy* for a pLTS $\langle S, \mathsf{Act}_\tau, \rightarrow\rangle$ is a function $\mathscr{P} : H \rightarrow \mathscr{D}(S)$ such that if $\mathscr{P}(h) = p{\cdot}\Delta$ with $p \in (0,1]$ and $\Delta \in \mathscr{D}_1(S)$ then there is a transition $\overline{last(h)} \xrightarrow{\tau} \Delta$.

A policy $\mathscr{P}$, with $\mathscr{P}(h) = p.\Delta$, says that when we are in a state $s = last(h)$ and our history is h, with probability p we proceed to the distribution Δ, and with probability $1{-}p$ we remain permanently in s. A policy is *static* if instead of taking probabilistic combinations, it specifies just one (or none) of the outgoing transitions of s; that is, in Definition 3.25 we require that either $\mathscr{P}(h) = \varepsilon$ or $\mathscr{P}(h) = \Delta \in \mathscr{D}_1(S)$ with $last(h) \xrightarrow{\tau} \Delta$.

We consider several types of policies, depended on the choice of H. A *history-independent* policy [22] is one that does not depend on histories of states; take $H = S$ and *last* the identity function. For a *history-dependent* policy [22] take $H = S^*$. In MDPs, there is an *initial* distribution Δ, and a *history* h *of a state* s is defined as a sequence of states $s_0, s_1, \ldots, s_k$ such that $s_0 \in \lceil\Delta\rceil$ and for $i = 1, \ldots, k$ there are $\Theta_i \in \mathscr{D}_1(S)$ with $s_{i-1} \xrightarrow{\tau} \Theta_i$ and $s_i \in \lceil\Theta_i\rceil$. Here $last(h) := s_k = s$. We define $\mathscr{P}^*(h)$ as the probability that in the run specified by $\mathscr{P}$ we initially visit the sequence of states $h \in H$:

$$\mathscr{P}^*(s) := \Delta(s) \qquad\qquad \mathscr{P}^*(hs) := \mathscr{P}^*(h){\cdot}\mathscr{P}(h)(s)\ .$$

Furthermore, let the *length* $|h| \in \mathbb{N}$ of h be given by $|s_0, s_1, \ldots, s_k| = k$. We now formalise the *run* induced by a history-dependent policy $\mathscr{P}$ from an initial distribution Δ as the weak derivation $\Delta \Longrightarrow \Delta'$, where for $k \in \mathbb{N}$ the Δ_k, $\Delta_k^\times$ and $\Delta_k^\rightarrow$ of Definition 3.13 are given by

$$\Delta_k(s) := \sum_{\{h \mid last(h)=s \wedge |h|=k\}} \mathscr{P}^*(h) \qquad \Delta_k^\rightarrow(s) := \sum_{\{h \mid last(h)=s \wedge |h|=k\}} \mathscr{P}^*(h){\cdot}|\mathscr{P}(h)| \qquad \Delta_k^\times(s) := \sum_{\{h \mid last(h)=s \wedge |h|=k\}} \mathscr{P}^*(h){\cdot}(1{-}|\mathscr{P}(h)|)\ .$$

Note that $\Delta_k^\rightarrow := \sum_{s\in S} \sum_{\{h \mid last(h)=s \wedge |h|=k\}} \mathscr{P}^*(h)){\cdot}|\mathscr{P}(h)|{\cdot}\overline{s} = \sum_{\{h \mid |h|=k\}} \mathscr{P}^*(h){\cdot}|\mathscr{P}(h)|{\cdot}\overline{last(h)}$. Furthermore, by Definition 3.25, for all $h \in H$ we have $\overline{last(h)} \xrightarrow{\tau} p^{-1}{\cdot}\mathscr{P}(h)$ with $p = |\mathscr{P}(h)|$, and thus

$$|\mathscr{P}(h)|{\cdot}\overline{last(h)} \xrightarrow{\tau} \mathscr{P}(h) = \sum_{s\in S} \mathscr{P}(h)(s){\cdot}\overline{s}\ .$$

Hence, Lemma 3.5 yields

$$\Delta_k^\rightarrow \xrightarrow{\tau} \sum_{\{h \mid |h|=k\}} \mathscr{P}^*(h){\cdot}\sum_{s\in S} \mathscr{P}(h)(s){\cdot}\overline{s} = \sum_{\{hs \mid |h|=k\}} \mathscr{P}^*(hs){\cdot}\overline{s} = \textstyle\sum_{s\in S} \displaystyle\sum_{\{h \mid last(h)=s \wedge |h|=k+1\}} \mathscr{P}^*(h){\cdot}\overline{s} = \sum_{s\in S} \Delta_{k+1}(s){\cdot}\overline{s} = \Delta_{k+1}\ .$$

Since also $\Delta_0 = \Delta$ and $\Delta_k = \Delta_k^\times + \Delta_k^\rightarrow$, this yields a weak derivation indeed. We denote it by $\Delta \stackrel{\mathscr{P}}{\Longrightarrow} \Delta'$.

So each history-dependent policy induces a weak derivation. We will complete the promised correspondence between policies and derivations by showing that, conversely, each weak derivation can be induced by a history-dependent policy. In fact we obtain a stronger result: each weak derivation is already induced by a special kind of history-dependent policy, which we call a *time-dependent policy*. If follows that the extra generality of history-dependent over time-dependent policies is not needed for the purpose of specifying runs of pLTSs.

A time-dependent policy is obtained by taking $H = S \times \mathbb{N}$ with $last(s,k) = s$. Here (s,k) merely says that one has reached state s after exactly k transitions.[3] Obviously, each time-dependent policy $\mathscr{P}$ can be seen as special kind of history-dependent policy $\mathscr{P}^{\mathrm{hd}}$, defined by $\mathscr{P}^{\mathrm{hd}}(h) := \mathscr{P}(last(h), |h|)$.

Definition 3.26 Given a time-dependent policy $\mathscr{P}$, we formalise the *run* it induces from an initial distribution Δ as the weak derivation $\Delta \Longrightarrow \Delta'$, where for $k \in \mathbb{N}$ the Δ_k, $\Delta_k^\times$ and $\Delta_k^\rightarrow$ of Definition 3.13 are given by

$$\Delta_0 := \Delta \qquad \Delta_k^\times(s) := \Delta_k(s){\cdot}(1{-}|\mathscr{P}(s,k)|) \qquad \Delta_k^\rightarrow(s) := \Delta_k(s){\cdot}|\mathscr{P}(s,k)| \qquad \Delta_{k+1} := \sum_{s\in\lceil\Delta\rceil_k} \Delta_k(s){\cdot}\mathscr{P}(s,k).$$

Since for all $k \in \mathbb{N}$ we have $\Delta_k = \Delta_k^\times + \Delta_k^\rightarrow$ and $\Delta_k^\rightarrow \xrightarrow{\tau} \Delta_{k+1}$, this specifies weak derivation indeed. Again, we denote it by $\Delta \stackrel{\mathscr{P}}{\Longrightarrow} \Delta'$. We now show that this induced weak derivation agrees with the one defined for history-dependent policies.

Proposition 3.27 If $\mathscr{P}$ is a time-dependent policy, then $\Delta \stackrel{\mathscr{P}}{\Longrightarrow} \Delta'$ iff $\Delta \stackrel{\mathscr{P}^{\mathrm{hd}}}{\Longrightarrow} \Delta'$.

[3] As our formalism doesn't model time explicitly, the number of transitions performed so far could serve as a crude approximation of time.

Proof: It suffices to derive the four defining equations of Definition 3.26 from the definition of $\Delta \stackrel{\mathscr{P}^{\rm hd}}{\Longrightarrow} \Delta'$. For the first three this is immediate, using that $\mathscr{P}^{\rm hd}(h) = \mathscr{P}(\mathit{last}(h), |h|)$. Furthermore, for all $k \in \mathbb{N}$ and $t \in S$,

$$\begin{aligned}
\textstyle\sum_{s\in\lceil\Delta\rceil_k} \Delta_k(s)\cdot\mathscr{P}(s,k)(t) &= \textstyle\sum_{s\in S}\left(\sum_{\{h|\mathit{last}(h)=s\wedge|h|=k\}} \mathscr{P}^{\rm hd*}(h)\right)\cdot\mathscr{P}(s,k)(t) \\
&= \textstyle\sum_{s\in S}\sum_{\{h|\mathit{last}(h)=s\wedge|h|=k\}} \left(\mathscr{P}^{\rm hd*}(h)\cdot\mathscr{P}(\mathit{last}(h),|h|)(t)\right) \\
&= \textstyle\sum_{\{h||h|=k\}} (\mathscr{P}^{\rm hd*}(h)\cdot\mathscr{P}^{\rm hd}(h)(t)) \\
&= \textstyle\sum_{\{h||h|=k\}} \mathscr{P}^{\rm hd*}(ht) \\
&= \textstyle\sum_{\{h|\mathit{last}(h)=t\wedge|h|=k+1\}} \mathscr{P}^{\rm hd*}(h) \\
&= \Delta_{k+1}(t)\ .
\end{aligned}$$
□

Theorem 3.28 For every weak derivation $\Delta \Longrightarrow \Delta'$ there exists a time-dependent policy $\mathscr{P}$ such that $\Delta \stackrel{\mathscr{P}}{\Longrightarrow} \Delta'$.

Proof: Let $\Delta \Longrightarrow \Delta'$. By Definition 3.13 there are Δ_k, $\Delta_k^{\times}$ and $\Delta_k^{\rightarrow}$ for all $k \geq 0$ such that

$$\Delta = \Delta_0, \qquad \Delta_k = \Delta_k^{\times} + \Delta_k^{\rightarrow}, \qquad \Delta_k^{\rightarrow} \stackrel{\tau}{\longrightarrow} \Delta_{k+1}, \qquad \Theta = \sum_{k=0}^{\infty} \Delta_k^{\times}.$$

By Proposition 3.10, $\Delta_k^{\rightarrow} \stackrel{\tau}{\longrightarrow} \Delta_{k=1}$ implies that there are distributions $\Delta_{k+1}^s \in \mathscr{D}_1(S)$ for $s \in \lceil\Delta_k^{\rightarrow}\rceil$, such that

$$\overline{s} \stackrel{\tau}{\longrightarrow} \Delta_{k+1}^s \ \text{ for each } s \in \lceil\Delta_k^{\rightarrow}\rceil \qquad \text{and} \qquad \Delta_{k+1} = \textstyle\sum_{s\in\lceil\Delta_k^{\rightarrow}\rceil} \Delta_k^{\rightarrow}(s)\cdot\Delta_{k+1}^s\ .$$

Now take $\mathscr{P}(s,k) := \frac{\Delta_k^{\rightarrow}(s)}{\Delta_k(s)}\cdot\Delta_{k+1}^s$. Then all four equations of Definition 3.26 are satisfied, so $\Delta \stackrel{\mathscr{P}}{\Longrightarrow} \Delta'$. □

4 Testing probabilistic processes

This section is divided into three. Applying a test to a process results in a nondeterministic, but possibly probabilistic, computation structure. The main conceptual issue is how to associate outcomes with these nondeterministic structures. In the first subsection we outline a general approach in which intuitively the nondeterministic choices are resolved implicitly in a dynamic manner. In the second section we describe an alternative approach in which we explicitly associate with a nondeterministic structure a set of deterministic *computations*, each of which determines a possible outcome. In the final section we show that although these approaches are formally quite different they lead to exactly the same testing preorders.

4.1 Applying a test to a process

We now retrace our earlier approach [4, 2] to the testing of probabilistic processes. A *test* is simply a process in the language pCSP, except that it may in addition use special *success* actions for reporting outcomes: these are drawn from a set Ω of fresh actions not already in Act_τ. We refer to the augmented language as pCSP^{Ω}. Formally a test T is some process from that language, and to apply test T to process P we form the process $T\ |_{\mathsf{Act}}\ P$ in which *all* visible actions of P must synchronise with T. The resulting composition is a process whose only possible actions are τ and the elements of Ω. We will define the result $\mathscr{A}(T,P)$ of applying the test T to the process P to be a set of testing outcomes, exactly one of which results from each resolution of the choices in $T\ |_{\mathsf{Act}}\ P$. Each *testing outcome* is an Ω-tuple of real numbers in the interval [0,1], i.e. a function $o : \Omega \to [0,1]$, and its ω-component $o(\omega)$, for $\omega \in \Omega$, gives the probability that the resolution in question will reach an ω-*success state*, one in which the success action ω is possible.

There are several ways to fill in details in this approach. Following [5], we first of all distinguish between *vector-based* testing, in which one allows countably many success actions, and *scalar* testing, in which there is only one success action and consequently outcomes are scalars rather than vectors. Scalar testing is employed in [6, 8, 25, 4], and vector-based testing in [24]. As in [2], our prime interest here is in scalar testing, but we use vector-based testing as an indispensable tool for establishing our results. To this end we employ a result from [5] saying that for finitary probabilistic processes, scalar and vector-based testing give rise to the very same testing preorders.

Secondly, following [5, 2] we distinguish between *state-based* and *action-based* testing. The former is what we described above: success actions are merely used as a method to define success states; a method that bypasses the need to formally introduce state predicates in the operational semantics of our language. In action-based testing, on the other hand, it is the actual execution of a success action that constitutes success, and $o(\omega)$ gives the probability that the resolution in question will perform the action ω. State-based testing is employed in [6, 8, 25, 4], and action-based testing in [24, 5]. In [2] it has been shown that for finite probabilistic processes (obtained by dropping the recursion construct from pCSP) the state-based and action-based testing preorders coincide. This allowed us to use action-based testing to obtain results about state-based testing. However, [2] also provides an example showing that for the language considered in the present paper, the two approaches are different, in particular that state-based must testing is more discriminating than action-based must testing. The same example also applies to the non-probabilistic world and, for finitely branching processes, it is the state-based must-testing preorder that coincides with the CSP refinement preorder based on failures and divergences [1, 10, 21]. It is in part for this reason that we employ state-based testing in the current paper.

Whereas state-based scalar testing as well as action-based scalar- and vector-based testing have been used previously in the literature, our use of state-based vector-based testing is new. Since we use this concept merely as a method for proving results about state-based scalar testing, we are not concerned about the generality of our tests for conceptual reasons; any notion of state-based vector-based testing that works in our proofs would be acceptable, as long as the special case of state-based scalar testing agrees with the definitions found in the literature. Here we restrict attention to tests in which no state is simultaneously an ω-success state for different values of ω. In fact, we can go further by ruling out all tests in which from one success state one can reach another one, with a different success value.

Definition 4.1 An Ω-*test* is a closed pCSP expression T, but allowing the enriched alphabet $\mathsf{Act}_\tau \cup \Omega$ of actions instead of just Act_τ, such that if $t \xrightarrow{\omega_1}$ and $u \xrightarrow{\omega_2}$ for $\omega_1, \omega_2 \in \Omega$ with t reachable from T and u from t, then $\omega_1 = \omega_2$.

Note that for the special case of state-based scalar testing the above restriction is void. In [5], working in an action-based framework, following [24], we did not put such a restriction in our definition of testing, but showed, in Appendix A: "One Success Never Leads to Another" that imposing it does not change the resulting testing preorders. Also note that the composition $T \mid_{\mathsf{Act}} P$ of an Ω-test T and a pCSP process P is again an Ω-test (i.e. satisfying the requirement of Definition 4.1).

Intuitively, the application of a test T to a process P has an outcome $o \in [0,1]^\Omega$ if we can imagine an army — with a continuum of soldiers — marching through our pLTS, starting from the distribution $[\![T \mid_{\mathsf{Act}} P]\!]$, of which, for $\omega \in \Omega$, a fraction $o(\omega) \in [0,1]$ eventually reaches an ω-success state. Each time a fragment of the army ends up in a non-success state, it splits up in arbitrary proportions among the outgoing transitions of that state (which are all labelled τ). If such a transition ends up in a distribution Δ then, for $s \in \lceil\Delta\rceil$, a fraction $\Delta(s)$ of the fragment that took that transition ends up in state s. The army begins its march by being distributed over the initial distribution $[\![T \mid_{\mathsf{Act}} P]\!]$ in the same vein. As soon as a fragment of the army reaches an ω-success state, it stops marching, and the size of that fragment is counted towards $o(\omega)$. Definition 4.1 ensures that in such a case there is no ambiguity about which success action the division contributes to. Definition 4.1 also ensures that there is no point in marching any further. The total success value $\sum_{\omega\in\Omega} o(\omega)$ must be in the interval [0,1]; it represents the fraction of the army eventually reaching a success state of any kind. The unsuccessful part of the army $1 - \sum_{\omega\in\Omega} o(\omega)$ represents the fraction that either got stuck in a *deadlock* state, one without outgoing transitions, or that will march forever. In general we get different outcomes $o \in \mathscr{A}(T,P)$ for each possible way a fragment in a non-success and non-deadlock state can partition itself among the outgoing transitions of that state.

We will now formalise this intuition by a definition of $\mathscr{A}(T,P)$. Our definition has three ingredients. First of all we normalise our pLTS by removing all τ-transitions that leave a success state. This way an ω-success state will only have outgoing transitions labelled ω. This prevents our army from scooting past a success state.

Definition 4.2 (ω-respecting) Let $\langle S, L, \rightarrow\rangle$ be a pLTS such that the set of labels L includes Ω. It is said to be ω-*respecting* whenever $s \xrightarrow{\omega}$, for any $\omega \in \Omega$, implies $s \not\xrightarrow{\tau}$.

It is straightforward to modify an arbitrary pLTS so that it becomes ω-respecting. Here we outline how this is done for our pLTS for pCSP.

Definition 4.3 (Pruning) Let $[\cdot]$ be the unary operator on Ω-test states given by the operational rules

$$\frac{s \xrightarrow{\omega} \Delta}{[s] \xrightarrow{\omega} [\Delta]} \;(\omega \in \Omega) \qquad\qquad \frac{s \stackrel{\omega}{\not\rightarrow} \;(\text{for all } \omega \in \Omega),\;\; s \xrightarrow{\alpha} \Delta}{[s] \xrightarrow{\alpha} [\Delta]} \;(\alpha \in \mathsf{Act}_\tau)\;.$$

Just as $\Box$ and $|_A$, this operator extends as syntactic sugar to Ω-tests by distributing $[\cdot]$ over ${}_p\oplus$; likewise, it extends to distributions by $[\Delta]([s]) = \Delta(s)$. Clearly, this operator does nothing else than removing all outgoing transitions of a success state other than the ones labelled with $\omega \in \Omega$. Applying this operator, we can just as well envision our army to have started marching from the distribution $[\![[T \mid_{\mathsf{Act}} P]]\!]$; it will continue marching along τ-transitions for as long as τ-transitions are possible, and will halt in state s iff $s \stackrel{\tau}{\not\rightarrow}$, which is the case iff s is either a success or a deadlock state.

Next, using Definition 3.13, we characterise the set of subdistributions Θ that can be reached by an army as envisioned above at the end of its march from $[\![[T \mid_{\mathsf{Act}} P]]\!]$. In general Θ need not be a total distribution: the mass $|\Theta|$ represents the fraction of the army that eventually stops marching and thus reaches Θ. The remaining fraction $1-|\Theta|$ of the army marches on forever. A march of an army as described above can be modelled perfectly by a weak transition $[\![[T \mid_{\mathsf{Act}} P]]\!] \Longrightarrow \Theta$ as defined in Section 3.2. The end subdistribution Θ of this march has the property that there is no nontrivial weak transition $\Theta \Longrightarrow$. System states with this property are traditionally called *stable*.

Definition 4.4 (Extreme derivatives) A state s in a pLTS is called *stable* if $s \stackrel{\tau}{\not\rightarrow}$, and a subdistribution Θ is called *stable* if every state in its support is stable. We write $\Delta \Longrightarrow\!\!\succ \Theta$ whenever $\Delta \Longrightarrow \Theta$ and Θ is stable, and call Θ an *extreme* derivative of Δ.

Referring to Definition 3.13, we see this means that in the extreme derivation of Θ from Δ at every stage a state must move on if it can, so that every stopping component can contain only states which *must* stop: for $s \in \lceil \Delta_k^{\rightarrow} + \Delta_k^{\times} \rceil$ we have $s \in \lceil \Delta_k^{\times} \rceil$ if *and now also* only if $s \stackrel{\tau}{\not\rightarrow}$. Moreover if the pLTS is ω-respecting then whenever $s \in \lceil \Delta_k^{\rightarrow} \rceil$, that is whenever it marches on, it is not successful, i.e. $s \stackrel{\omega}{\not\rightarrow}$ for every $\omega \in \Omega$.

Lemma 4.5 (Existence and uniqueness of extreme derivatives)

(i) For every subdistribution Δ there exists some (stable) Δ' such that $\Delta \Longrightarrow\!\!\succ \Delta'$.

(ii) In a deterministic pLTS we have that $\Delta \Longrightarrow\!\!\succ \Delta'$ and $\Delta \Longrightarrow\!\!\succ \Delta''$ implies $\Delta' = \Delta''$.

Proof: We construct a derivation as in Definition 3.13 of a stable Δ' by defining the components $\Delta_k, \Delta_k^{\times}$ and $\Delta_k^{\rightarrow}$ using induction on k. Let us assume that the subdistribution Δ_k has been defined; in the base case $k = 0$ this is simply Δ. The decomposition of this Δ_k into the components $\Delta_k^{\times}$ and $\Delta_k^{\rightarrow}$ is carried out by defining the former to contain precisely those states which must stop, i.e. those s for which $s \stackrel{\tau}{\not\rightarrow}$. Formally $\Delta_k^{\times}$ is determined by:

$$\Delta_k^{\times}(s) = \begin{cases} \Delta_k(s) & \text{if } s \stackrel{\tau}{\not\rightarrow} \\ 0 & \text{otherwise.} \end{cases}$$

Then $\Delta_k^{\rightarrow}$ is given by the *remainder* of Δ_k, namely those states which can perform a τ action:

$$\Delta_k^{\rightarrow}(s) = \begin{cases} \Delta_k(s) & \text{if } s \xrightarrow{\tau} \\ 0 & \text{otherwise.} \end{cases}$$

Note that these definitions divide the support of Δ_k into two disjoints sets, namely the support of $\Delta_k^{\times}$ and the support of $\Delta_k^{\rightarrow}$. Moreover by construction we know that $\Delta_k^{\rightarrow} \xrightarrow{\tau} \Theta$ for some Θ; we let Δ_{k+1} be an arbitrary such Θ.

This completes our definition of an extreme derivative as in Definition 3.13 and so we have established (i).

For (ii) we observe that in a deterministic pLTS the above choice of Δ_{k+1} is unique, so that the whole derivative construction becomes unique. $\Box$

It is worth pointing out that the use of subdistributions, rather than distributions, is essential here. If Δ diverges, that is if there is an infinite sequence of derivations $\Delta \xrightarrow{\tau} \Delta_1 \xrightarrow{\tau} \ldots \Delta_k \xrightarrow{\tau} \ldots$, then one extreme derivative of Δ is the empty subdistribution ε. For example the only transition of $\mathsf{rec}\, x.\, x$ is $\mathsf{rec}\, x.\, x \xrightarrow{\tau} \overline{\mathsf{rec}\, x.\, x}$, and thus $\overline{\mathsf{rec}\, x.\, x}$ diverges; ε is its unique extreme derivative.

The final ingredient in the definition of the set of outcomes $\mathscr{A}(T, P)$ is the outcome of a particular extreme derivative Θ. All states $s \in \lceil \Theta \rceil$ either satisfy $s \xrightarrow{\omega}$ for a unique $\omega \in \Omega$, or have $s \not\rightarrow$.

Definition 4.6 (Outcomes) The outcome $\$\Theta \in [0,1]^\Omega$ of a stable subdistribution Θ is given by $\$\Theta(\omega) = \sum_{s \in \lceil \Theta \rceil,\ s \xrightarrow{\omega}} \Theta(s)$.

Putting all three ingredients together, we arrive at a definition of $\mathscr{A}(T, P)$:

Definition 4.7 Let P be a pCSP process and T an Ω-test. Then $\mathscr{A}(T, P) = \{\$\Theta \mid [\![[T \mid_{\mathsf{Act}} P]]\!] \Longrightarrow\!\!\!\succ \Theta\}$.

The role of pruning in the above definition can be seen via the following example.

Example 4.8 Let $P = a.b$ and $T = a.(b \;\Box\; \omega)$. The pLTS generated by applying T to P can be described by the process $\tau.(\tau \;\Box\; \omega)$. Now $[\![T \mid_{\mathsf{Act}} P]\!]$ has a unique extreme derivative $[\![\, \mathbf{0} \,]\!]$, whereas $[\![[T \mid_{\mathsf{Act}} P]]\!]$ has a unique extreme derivative $[\![\omega]\!]$. The outcome in $\mathscr{A}(T, P)$ shows that process P passes test T with probability 1, which is what we expect for state-based testing, which we use in this paper. Without pruning we would get an outcome saying that P passes T with probability 0, which would be what is expected for action-based testing. $\Box$

As this example is nonprobabilistic, it also illustrates how pruning enables the standard notion of nonprobabilistic testing to be captured in this way.

We compare two vectors of probabilities component-wise, and two sets of vectors of probabilities via the Hoare- and Smyth preorders:

$$
\begin{array}{ll}
O_1 \leq_{\mathrm{Ho}} O_2 & \text{if for every } o_1 \in O_1 \text{ there exists some } o_2 \in O_2 \text{ such that } o_1 \leq o_2 \\
O_1 \leq_{\mathrm{Sm}} O_2 & \text{if for every } o_2 \in O_2 \text{ there exists some } o_1 \in O_1 \text{ such that } o_1 \leq o_2 \,.
\end{array}
$$

This gives us our definition of the may- and must-testing preorders; they are decorated with $\cdot^\Omega$ for the repertoire Ω of testing actions they employ.

Definition 4.9 (Probabilistic testing preorders) Given two pCSP processes P and Q,

1. $P \sqsubseteq^\Omega_{\mathrm{pmay}} Q$ if for every Ω-test T, $\mathscr{A}(T, P) \leq_{\mathrm{Ho}} \mathscr{A}(T, Q)$;
2. $P \sqsubseteq^\Omega_{\mathrm{pmust}} Q$ if for every Ω-test T, $\mathscr{A}(T, P) \leq_{\mathrm{Sm}} \mathscr{A}(T, Q)$.

These preorders are abbreviated to $P \sqsubseteq_{\mathrm{pmay}} Q$, and $P \sqsubseteq_{\mathrm{pmust}} Q$, when $|\Omega| = 1$, and there kernels are denoted by $\simeq_{\mathrm{pmay}}$ and $\simeq_{\mathrm{pmust}}$ respectively.

Here are two examples of these preorders.

Example 4.10 Consider the process $Q_1 = \mathsf{rec}\, x.\, (\tau.x \;{}_{\frac{1}{2}}\!\oplus\; a)$, which was already discussed in the introduction, Figure 1(a). When we apply the test $T = a.\omega$ to it we get the pLTS-fragment in Figure 3(c), which is deterministic and unaffected by pruning; from part (ii) of Lemma 4.5 it follows that $T \mid_{\mathsf{Act}} Q_1$ has a unique extreme derivative Θ. Moreover Θ can be calculated to be

$$\sum_{k \geq 1} \frac{1}{2^k} \cdot \overline{s_3},$$

which simplifies to the distribution $\overline{s_3}$. Therefore, $\mathscr{A}(T, Q_1) = \{\$s_3\} = \{\vec{\omega}\}$, where $\vec{\omega} : \Omega \to [0,1]$ is the Ω-tuple with $\vec{\omega}(\omega) = 1$ and $\vec{\omega}(\omega') = 0$ for all $\omega' \neq \omega$. This is the same set of results gained by applying T to a on its own; and in fact it is possible to show that this holds for all tests, giving

$$Q_1 \simeq_{\mathrm{pmay}} a \qquad\qquad Q_1 \simeq_{\mathrm{pmust}} a \,.$$

$\Box$

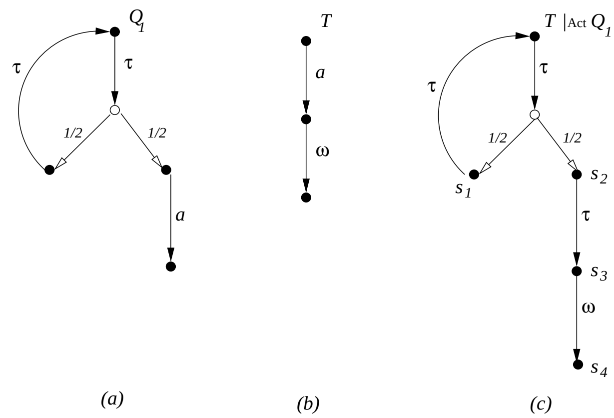


Figure 3: Testing the process Q_1

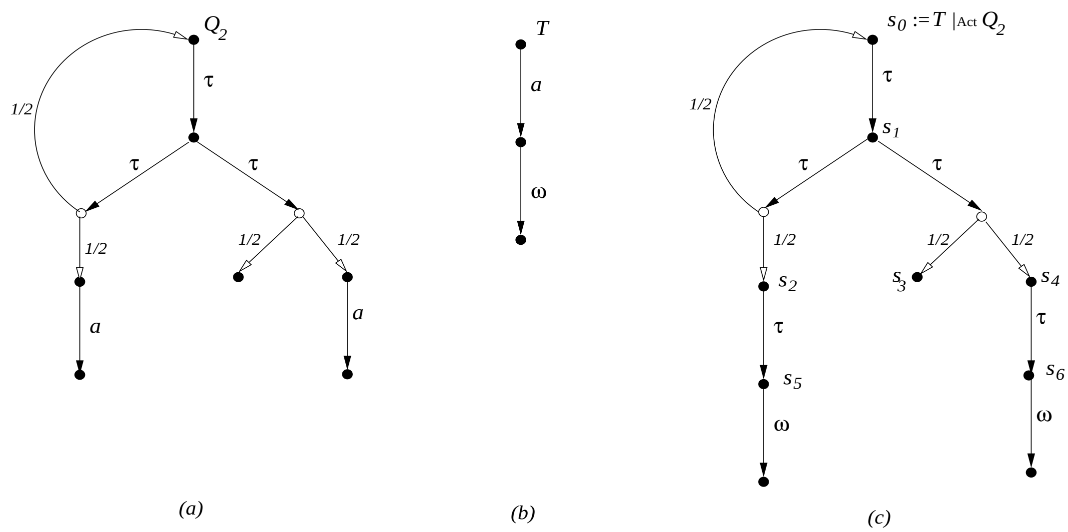


Figure 4: Testing the process Q_2

Example 4.11 Consider the process $Q_2 = \mathsf{rec}\, x.\, (\tau.(x \;{}_{\frac{1}{2}}\!\oplus\; a) \;\Box\; \tau.(\mathbf{0} \;{}_{\frac{1}{2}}\!\oplus\; a))$ and the application of the same test $T = a.\omega$ to it, as outlined in Figure 4. Since there is only one success action ω, the testing outcomes can be regarded as scalars in $[0,1]$ — that is, we write p for $p \cdot \vec{\omega}$, with $p \in [0,1]$.

Consider any extreme derivative Δ' from $\overline{s_0} = [\![T \mid_{\mathsf{Act}} Q_2]\!]$; note that here again pruning has no effect. Using the notation of Definition 3.13, it is clear that $\Delta_0^\times$ and $\Delta_0^\rightarrow$ must be ε and $\overline{s_0}$ respectively. Similarly, $\Delta_1^\times$ and $\Delta_1^\rightarrow$ must be ε and $\overline{s_1}$ respectively. But s_1 is a nondeterministic state, having two possible transitions:

(i) $s_1 \xrightarrow{\tau} \Theta_0$ where Θ_0 has support $\{s_0, s_2\}$ and assigns each of them the weight $\frac{1}{2}$

(ii) $s_1 \xrightarrow{\tau} \Theta_1$ where Θ_1 has the support $\{s_3, s_4\}$, again diving the mass equally among them.

So there are many possibilities for Δ_2; Lemma 3.5 shows that in fact Δ_2 can be of the form

$$p \cdot \Theta_0 + (1{-}p) \cdot \Theta_1 \tag{9}$$

for any choice of $p \in [0,1]$.

Let us consider one possibility, an extreme one where p is chosen to be 0; only the transition (ii) above is used. Here $\Delta_2^\rightarrow$ is the subdistribution $\frac{1}{2}\overline{s_4}$, and $\Delta_k^\rightarrow = \varepsilon$ whenever $k > 2$. A simple calculation shows that in this case the extreme derivative generated is $\Theta_1^e = \frac{1}{2}\overline{s_3} + \frac{1}{2}\overline{s_6}$ which implies that $\frac{1}{2} \in \mathscr{A}(T, Q_2)$.

Another possibility for Δ_2 is Θ_0, corresponding to the choice of $p{=}1$ in (9) above. Continuing with this derivation leads to Δ_3 being $\frac{1}{2} \cdot \overline{s_1} + \frac{1}{2} \cdot \overline{s_5}$; thus $\Delta_3^\times = \frac{1}{2} \cdot \overline{s_5}$ and $\Delta_3^\rightarrow = \frac{1}{2} \cdot \overline{s_1}$. Now in the generation of Δ_4 from $\Delta_3^\rightarrow$ once more we have to resolve a transition from the nondeterministic state s_1, by choosing some arbitrary $p \in [0,1]$ in (9). Suppose we chose $p = 1$ every time, completely ignoring transition (ii) above. Then the extreme derivative generated is

$$\Theta_0^e = \sum_{k \geq 1} \frac{1}{2^k} \cdot \overline{s_5}$$

which simplifies to the distribution $\overline{s_5}$. This in turn means that $1 \in \mathscr{A}(T, Q_2)$.

We have seen two possible derivations of extreme derivatives from $\overline{s_0}$. But there are many others. In general whenever $\Delta_k^\rightarrow$ is of the form $q \cdot \overline{s_1}$ we have to resolve the nondeterminism by choosing a $p \in [0,1]$ in (9) above; moreover each such choice is independent. However, it will follow from later results, specifically Corollary 6.10, that every extreme derivative Δ' of $\overline{s_0}$ is of the form

$$q \cdot \Theta_0^e + (1{-}q) \cdot \Theta_1^e$$

for some choice of $q \in [0,1]$; this is explained in Example 6.11. Consequently it follows that $\mathscr{A}(T, Q_2) = [\frac{1}{2}, 1]$.

Since $\mathscr{A}(T, a) = \{1\}$ it follows that

$$\mathscr{A}(T, a) \;\leq_{\mathrm{Ho}}\; \mathscr{A}(T, Q_2) \qquad\qquad \mathscr{A}(T, Q_2) \;\leq_{\mathrm{Sm}}\; \mathscr{A}(T, a)\;.$$

Again it is possible to show that these inequalities result from any test T and that therefore we have

$$a \sqsubseteq_{\text{pmay}} Q_2 \qquad\qquad Q_2 \sqsubseteq_{\text{pmust}} a\;. \qquad \Box$$

4.2 Using explicit resolutions

The derivation of extreme derivatives, via the schema in Definition 3.13, involves the systematic dynamic resolution of nondeterministic states, in each transition from $\Delta_k^\rightarrow$ to Δ_{k+1}. In the literature various mechanisms have been proposed for making these choices; for example *policies* are used in [22], adversaries in [15], schedulers in [23], ... Here we concentrate not on any such mechanism but rather the results of their application. In general they reduce a nondeterministic structure, typically a pLTS, to a set of deterministic structures. To describe these deterministic structures we adapt the notion of *resolution*, defined in [Seg96,..][5] for probabilistic automata, to pLTSs.

Definition 4.12 (Resolutions) A *resolution* of a subdistribution $\Delta \in \mathscr{D}(S)$ in a pLTS $\langle S, \Omega_\tau, \rightarrow\rangle$ is a triple $\langle R, \Theta, \rightarrow_R\rangle$ where $\langle R, \Omega_\tau, \rightarrow_R\rangle$ is a deterministic pLTS and $\Theta \in \mathscr{D}(R)$, such that there exists a *resolving function* $f \in R \rightarrow S$ satisfying

(i) $\mathrm{Img}_f(\Theta) = \Delta$

(ii) if $r \xrightarrow{\alpha}_R \Theta'$ for $\alpha \in \Omega_\tau$ then $f(r) \xrightarrow{\alpha} \mathrm{Img}_f(\Theta')$

(iii) if $f(r) \xrightarrow{\alpha}$ for $\alpha \in \Omega_\tau$ then $r \xrightarrow{\alpha}_R$.

The reader is referred to Section 2 of [5] for a detailed discussion of this concept of resolution, and the manner in which a resolution represents a run or *computation* of a process; in particular, in a resolution states in S are allowed to be resolved into distributions, and computation steps can be *probabilistically interpolated*.

Add paragraph on fully probabilistic

Schedulers

We now explain how to associate an outcome with a particular resolution, which in turn will associate a set of outcomes with a subdistribution in a pLTS. Given a deterministic pLTS $\langle R, \Omega_\tau, \rightarrow\rangle$, consider the functional $\mathscr{R} : (R \rightarrow [0,1]^\Omega) \rightarrow (R \rightarrow [0,1]^\Omega)$ defined by

$$\mathscr{R}(f)(r)(\omega) := \begin{cases} 1 & \text{if } r \xrightarrow{\omega} \\ 0 & \text{if } r \not\xrightarrow{\omega} \text{ and } r \not\xrightarrow{\tau} \\ \mathrm{Exp}_\Delta(f)(\omega) & \text{if } r \not\xrightarrow{\omega} \text{ and } r \xrightarrow{\tau} \Delta. \end{cases} \tag{10}$$

We view the unit interval $[0,1]$ ordered in the standard manner as a complete lattice; this induces the structure of a complete lattice on the product $[0,1]^\Omega$ and in turn on the set of functions $R \rightarrow [0,1]^\Omega$. The functional $\mathscr{R}$ is easily seen to be monotonic and therefore has a least fixed point, which we denote by $\mathbb{V}_{\langle R,\Omega_\tau,\rightarrow\rangle}$; this is abbreviated to $\mathbb{V}$ when the resolution in question is understood. Henceforth we write $\mathbb{V}(\Delta)$ for $\mathrm{Exp}_\Delta(\mathbb{V})$. Note that $\mathbb{V}(\sum_{i\in I}\Delta_i) = \sum_{i\in I}\mathbb{V}(\Delta_i)$.

Now let $\mathscr{A}^{\mathrm{r}}(T,P)$ denote the set of vectors

$$\{\, \mathbb{V}_{\langle R,\Omega_\tau,\rightarrow\rangle}(\Theta) \mid \langle R, \Theta, \rightarrow\rangle \text{ is a resolution of } [\![T \mid_{\mathsf{Act}} P]\!] \,\} \,.$$

Note that here we use resolutions of $[\![T \mid_{\mathsf{Act}} P]\!]$ rather than its pruning $[\![[T \mid_{\mathsf{Act}} P]]\!]$. This is because the functional $\mathscr{R}$, and therefore its least fixed point $\mathbb{V}$, has pruning built-in; that is $\mathscr{R}$ is defined so that $\mathbb{V}(s) = \mathbb{V}([s])$.

In Section 4.3 we will show that $\mathscr{A}^{\mathrm{r}}(T,P) = \mathscr{A}(T,P)$ for any test T and process P. Hence the testing preorders of Definition 4.9 can equivalently be defined in terms of $\mathscr{A}^{\mathrm{r}}$.

Example 4.13 (revisiting Example 4.10) The pLTS-fragment in Figure 3(c) is already deterministic, hence has essentially only one resolution, itself. Moreover the outcome $\mathrm{Exp}_{[\![T\|Q_1]\!]}(\mathbb{V}) = \mathbb{V}(T\|Q_1)$ associated with it is the least solution of the equation

$$\mathbb{V}(T\|Q_1) = \frac{1}{2}\cdot\mathbb{V}(T\|Q_1) + \frac{1}{2}\vec{\omega}$$

In fact this equation has a unique solution in $[0,1]^\Omega$, namely $\vec{\omega}$. Thus $\mathscr{A}^{\mathrm{r}}(T,Q_1) = \{\vec{\omega}\}$. □

Example 4.14 (revisiting Example 4.11) Here we reuse the notation of Example 4.11.

Consider the process $Q_2 = \mathsf{rec}\, x.\,(\tau.(x \;{}_{\frac{1}{2}}\!\oplus\; a) \;\Box\; \tau.(\mathbf{0} \;{}_{\frac{1}{2}}\!\oplus\; a))$ and the application of the test $T = a.\omega$ to it, as outlined in Figure 4. For each $k \geq 1$ the distribution $[\![T \mid_{\mathsf{Act}} Q_2]\!]$ has a resolution $\langle R_k, \Theta, \rightarrow_{R_k}\rangle$ such that $\mathbb{V}(\Theta) = (1-\frac{1}{2^k})$; intuitively it goes around the loop $(k-1)$ times before at last taking the right hand τ action. Thus $\mathscr{A}^{\mathrm{r}}(T,Q_2)$ contains $(1-\frac{1}{2^k})$ for every $k \geq 1$. But it also contains 1, because of the resolution which takes the left hand τ-move every time. Thus $\mathscr{A}^{\mathrm{r}}(T,Q_2)$ includes the set

$$\{(1-\frac{1}{2}),\ (1-\frac{1}{2^2}),\ldots,(1-\frac{1}{2^k}),\ldots\ 1\}$$

From later results it will follow that $\mathscr{A}^{\mathrm{r}}(T,Q_2)$ is actually the convex closure of this set, namely $[\frac{1}{2},1]$. □

4.3 Comparison

We have now seen two ways of associating sets of outcomes with the application of a test to a process. The first, in Section 4.1, uses extreme derivations in which nondeterministic choices are resolved dynamically as the derivation proceeds, while the second, in Section 4.2, associates with a test and a process a set of deterministic structures called resolutions. In this section we show that both approaches yield the same sets of outcomes.

We start by showing that resolution-based testing is insensitive to pruning. Let $\mathscr{A}^{\mathrm{rp}}(T,P)$ denote the set of vectors

$$\{\, \mathbb{V}_{\langle R,\Omega_\tau,\rightarrow\rangle}(\Theta) \mid \langle R,\Theta,\rightarrow\rangle \text{ is a resolution of } [\![[T \mid_{\mathsf{Act}} P]]\!] \,\} \;.$$

Proposition 4.15 For any test T and process P we have that $\mathscr{A}^{\mathrm{rp}}(T,P) = \mathscr{A}^{\mathrm{r}}(T,P)$.

Proof: "$\supseteq$": Let $\langle R,\Theta,\rightarrow_R\rangle$ be a resolution of $[\![T \mid_{\mathsf{Act}} P]\!]$. Then, following Definition 4.12, $\langle R,[\Theta],\rightarrow_R\rangle$ is a resolution of $[\![[T \mid_{\mathsf{Act}} P]]\!]$ and, by (10), $\mathbb{V}_{\langle R,\Omega_\tau,\rightarrow_R\rangle}([\Theta]) = \mathbb{V}_{\langle R,\Omega_\tau,\rightarrow_R\rangle}(\Theta)$.

"$\subseteq$": Let $\langle R,\Theta,\rightarrow_R\rangle$ be a resolution of $[\![[T \mid_{\mathsf{Act}} P]]\!]$ with resolving function f. We construct a resolution $\langle R',\Theta,\rightarrow'_R\rangle$ of $[\![T \mid_{\mathsf{Act}} P]\!]$ as a random extension of $\langle R,\Theta,\rightarrow_R\rangle$. Let $\langle S,\Omega_\tau,\rightarrow\rangle$ be the PLTS in which the distribution $[\![T \mid_{\mathsf{Act}} P]\!]$ exists. For every pair $(s,\alpha)\in S\times\Omega_\tau$ with $s \xrightarrow{\alpha}$ pick a distribution $\Psi^{(s,\alpha)} \in \mathscr{D}_1(S)$ such that $s \xrightarrow{\alpha} \Psi$. Now define $R' := R \mathbin{\dot\cup} (S\times\mathbb{N})$ and obtain $\rightarrow'_R$ from $\rightarrow_R$ by adding (A) a transition $(s,k) \xrightarrow{\alpha}{}'_R \Psi^{(s,\alpha)}_{k+1}$ for each $k\in\mathbb{N}$ and each $s\in S$ with $s\xrightarrow{\alpha}$, and (B) a transition $r \xrightarrow{\tau}{}'_R \Psi_0^{(f(r),\tau)}$ for each $r\in R$ with $f(r)\xrightarrow{\tau}$ as well as $f(r)\xrightarrow{\omega}$ for some $\omega\in\Omega$. Here $\Psi^{(s,\alpha)}_{k+1} \in \mathscr{D}_1(S\times\{k{+}1\})$ is given by $\Psi^{(s,\alpha)}_{k+1}(t,k{+}1) = \Psi^{(s,\alpha)}(t)$ for all $t\in S$. The resolving function f is extended by $f(s,k) := s$. Using Definition 4.12 it follows that $\langle R',\Theta,\rightarrow'_R\rangle$ is a resolution of $[\![T \mid_{\mathsf{Act}} P]\!]$ and, again by (10), $\mathbb{V}_{\langle R',\Omega_\tau,\rightarrow'_R\rangle}(\Theta) = \mathbb{V}_{\langle R,\Omega_\tau,\rightarrow_R\rangle}(\Theta)$. □

It remains to show that $\mathscr{A}(T,P) = \mathscr{A}^{\mathrm{rp}}(T,P)$ for any test T and process P, or, in other words, that

$$\{\$\Theta \mid \Delta \Longrightarrow\!\!\succ \Theta\} = \{\, \mathbb{V}_{\langle R,\Omega_\tau,\rightarrow\rangle}(\Theta) \mid \langle R,\Theta,\rightarrow\rangle \text{ is a resolution of } \Delta \,\}$$

for any distribution Δ is an ω-respecting pLTS $\langle S,\Omega_\tau,\rightarrow\rangle$.

First let us see how an extreme derivation can be viewed as a method for dynamically generating a resolution.

Proposition 4.16 (Resolutions from extreme derivatives) Let $\Delta \Longrightarrow\!\!\succ \Delta'$ in a pLTS $\langle S,\Omega_\tau,\rightarrow\rangle$. Then there is a resolution $\langle R,\Theta,\rightarrow_R\rangle$ of Δ, with resolving function f, such that $\Theta \Longrightarrow\!\!\succ_R \Theta'$ for some Θ' for which $\Delta' = \mathrm{Img}_f(\Theta')$.

Proof: Consider an extreme derivation of $\Delta \Longrightarrow\!\!\succ \Delta'$ as given in Definition 3.13 where all $\Delta^{\times}_k$ must be stable:

$$\Delta = \Delta_0, \qquad \Delta_k = \Delta^{\times}_k + \Delta^{\rightarrow}_k, \qquad \Delta^{\rightarrow}_k \xrightarrow{\tau} \Delta_{k+1}, \qquad \Delta' = \textstyle\sum_{k=0}^{\infty} \Delta^{\times}_k .$$

By Lemma 3.5, $\Delta^{\rightarrow}_k \xrightarrow{\tau} \Delta_{k=1}$ implies that there are states $s_{ik}\in S$ and distributions $\Delta_{i(k+1)} \in \mathscr{D}_1(S)$, such that

$$\Delta^{\rightarrow}_k = \textstyle\sum_{i\in I_k} p_{ik}\cdot\overline{s_{ik}}, \qquad s_{ik} \xrightarrow{\tau} \Delta_{i(k+1)} \text{ for each } i\in I_k \qquad \text{and} \qquad \Delta_{k+1} = \textstyle\sum_{i\in I_k} p_{ik}\cdot\Delta_{i(k+1)} \;.$$

Let $\Delta^{\times}_{ik}(s) := \begin{cases}\Delta_{ik}(s) & \text{if } s \stackrel{\tau}{\not\rightarrow} \\ 0 & \text{if } s \xrightarrow{\tau}\end{cases}$. Since $\Delta^{\times}_k(s) = \begin{cases}\Delta_k(s) & \text{if } s \stackrel{\tau}{\not\rightarrow} \\ 0 & \text{if } s \xrightarrow{\tau}\end{cases}$ it follows that $\Delta^{\times}_{k+1} = \sum_{i\in I_k} p_{ik}\cdot\Delta^{\times}_{i(k+1)}$.

We will now define the resolution $\langle R,\Theta,\rightarrow_R\rangle$ and the resolving function f. The set of states R is $(S\times\mathbb{N}) \cup \bigcup_{k\in\mathbb{N}}(I_k\times\{k\})$. The resolving function $f : R\rightarrow S$ maps $(s,k)\in S\times\mathbb{N}$ to s and $(i,k)\in I_k\times\{k\}$ to $s_{ik}\in S$. The second component k of a state counts how many transitions have fired already: each transition in $\rightarrow_R$ goes from a state (i,k) or (s,k) to a distribution over $(S\cup I_{k+1})\times\{k{+}1\}$.

Define the subdistributions $\Theta^{\times}_k \in \mathscr{D}(S\times\{k\})$ and $\Theta^{\rightarrow}_k \in \mathscr{D}(I_k\times\{k\})$ by $\Theta^{\times}_k(s,k) = \Delta^{\times}_k(s)$ and $\Theta^{\rightarrow}_k(i,k) = p_{ik}$. Let $\Theta_k := \Theta^{\times}_k + \Theta^{\rightarrow}_k$ and $\Theta := \Theta_0$. Furthermore, for all $k>0$ and $i\in I_{k-1}$, define $\Theta_{ik} \in \mathscr{D}((S\cup I_k)\times\{k\})$ by

$$\Theta_{ik}(s,k) = \Delta^{\times}_{ik}(s) \qquad\qquad \text{and} \qquad\qquad \Theta_{ik}(j,k) = p_{jk}\cdot\frac{\Delta_{ik}(s_{jk})}{\Delta_k(s_{jk})}$$

for $j\in I_k$. We introduce the transitions $(i,k) \xrightarrow{\tau}_R \Theta_{i(k+1)}$ for $k\geq 0$ and $i\in I_k$. Moreover, for each state $s\in S$ and label $\alpha\in\mathsf{Act}_\tau$ such that $s\xrightarrow{\alpha}$, pick a transition $s\xrightarrow{\alpha}\Psi$, and add the transition $(s,k)\xrightarrow{\alpha}_R \Psi_{k+1}$ to $\rightarrow_R$, for all $k\in\mathbb{N}$. Here Ψ_{k+1} is the distribution with $\Psi_{k+1}(t,k{+}1) = \Psi(t)$ for all $t\in S$. Likewise, for each $k\in\mathbb{N}$, $i\in I_k$ and

$\omega \in \Omega$ such that $s_{ik} \xrightarrow{\omega}$, pick a transition $s_{ik} \xrightarrow{\omega} \Psi$, and add the transition $(i,k) \xrightarrow{\omega}_R \Psi_{k+1}$ to $\rightarrow_R$. This ends the definition of the resolution $\langle R, \Theta, \rightarrow_R \rangle$ and the resolving function f. By construction, $\langle R, \Omega_\tau, \rightarrow_R \rangle$ is a deterministic pLTS. We now check that f satisfies the requirements for a resolving function of Definition 4.12.

(i) $$\text{Img}_f(\Theta_k)(s) = \Theta_k(s,k) + \sum_{s_{ik}=s} \Theta_k(i,k) = \Theta_k^\times(s,k) + \sum_{s_{ik}=s} p_{ik} = \Delta_k^\times(s) + \Delta_k^\rightarrow(s) = \Delta_k(s)$$
for all $s \in S$, so $\text{Img}_f(\Theta_k) = \Delta_k$, and in particular $\text{Img}_f(\Theta) = \Delta$.

(ii) Let $r \xrightarrow{\alpha}_R \Gamma$ for $\alpha \in \Omega_\tau$. In case $r = (s,k)$ it must be that $\Gamma = \Psi_{k+1}$ and $f(r) = s \xrightarrow{\alpha} \Phi = \text{Img}_f(\Psi_{k+1})$. Likewise, in case $r = (i,k)$ and $\alpha \in \Omega$ it must be that $\Gamma = \Psi_{k+1}$ and $f(r) = s_{ik} \xrightarrow{\alpha} \Phi = \text{Img}_f(\Psi_{k+1})$. The remaining case is $r = (i,k)$, $\alpha = \tau$ and $\Gamma = \Theta_{i(k+1)}$. Then $f(r) = s_{ik} \xrightarrow{\tau} \Delta_{i(k+1)}$, so it suffices to show that $\text{Img}_f(\Theta_{ik}) = \Delta_{ik}$ for all $k \in \mathbb{N}$ and $i \in I_k$. For any $s \in S$ we have
$$\text{Img}_f(\Theta_{ik})(s) = \Theta_{ik}(s,k) + \sum_{s_{jk}=s} \Theta_{ik}(j,k) = \Delta_{ik}^\times(s) + \sum_{s_{jk}=s} p_{jk} \cdot \frac{\Delta_{ik}(s_{jk})}{\Delta_k(s_{jk})} = \Delta_{ik}^\times(s) + \frac{\Delta_{ik}(s)}{\Delta_k(s)} \cdot \sum_{s_{jk}=s} p_{jk} \;.$$
In case $s \not\xrightarrow{\tau}$ we have $s_{jk} = s$ for no $j \in I_k$, so $\text{Img}_f(\Theta_{ik})(s) = \Delta_{ik}^\times(s) = \Delta_{ik}(s)$.
In case $s \xrightarrow{\tau}$ we have $\Delta_{ik}^\times(s) = 0$ and $\sum_{s_{jk}=s} p_{jk} = \Delta_k^\rightarrow(s) = \Delta_k(s)$, so again $\text{Img}_f(\Theta_{ik})(s) = \Delta_{ik}(s)$.

(iii) Let $f(r) \xrightarrow{\alpha}$ for $\alpha \in \Omega_\tau$. By construction there is a Φ_{k+1} such that $r \xrightarrow{\alpha}_R \Phi_{k+1}$.

Hence $\langle R, \Theta, \rightarrow_R \rangle$ is a resolution of Δ. We have:
$$\sum_{i \in I_k} p_{ik} \cdot \Theta_{i(k+1)}(s,k{+}1) = \sum_{i \in I_k} p_{ik} \cdot \Delta_{i(k+1)}^\times(s) = \Delta_{k+1}^\times(s) = \Theta_{k+1}^\times(s,k{+}1) = \Theta_{k+1}(s,k{+}1)$$
$$\sum_{i \in I_k} p_{ik} \cdot \Theta_{i(k+1)}(j,k{+}1) = \sum_{i \in I_k} p_{ik} \cdot p_{j(k+1)} \cdot \frac{\Delta_{i(k+1)}(s_{j(k+1)})}{\Delta_{k+1}(s_{j(k+1)})} = p_{j(k+1)} = \Theta_{k+1}^\rightarrow(j,k{+}1) = \Theta_{k+1}(j,k{+}1).$$
Hence $\Theta_{k+1} = \sum_{i \in I_k} p_{ik} \cdot \Theta_{i(k+1)}$. Since also $\Theta_k^\rightarrow = \sum_{i \in I_k} p_{ik} \cdot \overline{(i,k)}$ and $(i,k) \xrightarrow{\tau}_R \Theta_{i(k+1)}$, Lemma 3.5 yields $\Theta_k^\rightarrow \xrightarrow{\tau}_R \Theta_{k+1}$. Let $\Theta' = \sum_{k=0}^\infty \Theta_k^\times$. Then, by Definition 3.13, $\Theta \Longrightarrow\!\!\succ_R \Theta'$.

By construction $\text{Img}_f(\Theta_k^\times) = \Delta_k^\times$ for all $k \in \mathbb{N}$. Hence $\text{Img}_f(\Theta') = \sum_{k=0}^\infty \text{Img}_f(\Theta_k^\times) = \sum_{k=0}^\infty \Delta_k^\times = \Delta'$. □

The converse is somewhat simpler.

Proposition 4.17 (Extreme derivatives from resolutions) Let $\langle R, \Theta, \rightarrow_R \rangle$ be a resolution of a subdistribution Δ in a pLTS $\langle S, \Omega_\tau, \rightarrow \rangle$ with resolving function f. Then $\Theta \Longrightarrow\!\!\succ_R \Theta'$ implies $\Delta \Longrightarrow\!\!\succ \text{Img}_f(\Theta')$.

Proof: The definition of Img_f implies that $\text{Img}_f(\sum_i p_i \cdot \Psi_i) = \sum_i p_i \cdot \text{Img}_f(\Psi_i)$. Furthermore $\Psi \xrightarrow{\tau} \Psi'$ implies $\text{Img}_f(\Psi) \xrightarrow{\tau} \text{Img}_f(\Psi')$. Namely, by Lemma 3.5, $\Psi \xrightarrow{\tau} \Psi'$ implies
$$\Psi = \textstyle\sum_{i \in I} p_i \cdot \overline{s_i}, \qquad s_i \xrightarrow{\tau} \Psi_i \text{ for each } i \in I \qquad \text{and} \qquad \Psi' = \textstyle\sum_{i \in I} p_i \cdot \Psi_i$$
which, using Definition 4.12, entails
$$\text{Img}_f(\Psi) = \textstyle\sum_{i \in I} p_i \cdot \overline{f(s_i)}, \quad f(s_i) \xrightarrow{\tau} \text{Img}_f(\Psi_i) \text{ for each } i \in I \quad \text{and} \quad \text{Img}_f(\Psi') = \textstyle\sum_{i \in I} p_i \cdot \text{Img}_f(\Psi_i).$$
Hence $\text{Img}_f(\Psi) \xrightarrow{\tau} \text{Img}_f(\Psi')$.

Now consider any derivation of $\Theta \Longrightarrow\!\!\succ_R \Theta'$ along the lines of Definition 3.13. By systematically applying the function f to the component subdistributions in this derivation we get a derivation $\text{Img}_f(\Theta) \Longrightarrow \text{Img}_f(\Theta')$, that is $\Delta \Longrightarrow \text{Img}_f(\Theta')$. To show that $\text{Img}_f(\Theta')$ is actually an extreme derivative it suffices to show that s is stable for every $s \in \lceil \text{Img}_f(\Theta') \rceil$. But if $s \in \lceil \text{Img}_f(\Theta') \rceil$ then by definition there is some $t \in \lceil \Theta' \rceil$ such that $s = f(t)$. Since $\Theta \Longrightarrow\!\!\succ_R \Theta'$ the state t must be stable. The stability of s now follows from requirement (iii) of Definition 4.12. □

Our next step is to relate the outcomes extracted from extreme derivatives to those extracted from the corresponding resolutions. This requires some analysis of the evaluation function $\mathbb{V}(-)$.

Definition 4.18 (Continuous functions) A chain in a complete lattice L is a sequence of elements $\{ c_n \mid n \geq 0 \}$ satisfying $c_i \leq c_{i+1}$. Obviously chains have least upper bounds which we denote by $\bigsqcup_{n\geq 0} c_n$. A function $f : L \rightarrow L$ is said to be *continuous* if it preserves the least upper bounds of chains:

$$f(\bigsqcup_{n\geq 0} c_n) = \bigsqcup_{n\geq 0} f(c_n) \ .$$

Lemma 4.19 The functional $\mathscr{R} : (R \rightarrow [0,1]^\Omega) \rightarrow (R \rightarrow [0,1]^\Omega)$ defined in (10) is continuous.

Proof: The proof is surprisingly difficult; see Lemma B.5 in Appendix B which shows the result in the special case that Ω is the singleton set $\{\omega\}$; the general case is similar. □

Continuity of $\mathscr{R}$ implies that its fixed point $\mathbb{V}$ can be captured by a chain of approximants. The functions $\mathbb{V}^n,\ n \geq 0$ are defined by induction on n:

$$\begin{aligned} \mathbb{V}^0(r)(\omega) &= 0 \\ \mathbb{V}^{n+1} &= \mathscr{R}(\mathbb{V}^n) \ . \end{aligned}$$

Again we write $\mathbb{V}^n(\Delta)$ for $\mathrm{Exp}_\Delta(\mathbb{V}^n)$. Now $\mathbb{V} = \bigsqcup_{n\geq 0} \ \mathbb{V}^n$. This is used in the following result.

Lemma 4.20 Let Δ be a subdistribution in an ω-respecting deterministic pLTS. If $\Delta \Longrightarrow\!\!\!\succ \Delta'$ then $\mathbb{V}(\Delta) = \mathbb{V}(\Delta')$.

Proof: Since the pLTS is ω-respecting we know that $s \xrightarrow{\tau} \Delta$ implies $s \not\xrightarrow{\omega}$ for any ω. Therefore, from the definition of the functional $\mathscr{R}$ we have that $s \xrightarrow{\tau} \Delta$ implies $\mathbb{V}^{n+1}(s) = \mathbb{V}^n(\Delta)$, whence by lifting and linearity we get

$$\text{If } \Theta \xrightarrow{\tau} \Theta' \text{ then } \mathbb{V}^{n+1}(\Theta) = \mathbb{V}^n(\Theta') \text{ for all } n \geq 0.$$

Now suppose $\Delta \Longrightarrow\!\!\!\succ \Delta'$. Then

$$\Delta = \Delta_0, \qquad \Delta_k = \Delta_k^\times + \Delta_k^\rightarrow, \qquad \Delta_k^\rightarrow \xrightarrow{\tau} \Delta_{k+1}, \qquad \Delta' = \sum_{k=0}^{\infty} \Delta_k^\times .$$

Using in the base case that $\mathbb{V}^0(\Theta)(\omega) = 0$ for every Θ, a straightforward induction on n yields

$$\mathbb{V}^n(\Delta) \quad = \quad \sum_{k=0}^{n} \mathbb{V}^{n-k}(\Delta_k^\times) \ . \tag{11}$$

Since $\Delta_k^\times$ is stable, we have $\mathbb{V}^m(\Delta_k^\times) = \mathbb{V}(\Delta_k^\times)$ for every $k, m \geq 0$. We conclude by reasoning

$$\begin{aligned}
\mathbb{V}(\Delta) &= \textstyle\bigsqcup_{n\geq 0} \mathbb{V}^n(\Delta) && \text{by continuity of } \mathscr{R} \\
&= \textstyle\bigsqcup_{n\geq 0} \sum_{k=0}^{n} \mathbb{V}^{n-k}(\Delta_k^\times) && \text{from (11) above} \\
&= \textstyle\bigsqcup_{n\geq 0} \sum_{k=0}^{n} \mathbb{V}^{n}(\Delta_k^\times) && \text{since } \mathbb{V}^{n-k}(\Delta_k^\times) = \mathbb{V}(\Delta_k^\times) = \mathbb{V}^n(\Delta_k^\times) \\
&= \textstyle\bigsqcup_{n\geq 0} \mathbb{V}^{n}(\sum_{k=0}^{n} \Delta_k^\times) && \text{by linearity of } \mathbb{V}^n \\
&= \mathbb{V}(\textstyle\bigsqcup_{n\geq 0} \sum_{k=0}^{n} \Delta_k^\times) && \text{by continuity of } \mathscr{R} \\
&= \mathbb{V}(\textstyle\sum_{k=0}^{\infty} \Delta_k^\times) \\
&= \mathbb{V}(\Delta') \ .
\end{aligned}$$

□

We are now ready to compare the two methods for calculating the set of outcomes associated with a subdistribution:

- using extreme derivatives and the reward function \$ from Definition 4.6
- using resolutions and the evaluation function $\mathbb{V}$ from page 23.

Theorem 4.21 In an ω-respecting pLTS $\langle S, \Omega_\tau, \rightarrow \rangle$, the following statements hold.

(a) If $\Delta \Longrightarrow\!\!\!\succ \Delta'$ then there is a resolution $\langle R, \Theta, \rightarrow_R \rangle$ of Δ such that $\mathbb{V}_{\langle R, \Omega_\tau, \rightarrow_R \rangle}(\Theta) = \$\Delta'$.

(b) For any resolution $\langle R, \Theta, \rightarrow_R \rangle$ of Δ, there exists a Δ' such that $\Delta \Longrightarrow\!\!\!\succ \Delta'$ and $\mathbb{V}_{\langle R, \Omega_\tau, \rightarrow_R \rangle}(\Theta) = \$\Delta'$.

Proof: Suppose $\Delta \Longrightarrow\!\!\!\!\!\!\succ \Delta'$. By Proposition 4.16, there is a resolution $\langle R, \Theta, \to_R\rangle$ of Δ with resolving function f and a subdistribution Θ' such that $\Theta \Longrightarrow\!\!\!\!\!\!\succ \Theta'$ and $\Delta' = \mathrm{Img}_f(\Theta')$. By Lemma 4.20, we have $\mathbb{V}(\Theta) = \mathbb{V}(\Theta')$. Since Θ' is an extreme derivative, all the states s in its support are stable, so $\mathbb{V}(s)(\omega) = 0$ if $s \stackrel{\omega}{\not\longrightarrow}$, for all $\omega \in \Omega$. Hence

$$\mathbb{V}(\Theta')(\omega) = \mathrm{Exp}_{\Theta'}(\mathbb{V})(\omega) = \sum_{s \in \lceil \Theta' \rceil} \Theta'(s) \cdot \mathbb{V}(s)(\omega) = \sum_{s \in \lceil \Theta' \rceil,\ s \stackrel{\omega}{\longrightarrow}} \Theta'(s) = \$\Theta'(\omega)\ .$$

Furthermore, for all $t \in \lceil \Delta' \rceil$, $\Delta'(t) = \mathrm{Img}_f(\Theta')(t) = \sum_{f(s)=t} \Theta'(s)$, so, for all $\omega \in \Omega$,

$$\$\Delta'(\omega) = \sum_{t \in \lceil \Delta' \rceil,\ t \stackrel{\omega}{\longrightarrow}} \Delta'(t) = \sum_{t \in \lceil \Delta' \rceil,\ t \stackrel{\omega}{\longrightarrow}} \mathrm{Img}_f(\Theta')(t) = \sum_{t \in \lceil \Delta' \rceil,\ t \stackrel{\omega}{\longrightarrow}} \sum_{f(s)=t} \Theta'(s) = \sum_{s \in \lceil \Theta' \rceil,\ f(s) \stackrel{\omega}{\longrightarrow}} \Theta'(s) = \$\Theta'(\omega)\ ,$$

where in the last step we the use the property of resolutions that $f(s) \stackrel{\omega}{\longrightarrow}$ iff $s \stackrel{\omega}{\longrightarrow}$. It follows that $\mathbb{V}(\Theta) = \$\Delta'$.

To prove part (b), suppose that $\langle R, \Theta, \to_R\rangle$ is a resolution of Δ with resolving function f, so that $\Delta = \mathrm{Img}_f(\Theta)$. We know from Lemma 4.5 that there exists a (unique) subdistribution Θ' such that $\Theta \Longrightarrow\!\!\!\!\!\!\succ \Theta'$. By Proposition 4.17 we have that $\Delta \Longrightarrow\!\!\!\!\!\!\succ \mathrm{Img}_f(\Theta')$. The same arguments as in the other direction show that $\mathbb{V}(\Theta) = \$(\mathrm{Img}_f(\Theta'))$. □

Corollary 4.22 For any test T and process P we have that $\mathscr{A}^{\mathrm{r}}(T,P) = \mathscr{A}(T,P)$. □

5 An alternative approach to scalar testing

In the previous section our approach to testing involved two steps:

(1) For each test T and process P calculate a set of outcomes $\mathscr{A}(T,P)$; for scalar testing this is a subset of [0,1].

(2) For each pair of processes $P,\ Q$ compare the corresponding sets of outcomes $\mathscr{A}(T,P)$ and $\mathscr{A}(T,Q)$ for every test T.

But our methods for comparing sets of outcomes does not necessarily require us to calculate the entire set of outcomes. For closed sets $O_1,\ O_2 \in 2^{[0,1]}$ it is easy to check that

$$\begin{array}{lcl} O_1 \leq_{\mathrm{Ho}} O_2 & \text{if and only if} & \sup(O_1) \leq \sup(O_2) \\ O_1 \leq_{\mathrm{Sm}} O_2 & \text{if and only if} & \inf(O_1) \leq \inf(O_2)\ . \end{array}$$

Here we propose an alternative approach to testing based on calculating directly the *sup*s and *inf*s of the possible outcomes. We restrict our attention to scalar testing, i.e. the case where tests are allowed to use a *single* success action ω only; thus $\Omega = \{\omega\}$.

5.1 Extremal testing

The functional $\mathscr{R}$ used to associate an outcome with a resolution, is defined, in (10) above, only for deterministic pLTSs. Here we consider generalisations to an arbitrary pLTS $\langle S, \Omega_\tau, \to\rangle$.

Define the functional $\mathscr{R}_{\mathrm{inf}} : (S \to [0,1]) \to (S \to [0,1])$ by:

$$\mathscr{R}_{\mathrm{inf}}(f)(s) = \begin{cases} 1 & \text{if } s \stackrel{\omega}{\longrightarrow} \\ 0 & \text{if } s \stackrel{\omega}{\not\longrightarrow} \text{ and } s \stackrel{\tau}{\not\longrightarrow} \\ \inf\{\, \mathrm{Exp}_\Delta(f) \mid s \stackrel{\tau}{\longrightarrow} \Delta \,\} & \text{if s } \stackrel{\omega}{\not\longrightarrow} \text{ and } s \stackrel{\tau}{\longrightarrow} \end{cases}$$

In a similar fashion we can define the functional $\mathscr{R}_{\mathrm{sup}} : (S \to [0,1]) \to (S \to [0,1])$ which uses the $\sup$ function in place of $\inf$. Both these functions are monotonic, and therefore have least fixed points, which we abbreviate to $\mathbb{V}_{\mathrm{inf}},\ \mathbb{V}_{\mathrm{sup}}$ respectively.

Now for a test T and a process P, we have two ways of defining the outcome of the application of T to P:

$$\begin{array}{c} \mathscr{A}^{\mathrm{e}}_{\mathrm{inf}}(T,P) = \mathbb{V}_{\mathrm{inf}}([\![T \mid_{\mathsf{Act}} P]\!]) \\ \mathscr{A}^{\mathrm{e}}_{\mathrm{sup}}(T,P) = \mathbb{V}_{\mathrm{sup}}([\![T \mid_{\mathsf{Act}} P]\!])\ . \end{array}$$

Here $\mathscr{A}^{\mathrm{e}}_{\mathrm{inf}}(T,P)$ returns a single probability p, estimating the minimal probability of success; it is a pessimistic estimate. On the other hand $\mathscr{A}^{\mathrm{e}}_{\mathrm{sup}}(T,P)$ is optimistic, in that it gives the maximal probability of success.

Definition 5.1

1. $P \sqsubseteq^e_{\text{pmay}} Q$ if for every test T, $\mathscr{A}^{\text{e}}_{\text{sup}}(T,P) \leq \mathscr{A}^{\text{e}}_{\text{sup}}(T,Q)$;
2. $P \sqsubseteq^e_{\text{pmust}} Q$ if for every test T, $\mathscr{A}^{\text{e}}_{\text{inf}}(T,P) \leq \mathscr{A}^{\text{e}}_{\text{inf}}(T,Q)$.

The kernels of these preorders are denoted by $\simeq^e_{\text{pmay}}$ and $\simeq^e_{\text{pmust}}$, respectively.

Example 5.2 The pLTS-fragment in Figure 3(c), obtained by applying the test $T = a.\omega$ to the process Q_1, is deterministic and hence all three functions $\mathbb{V}_{\text{sup}},\ \mathbb{V}_{\text{inf}},\ \mathbb{V}$ coincide, giving $\mathscr{A}^{\text{e}}_{\text{sup}}(T,P) = \mathscr{A}^{\text{e}}_{\text{inf}}(T,P) = 1$. It follows that $Q_1 \simeq^e_{\text{pmay}} a$ and $Q_1 \simeq^e_{\text{pmust}} a$. □

Example 5.3 (revisiting Example 4.11 again) Here we reuse the notation of Example 4.11.

Consider the pLTS-fragment from Figure 4(c) resulting from the application of the test $T = a.\omega$ to the process Q_2. It is easy to see that the function $\mathbb{V}_{\text{sup}}$ satisfies

$$\begin{aligned} \mathbb{V}_{\text{sup}}(s_0) &= \max\{\frac{1}{2},\ x\} \\ x &= \frac{1}{2} + \frac{1}{2} \cdot \mathbb{V}_{\text{sup}}(s_0) \end{aligned} \tag{12}$$

It is not difficult to show that these equations have a unique solution, namely $\mathbb{V}_{\text{sup}}(s_0) = 1$. Since $\mathbb{V}_{\text{sup}}([\![T \mid_{\mathsf{Act}} a]\!]) = 1$ one can conclude that

$$Q_2 \simeq^e_{\text{pmay}} a\ .$$

If $\max$ is replaced by $\min$ in (12) above then the resulting equations also have a unique solution, giving $\mathbb{V}_{\text{inf}}(s_0) = \frac{1}{2}$. It follows that

$$a \not\sqsubseteq^e_{\text{pmust}} Q_2$$

because $\mathbb{V}_{\text{inf}}([\![T \mid_{\mathsf{Act}} a]\!]) = 1$. However, $Q_2 \sqsubseteq^e_{\text{pmust}} a$. □

Lemma 5.4 Consider an arbitrary pLTS $\langle S, \Omega_\tau, \rightarrow \rangle$.

(a) Both functionals $\mathscr{R}_{\text{inf}}$ and $\mathscr{R}_{\text{inf}}$ are continuous.

(b) Both results functions $\mathbb{V}_{\text{inf}}$ and $\mathbb{V}_{\text{sup}}$ are continuous.

Proof: Again the proof of part (a) is non-trivial; see Lemma B.5 in Appendix B.2. However part (b) is an immediate consequence. □

So in analogy with the evaluation function $\mathbb{V}$ from Section 4.2 these results functions can be captured by a chain of approximants:

$$\mathbb{V}_{\text{inf}} = \bigsqcup_{n \in \mathbb{N}} {\mathbb{V}_{\text{inf}}}^n \qquad \text{and} \qquad \mathbb{V}_{\text{sup}} = \bigsqcup_{n \in \mathbb{N}} {\mathbb{V}_{\text{sup}}}^n \tag{13}$$

where ${\mathbb{V}_{\text{inf}}}^0(s) = {\mathbb{V}_{\text{sup}}}^0(s) = 0$ for every state $s \in S$, and

- ${\mathbb{V}_{\text{inf}}}^{(k+1)} = \mathscr{R}_{\text{inf}}({\mathbb{V}_{\text{inf}}}^k)$
- ${\mathbb{V}_{\text{sup}}}^{(k+1)} = \mathscr{R}_{\text{sup}}({\mathbb{V}_{\text{sup}}}^k)$

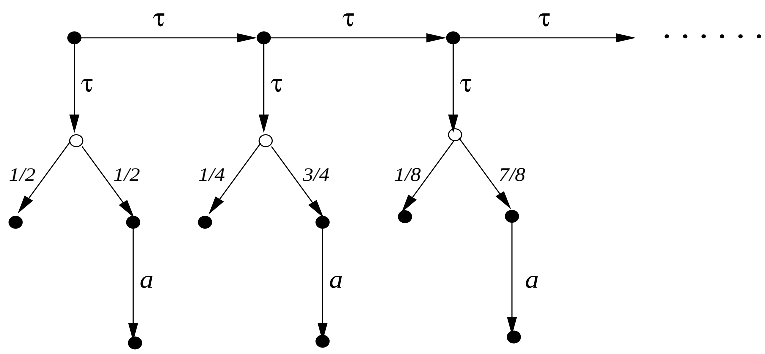


Figure 5: An infinite-state pLTS

5.2 Comparison with resolution-based testing

In this section we compare the two approaches of testing introduced in the previous two subsections. Our first result is that in the most general setting they lead to different testing preorders.

Example 5.5 Consider the infinite-state pLTS in Figure 5, which is defined as follows: in addition to the states a and $\mathbf{0}$ it has the infinite set $s_1, s_2, \ldots$, with each of these having two transitions:

- $s_k \xrightarrow{\tau} \overline{s_{k+1}}$
- $s_k \xrightarrow{\tau} [\![\, \mathbf{0} \,{}_{\frac{1}{2^k}}\!\oplus a]\!]$.

Now let us compare the state s_1 with the process a. With the test $a.\omega$, using resolutions, we get:

$$\begin{aligned} \mathscr{A}^{\mathrm{r}}(a.\omega, s_1) &= \updownarrow\{0, (1-\frac{1}{2}), \ldots, (1-\frac{1}{2^k}), \ldots\} \\ \mathscr{A}^{\mathrm{r}}(a.\omega, a) &= \{1\} \end{aligned} \tag{14}$$

which means that $a \not\sqsubseteq^{\Omega}_{\mathrm{pmay}} s_1$.

However when we use extremal testing, the test $a.\omega$ can not distinguish these processes. It is straightforward to see that $\mathbb{V}_{\mathsf{sup}}(a.\omega \mid_{\mathsf{Act}} a) = 1$. To see that $\mathbb{V}_{\mathsf{sup}}(a.\omega \mid_{\mathsf{Act}} s_1)$ also evaluates to 1, we let $x_k = \mathbb{V}_{\mathsf{sup}}(a.\omega \mid_{\mathsf{Act}} s_k)$, for all $k \geq 1$, and we have the following infinite equation system.

$$\begin{array}{rcl} x_1 &=& \max\{\frac{1}{2}, x_2\} \\ x_2 &=& \max\{1-\frac{1}{4}, x_3\} \\ &\vdots& \\ x_k &=& \max\{1-\frac{1}{2^k}, x_{k+1}\} \\ &\vdots& \end{array}$$

We have $x_k = 1$ for all $k \geq 1$ as the least solution of the above equation system.

With some more work one can go on to show that no test can distinguish between these processes using optimistic extremal testing, meaning that $a \sqsubseteq^{e}_{\mathrm{pmay}} s_1$.

□

In the remainder of this section we show that provided some finitary constraints are imposed on the pLTS extremal testing and resolution-based testing coincide; recall that here we are assuming that tests only use a single success action, $|\Omega| = 1$. First we examine *must* testing, which is easier than the *may* case; this in turn is treated in the following section.

5.2.1 Must testing

Here we show that provided we restrict our attention to finite-branching processes there is no difference between extremal *must* testing, and resolution-based *must* testing.

Let us consider a pLTS $\langle S, \Omega_\tau, \rightarrow\rangle$, obtained perhaps from applying a test T to a process P in $(T \mid_{\mathsf{Act}} P)$. We have two ways of obtaining a result for a distribution of states from S, by applying the function $\mathbb{V}_{\mathsf{inf}}$, or by using resolutions of the pLTS to realise $\mathbb{V}$. Our first result says that regardless of the actual resolution used, the value obtained from the latter will always dominate the former.
But first we need a technical lemma.

Lemma 5.6 Let $g : S \rightarrow [0,1],\ h : R \rightarrow [0,1]$ and $f : R \rightarrow S$ be three functions satisfying $g(f(r)) \leq h(r)$ for every $r \in R$. Then for every subdistribution Θ over R, $\mathrm{Exp}_\Delta(g) \leq \mathrm{Exp}_\Theta(h)$ where Δ denotes the subdistribution $\mathrm{Img}_f(\Theta)$.

Proof: A straightforward calculation. □

Proposition 5.7 *If $\langle R, \Theta, \rightarrow_R\rangle$ is a resolution of a subdistribution Δ then $Exp_\Delta(\mathbb{V}_{inf}) \leq Exp_\Theta(\mathbb{V})$.*

Proof: Let f denote the resolving function. First we show by induction on n that for every state $r \in R$

$$\mathbb{V}_{\mathsf{inf}}{}^n(f(r)) \leq \mathbb{V}^n(r) \qquad (15)$$

For $n = 0$, this is trivial. We consider the inductive step; note that by the previous lemma the inductive hypothesis implies that

$$\mathrm{Exp}_\Gamma(\mathbb{V}_{\mathsf{inf}}{}^n) \leq \mathrm{Exp}_\Theta(\mathbb{V}^n) \qquad (16)$$

for any pair of subdistributions satisfying $\Gamma = \mathrm{Img}_f(\Theta)$.

First if $r \xrightarrow{\omega}_R \Theta$, then $f(r) \xrightarrow{\omega}$, and thus $\mathbb{V}_{\mathsf{inf}}{}^{n+1}(f(r)) = 1 = \mathbb{V}^{n+1}(r)$. A similar argument applies if $r \not\longrightarrow$, that is $r \stackrel{\tau}{\not\longrightarrow}$ and $s \stackrel{\omega}{\not\longrightarrow}$. So the remaining possibility is that $r \xrightarrow{\tau}_R \Theta$ for some Θ, and $r \stackrel{\omega}{\not\longrightarrow}$, where we know $f(r) \xrightarrow{\tau} \mathrm{Img}_f(\Theta)$.

$$\begin{array}{rcll}
\mathbb{V}_{\mathsf{inf}}{}^{(n+1)}(f(r)) & = & min\{\mathrm{Exp}_\Delta(\mathbb{V}_{\mathsf{inf}}{}^n) \mid f(r) \xrightarrow{\tau} \Delta\} & \\
& \leq & \mathrm{Exp}_\Gamma(\mathbb{V}_{\mathsf{inf}}{}^n) & \text{where } \Gamma \text{ denotes } \mathrm{Img}_f(\Theta) \\
& \leq & \mathbb{V}^n(\Theta) & \text{by induction and (16) above} \\
& = & \mathbb{V}^{(n+1)}(r) &
\end{array}$$

Now by continuity we have from (15) that

$$\mathbb{V}_{\mathsf{inf}}(f(r)) \leq \mathbb{V}(r) \qquad (17)$$

The result now follows by the previous lemma, since if $\langle R, \Theta, \rightarrow_R\rangle$ is a resolution of a subdistribution Δ with resolving function f then by definition $\Delta = \mathrm{Img}_f(\Theta)$.

□

Our next result says that in any finite-branching computation structure we can find a resolution which realises the function $\mathbb{V}_{\mathsf{inf}}$. Moreover this resolution will be of a particularly simple form.

A resolution $\langle R, \Omega_\tau, \rightarrow_R\rangle$ is said to be *static* if its resolving function f_R is injective. Again we refer the reader to [5] for a discussion of power of this restriction. Static restrictions are particularly simple, in that they does not allow states to be resolved into distributions, or computation steps to be interpolated. Moreover they are very easy to describe.

Definition 5.8 A (static) *extreme policy* for a pLTS $\langle S, \Omega_\tau, \rightarrow\rangle$ is a partial function $\mathsf{epp} : S \rightharpoonup \mathscr{D}_1(S)$ satisfying:

(a) $s \xrightarrow{\omega}$ implies $s \xrightarrow{\omega} \mathsf{epp}(s)$

(b) otherwise, if $s \xrightarrow{\tau}$ then $s \xrightarrow{\tau} \mathsf{epp}(s)$

Intuitively an extreme policy epp determines a computation through the pLTS. But this set of possible computations, unlike resolutions as defined in Definition 4.12, are very restrictive. Policy epp decides at each state, once and for all, which of the available τ-choices to take; it does not interpolate, and since it is a function of the state, it makes the same choice on every visit. But there are two constraints:

(i) Condition (a) ensures an in-built preference for reporting success; if the state is successful the policy must also report success;

(ii) Condition (b), together with (a), means that $\mathsf{epp}(s)$ is defined whenever $s \longrightarrow$. This ensures that the policy cannot decide to stop at a state s if there is a possibility of proceeding from s; the computation must proceed, if it is possible to proceed.

We delay the formal definition of the computation determined by an extreme policy; see page 33. Here we are concerned with resolutions. An extreme policy epp determines a deterministic pLTS $\langle S, \Omega_\tau, \rightarrow_{\mathsf{epp}}\rangle$, where $\rightarrow_{\mathsf{epp}}$ is determined by $s \rightarrow_{\mathsf{epp}} \mathsf{epp}(s)$. Moreover for any subdistribution Δ over S it determines the obvious resolution $\langle S, \Delta, \rightarrow_{\mathsf{epp}}\rangle$, with the identity as the associated resolving function. Indeed it is possible to show that every static resolution is determined in this manner by some extreme policy.

Proposition 5.9 *Let Δ be any subdistribution in a finite-branching pLTS $\langle S, \Omega_\tau, \rightarrow\rangle$. Then there exists a static resolution of Δ, $\langle R, \Theta, \rightarrow_R\rangle$ such that $Exp_\Theta(\mathbb{V}) = Exp_\Delta(\mathbb{V}_{inf})$.*

Proof: We exhibit the required resolution by defining an extreme policy over S; in other words the resolution will take the form $\langle S, \Theta, \rightarrow_{\mathsf{epp}}\rangle$ for some extreme policy $\mathsf{epp}(-)$.

We say the extreme policy $\mathsf{epp}(-)$ is *min-seeking* if its domain is $\{\, s \in S \mid s \longrightarrow \}$ and it satisfies:

$$\text{if } s \overset{\omega}{\not\longrightarrow} \text{ but } s \overset{\tau}{\longrightarrow} \text{ then } \mathbb{V}_{\text{inf}}(\mathsf{epp}(s)) \leq \mathbb{V}_{\text{inf}}(\Delta) \;\text{ whenever } s \overset{\tau}{\longrightarrow} \Delta$$

Note that by design a min-seeking policy satisfies:

$$\text{if } s \overset{\omega}{\not\longrightarrow} \text{ but } s \overset{\tau}{\longrightarrow} \text{ then } \mathbb{V}_{\text{inf}}(s) = \mathbb{V}_{\text{inf}}(\mathsf{epp}(s)) \tag{18}$$

In a finite-branching pLTS it is straightforward to define a min-seeking extreme policy:

(i) If $s \overset{\omega}{\longrightarrow}$ then let $\mathsf{epp}(s)$ be any Δ such that $s \overset{\omega}{\longrightarrow} \Delta$.

(ii) Otherwise, if $s \overset{\tau}{\longrightarrow}$ let $\{\Delta_1, \ldots \Delta_n\}$ be the finite non-empty set $\{\, \Delta \mid s \overset{\tau}{\longrightarrow} \Delta \,\}$. Now let $\mathsf{epp}(s)$ be any Δ_k satisfying the property $\mathbb{V}_{\text{inf}}(\Delta_k) \leq \mathbb{V}_{\text{inf}}(\Delta_j)$ for every $1 \leq j \leq n$; at least one such Δ_k must exist.

We now show that the static resolution determined by such a policy, $\langle S, \Theta, \rightarrow_{\mathsf{epp}}\rangle$, satisfies the requirements of the proposition. For the sake of clarity let us write $\mathbb{V}_{\mathsf{epp}}(\Delta)$ for the value realised for Δ in this resolution.

We already know, from Proposition 5.7, that $\mathbb{V}_{\text{inf}}(\Delta) \leq \mathbb{V}_{\mathsf{epp}}(\Delta)$ and so we concentrate on the converse, $\mathbb{V}_{\mathsf{epp}}(\Delta) \leq \mathbb{V}_{\text{inf}}(\Delta)$. Recall that the function $\mathbb{V}_{\mathsf{epp}}$ is the least fixed point of the functional $\mathscr{R}$ defined in (10) above, and interpreted in the above resolution. So the result follows if we can show that the function $\mathbb{V}_{\text{inf}}$ is also a fixed point. Since $|\Omega| = 1$ this amounts to proving

$$\mathbb{V}_{\text{inf}}(s) = \begin{cases} 1 & \text{if } s \overset{\omega}{\longrightarrow} \\ 0 & \text{if } s \not\longrightarrow \\ \mathbb{V}_{\text{inf}}(\mathsf{epp}(s)) & \text{otherwise} \end{cases}$$

However this is a straightforward consequence of (18) above. □

Theorem 5.10 For finite-branching processes, $P \sqsubseteq^e_{\text{pmust}} Q$ if and only if $P \sqsubseteq_{\text{pmust}} Q$

Proof: This is a consequence of the two previous propositions. First suppose $P \sqsubseteq^e_{\text{pmust}} Q$. To show $P \sqsubseteq_{\text{pmust}} Q$ we must show that for any value v in $\mathscr{A}(T, Q)$, for any arbitrary test T, there exists some $v' \in \mathscr{A}(T, P)$ such that $v' \leq v$. The value v must be of the form $\mathbb{V}(\Theta_R)$, for some resolution $\langle R, \Theta_R, \rightarrow_R\rangle$ of $[[\![T \mid_{\mathsf{Act}} P]\!]]$. From Proposition 5.7 we know that $\mathbb{V}_{\text{inf}}([[\![T \mid_{\mathsf{Act}} Q]\!]]) \leq v$, and now from the hypothesis $P \sqsubseteq^e_{\text{pmust}} Q$ we have that $\mathbb{V}_{\text{inf}}([[\![T \mid_{\mathsf{Act}} P]\!]]) \leq v$. Now employing Proposition 5.9 we can find some other (static) resolution $\langle S, \Theta_S, \rightarrow_S\rangle$ of $[[\![Q \mid_{\mathsf{Act}} P]\!]]$ and such that $\mathbb{V}(\Theta_S) = \mathbb{V}_{\text{inf}}([[\![Q \mid_{\mathsf{Act}} P]\!]])$. So we can take the required v' to be $\mathbb{V}(\Theta_S)$.

The converse, $P \sqsubseteq_{\text{pmust}} Q$ implies $P \sqsubseteq^e_{\text{pmust}} Q$ is equally straightforward, and is left to the reader. □

5.2.2 May testing

Here we can try to apply the same proof strategy as in the previous section. The analogue to Proposition 5.7 goes through:

Proposition 5.11 *If $\langle R, \Theta, \rightarrow_R\rangle$ is a resolution of Δ then $Exp_\Theta(\mathbb{V}) \leq Exp_\Delta(\mathbb{V}_{sup})$.*

Proof: Similar to the proof of Proposition 5.7 □

However the proof strategy used in Proposition 5.9 cannot be used to show that $\mathbb{V}_{\text{sup}}$ can be realised by some static resolution, as the following example shows.

Example 5.12 In analogy with the definition used in the proof of Proposition 5.9, we say that an extreme policy $\mathsf{epp}(-)$ is *max-seeking* if its domain is precisely $\{\, s \in S \mid s \longrightarrow \}$, and

$$\text{if } s \stackrel{\omega}{\not\longrightarrow} \text{ but } s \stackrel{\tau}{\longrightarrow} \text{ then } \mathbb{V}_{\text{sup}}(\Delta) \leq \mathbb{V}_{\text{sup}}(\mathsf{epp}(s)) \quad \text{whenever } s \stackrel{\tau}{\longrightarrow} \Delta$$

This ensures that $\mathbb{V}_{\text{sup}}(s) = \mathbb{V}_{\text{sup}}(\mathsf{epp}(s))$, whenever $s \stackrel{\tau}{\longrightarrow}$ and $s \stackrel{\omega}{\not\longrightarrow}$, and again it is straightforward to define a max-seeking extreme policy in a finite-branching pLTS. However the resulting resolution does not in general realise the function $\mathbb{V}_{\text{sup}}$.

To see this, let us consider the (finite-branching) pLTS used in Example 5.5. Here in addition to the two states ω and $\mathbf{0}$ there is the infinite set $\{s_1, \ldots s_k, \ldots\}$ and the transitions

- $s_k \stackrel{\tau}{\longrightarrow} \overline{s_{k+1}}$
- $s_k \stackrel{\tau}{\longrightarrow} [\![\, \mathbf{0} \,{}_{\frac{1}{2^k}}\!\oplus \omega]\!]$.

One can calculate $\mathbb{V}_{\text{sup}}(s_k)$ to be 1 for every k, and a max-seeking extreme policy is determined by $\mathsf{epp}(s_k) = \overline{s_{k+1}}$; indeed this is essentially the only such policy. However this resolution associated with this policy does not realise $\mathbb{V}_{\text{sup}}$, as $\mathbb{V}_{\mathsf{epp}}(s_k) = 0$. □

Nevertheless we will show that if we restrict attention to finitary pLTSs, then there will always exist some static resolution which realises $\mathbb{V}_{\text{sup}}$. The proof relies on techniques used in Markov process theory [22], and unlike that of Proposition 5.9 is non-constructive; we simply prove that some such resolution exists, without actually showing how to construct it.

Theorem 5.13 Let Δ be any subdistribution in a finitary pLTS. Then there exists a static resolution of Δ, $\langle R, \Theta, \rightarrow_R\rangle$ such that $\text{Exp}_\Theta(\mathbb{V}) = \text{Exp}_\Delta(\mathbb{V}_{\text{sup}})$.

Proof: The proof is non-trivial and lengthy as it involves the development of *discounted* policies for pLTSs, based on discounted results-collecting functions like $\mathbb{V}^\delta$ and ${\mathbb{V}_{\text{sup}}}^\delta$ for discount factor δ. Although such techniques are relatively standard in the theory of Markov Decision Processes, see [22] for example, they are virtually unknown in concurrency theory. Consequently we relegate the proof to Appendix B; this enables us to give a detailed exposition without interfering with the overall flow of the paper. The exposition cumulates in Theorem B.8. □

Theorem 5.14 For finitary processes, $P \sqsubseteq^e_{\text{pmay}} Q$ if and only if $P \sqsubseteq_{\text{pmay}} Q$.

Proof: Similar to that of Theorem 5.10 but employing Theorem 5.13 in place of Proposition 5.9. □

6 Generating weak derivatives in a finitary pLTS

Now let us restrict our attention to finitary pLTSs, where the state space is $S = \{s_1, \ldots, s_n\}$. Here by definition the sets $\{\, \Theta \mid s \stackrel{\alpha}{\longrightarrow} \Theta \,\}$ are finite, for every state s and label α. This of course is no longer true for the weak arrows; the sets $\{\, \Theta \mid \overline{s} \stackrel{\alpha}{\Longrightarrow} \Theta \,\}$ are in general not finite, because of the infinitary nature of the weak derivative relation $\Longrightarrow$. The purpose of this section is to show that nevertheless they can be finitely represented, at least for finitary pLTSs.

This is explained in Section 6.1, and the ramifications are then explored in the following subsection. These include a very useful topological property of these sets of derivatives; they are *closed* in the sense (from analysis) of containing all its limit points where, in turn, limit depends on a Euclidean-style metric defining the distance between two distributions in a straightforward way. Another consequence is that we can find in any derivation that partially diverges (by no matter how small an amount) a point at which the divergence is *distilled* into a state which wholly diverges; we call this *distillation of divergence*.

6.1 Finite generability

A subdistribution over the finite state space S can now be viewed as a point in $\mathbb{R}^n$, and therefore a set of subdistributions, such as the set of weak derivatives $\{ \Delta \mid \overline{s} \Longrightarrow \Delta \}$ corresponds to a subset of $\mathbb{R}^n$. We endow $\mathbb{R}^n$ with the standard Euclidean metric and proceed to establish useful topological properties of such sets of subdistributions. Recall that a set $X \subseteq \mathbb{R}^n$ is (Cauchy) *closed* if for every Cauchy sequence $\{ x_n \mid n \geq 0 \}$ with limit x, if $x_n \in X$ for every $n \geq 0$ then x is also in X.

Lemma 6.1 If X is a finite subset of $\mathbb{R}^n$ then $\updownarrow X$ is closed.

Proof: Straightforward. □

In Definition 5.8 we gave a definition of extreme policies for pLTSs of the form $\langle S, \Omega_\tau, \rightarrow \rangle$ and showed how they determine resolutions. Here we generalise these to *derivative policies* and show that these generalised policies can also be used to generate arbitrary weak derivatives of subdistributions over S.

Definition 6.2 A (static) *derivative policy* for a pLTS $\langle S, \mathsf{Act}_\tau, \rightarrow \rangle$, is a partial function $\mathsf{dpp} : S \rightharpoonup \mathscr{D}_1(S)$ with the property that $\mathsf{dpp}(s) = \Delta$ implies $s \xrightarrow{\tau} \Delta$. If dpp is undefined at s, we write $\mathsf{dpp}(s)\uparrow$. Otherwise, we write $\mathsf{dpp}(s)\downarrow$.

A derivative policy dpp, as its name suggests, can be used to guide the derivation of a weak derivative. Suppose $\overline{s} \Longrightarrow \Delta$, using a derivation as given in Definition 3.13. Then we write $\overline{s} \Longrightarrow_{\mathsf{dpp}} \Delta$ whenever, for all $k \geq 0$,

(a) $\Delta_k^{\rightarrow}(s) = \begin{cases} \Delta_k(s) & \text{if } \mathsf{dpp}(s)\downarrow \\ 0 & \text{otherwise} \end{cases}$

(b) $\Delta_{(k+1)} = \sum_{s \in \lceil \Delta_k^{\rightarrow} \rceil} \Delta_k^{\rightarrow}(s) \cdot \mathsf{dpp}(s)$

Intuitively these conditions mean that the derivation of Δ from s is guided at each stage by the policy dpp:

- Condition (a) implies that the division of Δ_k into $\Delta_k^{\rightarrow}$, the subdistribution which will continue marching, and $\Delta_k^{\times}$, the subdistribution which will stop, is determined by the domain of the derivative policy dpp.
- Condition (b) ensures that the derivation of the next stage Δ_{k+1} from $\Delta_k^{\rightarrow}$ is determined by the action of the function dpp on the support of $\Delta_k^{\rightarrow}$.

Lemma 6.3 Let dpp be derivative policy in a pLTS. Then

(a) If $\overline{s} \Longrightarrow_{\mathsf{dpp}} \Delta$ and $\overline{s} \Longrightarrow_{\mathsf{dpp}} \Theta$ then $\Delta = \Theta$.

(b) For every state s there exists some Δ such that $\overline{s} \Longrightarrow_{\mathsf{dpp}} \Delta$.

Proof: To prove part (a) consider the derivation of $\overline{s} \Longrightarrow \Delta$ and $\overline{s} \Longrightarrow \Theta$ as in Definition 3.13, via the subdistributions $\Delta_k,\ \Delta_k^{\rightarrow},\ \Delta_k^{\times}$ and $\Theta_k,\ \Theta_k^{\rightarrow},\ \Theta_k^{\times}$ respectively. Because both derivations are guided by the same derivative policy dpp it is easy to show by induction on k that

$$\Delta_k = \Theta_k \qquad \Delta_k^{\rightarrow} = \Theta_k^{\rightarrow} \qquad \Delta_k^{\times} = \Theta_k^{\times}$$

from which $\Delta = \Theta$ follows immediately.

To prove (b) we use dpp to generate subdistributions $\Delta_k,\ \Delta_k^{\rightarrow},\ \Delta_k^{\times}$ for each $k \geq 0$ satisfying the constraints of Definition 3.13 and simultaneously those in Definition 6.2 above. The result will then follow by letting Δ be $\sum_{k \geq 0} \Delta_k^{\times}$. □

The net effect of this lemma is that a derivative policy dpp determines a *total* function from states to derivations. Let $\mathbf{Der}_{\mathsf{dpp}} : S \to \mathscr{D}_1(S)$ be defined by letting $\mathbf{Der}_{\mathsf{dpp}}(s)$ be the unique Δ such that $\overline{s} \Longrightarrow_{\mathsf{dpp}} \Delta$.

It should be clear that the use of derivative policies limits considerably the scope for deriving weak derivations. Each particular policy can only derive one weak derivative, and moreover in finitary pLTS there are only a finite number of derivative policies. Nevertheless we will show that this limitation is more apparent than real. In Section 5.2.1 we saw how the more restrictive extreme policies epp could in fact realise the maximum value attainable by any resolution of a finitely branching pLTS. Here we generalise this result by replacing resolutions with arbitrary weight functions.

Definition 6.4 [Weights and payoffs] A *weight function* is a function $\mathbf{w} : S \to [-1, 1]$. With $S = \{s_1, \ldots, s_n\}$ we often consider a weight function as the n-dimensional vector $\langle\, \mathbf{w}(s_1), ..., \mathbf{w}(s_n)\,\rangle$. In this way, we can use the notion $\mathbf{w} \centerdot \Delta$ to stand for the inner product of two vectors.

Given such a weight function, we define the payoff function $\mathbb{P}^{\mathbf{w}}_{\max} : S \to \mathbb{R}$ by

$$\mathbb{P}^{\mathbf{w}}_{\max}(s) \;=\; \sup\{\, \mathbf{w} \centerdot \Delta \mid \overline{s} \Longrightarrow \Delta \,\}$$

A priori these payoff functions for a given state s are determined by the set of weak derivatives of s. However the main result of this section is that they can in fact always be realised by derivative policies.

Theorem 6.5 (Realising payoffs) In a finitary pLTS, for every weight function **w** there exists some derivative policy dpp such that $\mathbb{P}^{\mathbf{w}}_{\max}(s) \;=\; \mathbf{w} \centerdot \mathbf{Der}_{\mathsf{dpp}}(s)$

Proof: As with Theorem 5.13 there is a temptation to give a constructive proof here, defining the effect of the required derivative policy dpp at state s by considering the application of the weight function **w** to both s and all of its derivatives - a finite set. However this is not possible, as the example below explains.

Instead the proof is non-constructive, requiring *discounted* policies. The overall structure of the proof is similar to that of Theorem 5.13, but the use of (discounted) derivative policies rather than extreme policies makes the details considerably different. Consequently the proof is spelled out in some detail in Appendix A, cumulating in Theorem A.15. □

Example 6.6 Let us say that a derivative policy dpp is max-seeking with respect to a weight function **w** if for all $s \in S$ the following requirements are met.

1. If $\mathsf{dpp}(s)\uparrow$ then $\mathbf{w}(s) \geq \mathbb{P}^{\mathbf{w}}_{\max}(\Delta_1)$ for all $s \xrightarrow{\tau} \Delta_1$.
2. If $\mathsf{dpp}(s)\downarrow = \Delta$ then
   (a) $\mathbb{P}^{\mathbf{w}}_{\max}(\Delta) \geq \mathbf{w}(s)$ and
   (b) $\mathbb{P}^{\mathbf{w}}_{\max}(\Delta) \geq \mathbb{P}^{\mathbf{w}}_{\max}(\Delta_1)$ for all $s \xrightarrow{\tau} \Delta_1$.

What a max-seeking policy does is to evaluate $\mathbb{P}^{\mathbf{w}}_{\max}$ in advance, for a given weight function **w**, and then label each state s with the payoff value $\mathbb{P}^{\mathbf{w}}_{\max}(s)$. The policy at any state s is then to compare $\mathbf{w}(s)$ with the expected label values $\mathbb{P}^{\mathbf{w}}_{\max}(\Delta')$ (i.e. $\mathrm{Exp}_{\Delta'}(\mathbb{P}^{\mathbf{w}}_{\max})$) for each outgoing transition $s \xrightarrow{\tau} \Delta'$ and then to select the greatest among all those values. Note that for the policy to be well defined, we require that the pLTS under consideration is finitely branching.

In case that seems obvious, we now consider the pLTS in Figure 6 and let us apply the above definition of max-seeking policies to the weight function given by $\mathbf{w}(s_0) = 0,\ \mathbf{w}(s_1) = 1$. For both states a payoff of 1 is attainable eventually, thus $\mathbb{P}^{\mathbf{w}}_{\max}(s_0) = \mathbb{P}^{\mathbf{w}}_{\max}(s_1) = 1$, because we have $s_0 \Longrightarrow \overline{s_1}$ and $s_1 \Longrightarrow \overline{s_1}$. Hence, both states will be $\mathbb{P}^{\mathbf{w}}_{\max}$-labelled with 1. At state s_0 the policy then makes a choice among three options: (1) to stay unmoved, yielding immediate payoff $\mathbf{w}(s_0) = 0$; (2) to take the transition $s_0 \xrightarrow{\tau} \overline{s_0}$; (3) to take the transition $s_0 \xrightarrow{\tau} \overline{s_0}\,{}_{1/2}\!\oplus \overline{s_1}$. Clearly one of the latter two is chosen — but which? If it is the second, then indeed the maximum payoff 1 can be achieved. If it is the first, then in fact the overall payoff will be 0 because of divergence, so the policy would fail to attain the maximum payoff 1.

However, for properly discounted max-seeking policies, we show in Proposition A.12 that they always attain the maximum payoffs. □

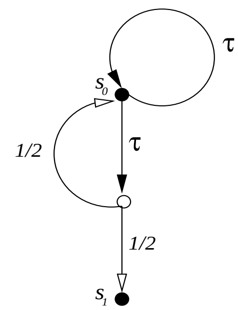


Figure 6: Max-seeking policies

With Theorem 6.5 at hand, we are in the position to prove the main result of this section, which says that in a finitary pLTS the set of weak derivatives from any state s, $\{\Delta \mid \overline{s} \Longrightarrow \Delta\}$, is generable by the convex closure of a finite set. But we first need a tool, the *Separating Hyperplane Lemma* from discrete geometry [18, Theorem 1.2.4 paraphrased].

Lemma 6.7 Let E and F be two convex- and Cauchy-closed subsets of the n-dimensional Euclidean space; assume that they are disjoint and that at least one of them is bounded. Then there is a hyperplane that strictly separates them.

Here a *hyperplane* is a set of the form $\{g \in \mathbb{R}^n \mid h \centerdot g = c\}$ for certain $h \in \mathbb{R}^n$ (the *normal* of the hyperplane) and $c \in \mathbb{R}$, and such a hyperplane *strictly separates* E and F if for all $e \in E$ and $f \in F$ we have $h \centerdot e < c < h \centerdot f$ or $h \centerdot e > c > h \centerdot f$.

Theorem 6.8 (Finite generability) Let $P = \{\mathsf{dpp}_1,\ \ldots,\ \mathsf{dpp}_k\}$ be the finite set of derivative policies in a finitary pLTS. Then $\overline{s} \Longrightarrow \Delta$ implies $\Delta \in \updownarrow\{\, \mathbf{Der}_{\mathsf{dpp}_i}(s) \mid 1 \le i \le k\,\}$.

Proof: For convenience let X denote the set $\updownarrow\{\, \mathbf{Der}_{\mathsf{dpp}_i}(s) \mid 1 \le i \le k\,\}$. Suppose, for a contradiction, that $\overline{s} \Longrightarrow \Delta$ for some Δ not in X. Recall that we are assuming that the underlying state space is $S = \{x_1, \ldots x_n\}$ so that X is a subset of $\mathbb{R}^n$. It is trivially bounded by $[-1, 1]^n$, and by definition it is convex-closed; by Lemma 6.1 it follows that X is (Cauchy) closed.

By the Separating Hyperplane Lemma, Lemma 6.7, Δ can be separated from X by a hyperplane H. What this means is that there is some function $\mathbf{w}_H : S \to \mathbb{R}$ and constant $c \in \mathbb{R}$ such that either

(a) $\mathbf{w}_H \centerdot \Theta < c$ for all $\Theta \in X$ and $\mathbf{w}_H \centerdot \Delta > c$

(b) or, $\mathbf{w}_H \centerdot \Theta > c$ for all $\Theta \in X$ and $\mathbf{w}_H \centerdot \Delta < c$

In fact from case (b) we can obtain case (a) by negating both the constant c and the components of the function $\mathbf{w}_H$; so we can assume (a) to be true. Moreover by scaling with respect to the largest $\mathbf{w}_H(s_i),\ 1 \le i \le n$, we can assume that $\mathbf{w}_H$ is actually a weight function.

In particular (a) means that $\mathbf{w}_H \centerdot \mathbf{Der}_{\mathsf{dpp}_i}(s) < c$, and therefore that $\mathbf{w}_H \centerdot \mathbf{Der}_{\mathsf{dpp}_i}(s) < \mathbf{w}_H \centerdot \Delta$, for each of derivative polices dpp_i. But this contradicts Theorem 6.5 which claims that there must be some $1 \le i \le n$ such that $\mathbf{w}_H \centerdot \mathbf{Der}_{\mathsf{dpp}_i}(s) = \mathbb{P}^{\mathbf{w}_H}_{\max}(s) \ge \mathbf{w}_H \centerdot \Delta$. $\square$

Note that by definition $\overline{s} \Longrightarrow \mathbf{Der}_{\mathsf{dpp}}(s)$ for every derivative policy dpp. So it follows immediately from this theorem that in a finitary pLTS the set of weak derivatives from the distribution $\overline{s}$ is exactly the convex closure of the finite set $\{\mathbf{Der}_{\mathsf{dpp}_1}(s),\ \ldots,\ \mathbf{Der}_{\mathsf{dpp}_n}\}$.

Extreme policies, as given in Definition 5.8, are particular kinds of derivative policies, designed for pLTSs of the form $\langle R, \Omega_\tau, \to_R\rangle$. The significant constraint on extreme policies is that for any state s if $s \xrightarrow{\tau}$ then $\mathsf{epp}(s)$ must be defined. As a consequence in the computation determined by epp if a state can contribute to the computation at any stage it must contribute.

Lemma 6.9 Let epp be any extreme policy. Then $\overline{s} \Longrightarrow_{\mathsf{epp}} \Delta$ implies $\overline{s} \Longrightarrow\!\!\!\succ \Delta$.

Proof: Consider the derivation of Δ as in Definition 3.13, and determined by the extreme policy epp. Since $\Delta = \sum_{k\geq 0} \Delta_k^{\times}$ it is sufficient to show that each $\Delta_k^{\times}$ is stable, that is $s \xrightarrow{\tau}$ implies $s \notin \lceil \Delta_k^{\times} \rceil$.

Since epp is an extreme policy, Definition 5.8 ensures that $\mathsf{epp}(s)$ is defined. From the definition of a computation, Definition 3.13, we know $\Delta_k = \Delta_k^{\rightarrow} + \Delta_k^{\times}$ and since the computation is guided by the policy epp we have that $\Delta_k^{\rightarrow}(s) = \Delta_k(s)$. An immediate consequence is that $\Delta_k^{\times}(s) = 0$. □

As a consequence the finite generability result, Theorem 6.8, specialises to extreme derivatives.

Corollary 6.10 Let $\{\mathsf{epp}_1, ..., \mathsf{epp}_k\}$ be the finite set of extreme policies of a finitary ω-respecting pLTS $\langle S, \Omega_\tau, \rightarrow_R \rangle$. Then $\overline{s} \Longrightarrow\!\!\!\succ \Delta$ if and only if $\Delta \in \updownarrow\{\, \mathbf{Der}_{\mathsf{epp}_i}(s) \mid 1 \leq i \leq k \,\}$.

Proof: One direction follows immediately from the previous lemma. Conversely suppose $\overline{s} \Longrightarrow\!\!\!\succ \Delta$. By Theorem 6.8 $\Delta = \sum_{1\leq i\leq n} p_i \cdot \mathbf{Der}_{\mathsf{dpp}_i}(s)$ for some finite collection of derivative policies dpp_i, where we can assume that each $p_i \geq 0$. Because Δ is stable, that is $s \not\xrightarrow{\tau}$ for every $s \in \lceil \Delta \rceil$, we show that each derivative policy dpp_i can be transformed into an extreme policy epp_i such that $\mathbf{Der}_{\mathsf{epp}_i}(s) = \mathbf{Der}_{\mathsf{dpp}_i}(s)$, from which the result will follow.

First note it is sufficient to define epp_i on the set of states t accessible from s via the policy dpp_i; on the remaining states in S epp_i can be defined arbitrarily, so as to satisfy the requirements of Definition 5.8. So consider the derivation of $\mathbf{Der}_{\mathsf{dpp}_i}(s)$ as in Definition 3.13, determined by dpp_i and suppose $t \in \Delta_k$ for some $k \geq 0$. There are three cases:

(i) Suppose $t \xrightarrow{\tau}$. Since Δ is stable we know $t \notin \lceil \Delta_k^{\times} \rceil$, and therefore by definition $\mathsf{dpp}_i(t)$ is defined. So in this case we let $\mathsf{epp}_i(t)$ be the same as $\mathsf{dpp}_i(t)$.

(ii) Suppose $t \xrightarrow{\omega}$, in which case, since the pLTS is ω-respecting, we know $t \not\xrightarrow{\tau}$, and therefore $\mathsf{dpp}_i(t)$ is not defined. Here we choose $\mathsf{epp}_i(t)$ arbitrarily so as to satisfy $t \xrightarrow{\omega} \mathsf{epp}_i(t)$.

(iii) Otherwise we leave $\mathsf{epp}(t)$ undefined.

By definition epp_i is an extreme policy since it satisfies conditions (a) and (b) in Definition 5.8, and by construction $\mathbf{Der}_{\mathsf{epp}_i}(s) = \mathbf{Der}_{\mathsf{dpp}_i}(s)$. □

This corollary gives a useful method for calculating the set of extreme derivatives of a given state, and therefore of the result of applying a test to a process.

Example 6.11 Consider again Figure 4, discussed in Example 4.11, where we have the ω-respecting pLTS obtained by applying the test $a.\omega$ to the process Q_2. There are only two extreme policies for this pLTS, denoted by epp_0 and epp_1. They differ only for the state s_1, with $\mathsf{epp}_0(s_1) = \Theta_0$ and $\mathsf{epp}_1(s_1) = \Theta_1$. The discussion in Example 4.11 explained how

$$\mathbf{Der}_{\mathsf{epp}_0}(s_1) = \overline{\omega} \qquad \mathbf{Der}_{\mathsf{epp}_1}(s_1) = \frac{1}{2}\overline{s_3} + \frac{1}{2}\overline{\omega}$$

By Corollary 6.10 we know that every possible extreme derivative of $[\![T \mid_{\mathsf{Act}} Q_2]\!]$ takes the form

$$q \cdot \overline{\omega} + (1-q) \cdot (\frac{1}{2}\overline{s_3} + \frac{1}{2}\overline{\omega})$$

for some $0 \leq q \leq 1$. Since $\$(\overline{\omega}) = 1$ and $\$(\frac{1}{2}\overline{s_3} + \frac{1}{2}\overline{\omega}) = \frac{1}{2}$ it follows that $\mathscr{A}(T, Q_2) = [\frac{1}{2}, 1]$. □

6.2 Consequences

In this section we outline two major consequences of Theorem 6.8, which informally means that the set of weak derivatives from a given state is the convex-closure of a finite set. The first is straightforward, and is explained in the following two results.

Lemma 6.12 (Closure of $\Longrightarrow$) For any state s in a finitary pLTS the set of derivatives $\{\, \Delta \mid \overline{s} \Longrightarrow \Delta \,\}$ is closed and convex.

Proof: Let $\mathsf{dpp}_1, ..., \mathsf{dpp}_n$ ($n \geq 1$) be all the derivative policies in the finitary pLTS. Consider two sets $C = \updownarrow\{\, \mathbf{Der}_{\mathsf{dpp}_i}(s) \mid 1 \leq i \leq n \,\}$ and $D = \{\Delta' \mid \Delta \Longrightarrow \Delta'\}$. By Theorem 6.8 we have $D \subseteq C$. On the other hand, it is easy to see that D is convex and thus $C \subseteq D$. Therefore, D coincides with C, the convex closure of a finite set. By Lemma 6.1, it is also Cauchy closed. □

The restriction here to finitary pLTSs is essential, as the following examples demonstrate.

Example 6.13 Consider the finite state but infinitely branching pLTS containing three states $s_1,\ s_2,\ s_3$ and the countable set of transitions given by

$$s_1 \xrightarrow{\tau} (\overline{s_2}\ {}_{\frac{1}{2^n}}\!\oplus \overline{s_3}) \qquad n \geq 1$$

For convenience let Δ_n denote the distribution $(\overline{s_2}\ {}_{\frac{1}{2^n}}\!\oplus \overline{s_3})$. Then $\{\, \Delta_n \mid n \geq 1 \,\}$ is a Cauchy sequence with limit $\overline{s_3}$. Trivially the set $\{\, \Delta \mid \overline{s_1} \Longrightarrow \Delta \,\}$ contains every Δ_n, but it does not contain the limit of the sequence, thus it is not closed. □

Example 6.14 By adapting Example 6.13, we obtain the following pLTS which is finitely branching but has infinitely many states. Let t_1 and t_2 be two distinct states. Moreover, for each $n \geq 1$ there is a state s_n with two outgoing transitions: $s_n \xrightarrow{\tau} \overline{s_{n+1}}$ and $s_n \xrightarrow{\tau} \overline{t_1}\ {}_{\frac{1}{2^n}}\!\oplus \overline{t_2}$. Let Δ_n denote the distribution $\overline{t_1}\ {}_{\frac{1}{2^n}}\!\oplus \overline{t_2}$. Then $\{\Delta_n \mid n \geq 1\}$ is a Cauchy sequence with limit $\overline{t_2}$. The set $\{\, \Delta \mid \overline{s_1} \Longrightarrow \Delta \,\}$ is not closed because it contains each Δ_n but not the limit $\overline{t_2}$. □

Corollary 6.15 [Closure of $\overset{a}{\Longrightarrow}$] For any state s in a finitary pLTS the set $\{\Delta \mid \overline{s} \overset{a}{\Longrightarrow} \Delta\}$ is closed and convex.

Proof: We first introduce a preliminary concept. We say a subset $D \subseteq \mathscr{D}(S)$ is *finitely generable* whenever there is some finite set $F \subseteq \mathscr{D}(S)$ such that $D = \updownarrow F$. A relation $\mathscr{R} \subseteq X \times \mathscr{D}(S)$ is *finitely generable* if for every x in X the set $x\cdot \mathscr{R}$ is finitely generable. We observe that

(i) If a set is finitely generable, then it is closed and convex.

(ii) If $\mathscr{R}_1, \mathscr{R}_2 \subseteq \mathscr{D}(S) \times \mathscr{D}(S)$ are finitely generable then so is their composition $\mathscr{R}_1; \mathscr{R}_2$.

The first property is a direct consequence of the definition of finite generability. To prove the second property, we let $\mathscr{B}^i_\Phi$ be a finite set of subdistributions such that $\Phi\cdot \mathscr{R}_i = \updownarrow \mathscr{B}^i_\Phi$ for $i = 1, 2$. Then one can check that

$$\Delta\cdot \mathscr{R}_1; \mathscr{R}_2 \;=\; \updownarrow \cup \{\, \mathscr{B}^2_\Theta \mid \Theta \in \mathscr{B}^1_\Delta \,\}$$

which implies that finite generability is preserved under composition of relations.

Notice that the relation $\overset{a}{\Longrightarrow}$ is a composition of three stages: $\Longrightarrow; \xrightarrow{a}; \Longrightarrow$. In the proof of Lemma 6.12 we have shown that $\Longrightarrow$ is finitely generable. In a finitary pLTS, the relation $\xrightarrow{a}$ is also finitely generable. It follows from property (ii) that $\overset{a}{\Longrightarrow}$ is finitely generable. By property (i) we have that $\overset{a}{\Longrightarrow}$ is closed and convex. □

Corollary 6.16 In a finitary pLTS, the relation $\overset{a}{\Longrightarrow}$ is the lifting of the closed and convex relation $\Longrightarrow_S \xrightarrow{a} \Longrightarrow$, where $s \Longrightarrow_S \Delta$ means $\overline{s} \Longrightarrow \Delta$.

Proof: The relation $\Longrightarrow_S \xrightarrow{a} \Longrightarrow$ is $\overset{a}{\Longrightarrow}$ restricted to point distributions. We have shown that $\overset{a}{\Longrightarrow}$ is closed and convex in Corollary 6.15. Therefore, $\Longrightarrow_S \xrightarrow{\mu} \Longrightarrow$ is closed and convex. Its lifting coincides with $\overset{\mu}{\Longrightarrow}$, which can be shown by some arguments analogous to those in the proof of Proposition 3.21. □

The second consequence of Theorem 6.8 concerns the manner in which divergent computations arise in pLTSs. Consider again the infinite state pLTS given in Example 3.17. There is no state s which wholly diverges, that is satisfying $s \Longrightarrow \varepsilon$, yet there are many partially divergent computations. In fact for every $k \geq 2$ we have $s_k \Longrightarrow \frac{1}{k}\overline{a}$. This can not arise in a finitary pLTS; if there is any partial derivation in a finitary pLTS, $\Delta \Longrightarrow \Delta'$ with $|\Delta| > |\Delta'|$, then there is some state in the pLTS which wholly diverges.

We say a pLTS is *convergent* if $\overline{s} \Longrightarrow \varepsilon$ for no state $s \in S$.

Lemma 6.17 Let Δ be a subdistribution in a *finite-state*, *convergent* and *deterministic* pLTS. If $\Delta \Longrightarrow \Delta'$ then $|\Delta| = |\Delta'|$.

Proof: Since the pLTS is convergent, then $\overline{s} \Longrightarrow \varepsilon$ for no state $s \in S$. In other words, each τ sequence from a state s is finite and ends with a distribution which cannot enable a τ transition. In a deterministic pLTS, each state has at most one outgoing transition. So from each s there is a unique τ sequence with length $n_s \geq 0$.

$$\overline{s} \xrightarrow{\tau} \Delta_1 \xrightarrow{\tau} \Delta_2 \xrightarrow{\tau} \cdots \xrightarrow{\tau} \Delta_{n_s} \not\xrightarrow{\tau}$$

Let p_s be $\Delta_{n_s}(s')$ where s' is any state in the support of Δ_{n_s}. We set

$$\begin{array}{rcl} n & = & \max\{n_s \mid s \in S\} \\ p & = & \min\{p_s \mid s \in S\} \end{array}$$

where n and p are well defined as S is a finite set since we are considering a finite-state pLTS. Now let $\Delta \Longrightarrow \Delta'$ be any weak derivation constructed by a collection of $\Delta_k^{\rightarrow}, \Delta_k^{\times}$ such that

$$\begin{array}{rcl} \Delta & = & \Delta_0^{\rightarrow} + \Delta_0^{\times} \\ \Delta_0^{\rightarrow} & \xrightarrow{\tau} & \Delta_1^{\rightarrow} + \Delta_1^{\times} \\ & \vdots & \\ \Delta_k^{\rightarrow} & \xrightarrow{\tau} & \Delta_{k+1}^{\rightarrow} + \Delta_{k+1}^{\times} \\ & \vdots & \end{array}$$

with $\Delta' = \sum_{k=0}^{\infty} \Delta_k^{\times}$. From each $\Delta_{kn+i}^{\rightarrow}$ with $k, i \in \mathbb{N}$, the block of n steps of τ transition leads to $\Delta_{(k+1)n+i}^{\rightarrow}$ such that $|\Delta_{(k+1)n+i}^{\rightarrow}| \leq |\Delta_{kn+i}^{\rightarrow}|(1-p)$. It follows that

$$\begin{array}{rcl} \sum_{j=0}^{\infty} |\Delta_j^{\rightarrow}| & = & \sum_{i=0}^{n-1} \sum_{k=0}^{\infty} |\Delta_{kn+i}^{\rightarrow}| \\ & \leq & \sum_{i=0}^{n-1} \sum_{k=0}^{\infty} |\Delta_i^{\rightarrow}|(1-p)^k \\ & = & \sum_{i=0}^{n-1} |\Delta_i^{\rightarrow}| \frac{1}{p} \\ & \leq & |\Delta_0^{\rightarrow}| \frac{n}{p} \end{array}$$

Therefore, we have that $\lim_{k\to\infty} \Delta_k^{\rightarrow} = 0$, which in turn means that $|\Delta'| = |\Delta|$. □

Corollary 6.18 [Zero-one law, deterministic case] If for some static derivative policy dpp over a finite-state pLTS there is for some s a derivation $\overline{s} \Longrightarrow_{\mathsf{dpp}} \Delta'$ with $|\Delta'| < 1$ then in fact for some (possibly different) state s_ε we have $s_\varepsilon \Longrightarrow_{\mathsf{dpp}} \varepsilon$.

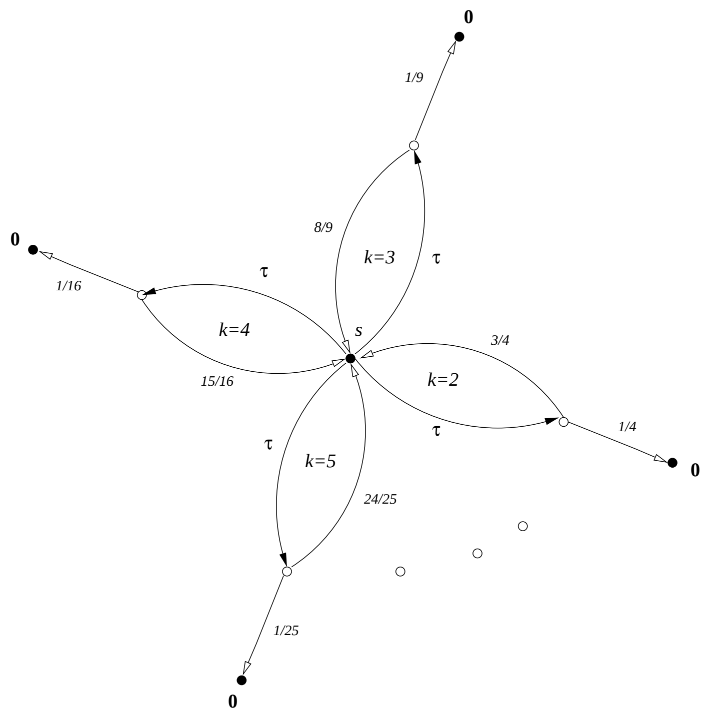
Proof: Suppose that for no state s do we have $\overline{s} \Longrightarrow_{\mathsf{dpp}} \varepsilon$. Then the pLTS induced by dpp is convergent. Since it is obviously finite-state and deterministic, we apply Lemma 6.17 and obtain $|\Delta'| = |\overline{s}| = 1$, contradicting the assumption that $|\Delta'| < 1$. Therefore, there must exist some state s_ε which wholly diverges. □

Although it is possible to have processes that diverge with some probability strictly between zero and one, in a finitary system it is possible to "distill" that divergence in the sense that in many cases we can limit our analyses to processes that either wholly diverge (can do so with probability one) or wholly converge (can diverge only with probability zero). This property is based on the zero-one law for finite-state probabilistic systems, and in this section we present the aspects of it that we need here.

Lemma 6.19 [Distillation of divergence, deterministic case]

If for some state s and static derivative policy dpp over a finite-state pLTS there is a derivation $\overline{s} \Longrightarrow_{\mathsf{dpp}} \Delta'$ then there is a probability p and full distributions $\Delta_1', \Delta_\varepsilon'$ such that $\overline{s} \Longrightarrow (\Delta_1' \,{}_p\!\oplus \Delta_\varepsilon')$ and $\Delta' = p \cdot \Delta_1'$ and $\Delta_\varepsilon' \Longrightarrow \varepsilon$.

Proof: (*Schema*) We modify dpp so as to obtain a static policy dpp' by setting $\mathsf{dpp}'(t) = \mathsf{dpp}(t)$ except when $\overline{t} \Longrightarrow_{\mathsf{dpp}} \varepsilon$, in which case we set $\mathsf{dpp}'(t) \uparrow$. The new policy determines a unique weak derivation $\Delta \Longrightarrow_{\mathsf{dpp}'} \Delta''$ for some subdistribution Δ'', and induces a sub-pLTS from the pLTS induced by dpp. Note that the sub-pLTS is deterministic and convergent. By Lemma 6.17, we know that $|\Delta''| = |\overline{s}| = 1$. We split Δ'' up into $\Delta_1'' + \Delta_\varepsilon''$ so that each state in



There are two states s and **0**. To diverge from s with probability $1-1/k$, start at "petal" k and take successive τ-loops anti-clockwise from there.
Yet, although divergence with arbitrarily high probability is present, complete probability-1 divergence is nowhere possible. Either infinite states or infinite branching is necessary for this anomaly.

Figure 7: Infinitely branching flower.

$\lceil\Delta''_\varepsilon\rceil$ is wholly divergent under policy dpp and Δ''_1 is supported by all other states. From Δ''_ε the policy dpp determines the weak derivation $\Delta''_\varepsilon \Longrightarrow_{\mathsf{dpp}} \varepsilon$. Combining the two weak derivations we have $\overline{s} \Longrightarrow_{\mathsf{dpp}'} \Delta''_1 + \Delta''_\varepsilon \Longrightarrow_{\mathsf{dpp}} \Delta''_1$. As we only divide the original weak SDP-derivation into two stages, and do not change the τ transition from each state, the final subdistribution will not change, thus $\Delta''_1 = \Delta'$. Finally we determine p, Δ'_1 and Δ'_ε by letting $p = |\Delta'|$, $\Delta'_1 = \frac{1}{p}\Delta'$ and $\Delta'_\varepsilon = \frac{1}{1-p}\Delta''_\varepsilon$. □

Theorem 6.20 [Distillation of divergence, general case] For any s, Δ' in a finitary pLTS with $\overline{s} \Longrightarrow \Delta'$ there is a probability p and full distributions $\Delta'_1, \Delta'_\varepsilon$ such that $\overline{s} \Longrightarrow (\Delta'_1 \,{}_p\!\oplus \Delta'_\varepsilon)$ and $\Delta' = p \cdot \Delta'_1$ and $\Delta'_\varepsilon \Longrightarrow \varepsilon$.

Proof: Let $\{\mathsf{dpp}_i \mid i \in I\}$ (I is a finite index set) be all the static derivative policies in the finitary pLTS. Each policy determines a weak derivation $\overline{s} \Longrightarrow_{\mathsf{dpp}_i} \Delta'_i$. From Theorem 6.8 we know that if $\overline{s} \Longrightarrow \Delta'$ then $\Delta' = \sum_{i\in I} p_i \Delta'_i$ for some p_i with $\sum_{i \in I} p_i = 1$. By Lemma 6.19, for each $i \in I$, there is a probability q_i and full distributions $\Delta'_{i,1}$, $\Delta'_{i,\varepsilon}$ such that $\overline{s} \Longrightarrow (\Delta'_{i,1} \,{}_{q_i}\!\oplus \Delta'_{i,\varepsilon})$, $\Delta'_i = q_i \cdot \Delta'_{i,1}$, and $\Delta'_{i,\varepsilon} \Longrightarrow \varepsilon$. Finally we determine p, Δ'_1 and Δ'_ε by letting $p = |\sum_{i\in I} p_i q_i \cdot \Delta'_{i,1}|$, $\Delta'_1 = \frac{1}{p}\Delta'$, and $\Delta'_\varepsilon = \frac{1}{1-p}\sum_{i \in I} p_i(1-q_i)\Delta'_{i,\varepsilon}$. They satisfy our requirements just by noting that $\overline{s} \Longrightarrow \sum_{i\in I} p_i(\Delta'_{i,1} \,{}_{q_i}\!\oplus \Delta'_{i,\varepsilon}) = \Delta'_1 \,{}_p\!\oplus \Delta'_\varepsilon$ □

The requirement on the pLTS to be finitary is essential for this distillation of divergence, as we explain in the following examples.

Example 6.21 [Revisiting Example 3.17] The pLTS in Example 3.17 is an infinite state system over states s_k for all $k \geq 2$, where the probability of convergence is $1/k$ from any state s_k, thus a situation where distillation of divergence fails because all the states partially diverge, yet there is no single state which wholly diverges. □

Example 6.22 Consider the finite state but infinitely branching pLTS described in Figure 7; this consists of two states s and $\mathbf{0}$ together with a k-indexed set of transitions

$$s \xrightarrow{\tau}_k (\llbracket \mathbf{0} \rrbracket_{1/k^2}\oplus \overline{s}) \quad \text{for } k \geq 2, \tag{19}$$

This pLTS is obtained from the infinite state pLTS described in Example 3.17 by identifying all of the states s_i and replacing the state a with $\mathbf{0}$.

As we have seen, by taking transitions $s \xrightarrow{\tau}_k \cdot \xrightarrow{\tau}_{k+1} \cdot \xrightarrow{\tau}_{k+2} \cdots$ we have $\overline{s} \Longrightarrow \frac{1}{k} \cdot \overline{\mathbf{0}}$ for any $k \geq 2$; but crucially $\overline{s} \not\Longrightarrow \varepsilon$. Since trivially $\overline{\mathbf{0}} \not\Longrightarrow \varepsilon$ there is no full distribution Δ such that $\Delta \Longrightarrow \varepsilon$.

Now to contradict the distillation of divergence for this pLTS note that $\overline{s} \Longrightarrow \frac{1}{2} \cdot \overline{\mathbf{0}}$, but this derivation cannot be factored in the required manner to $\overline{s} \Longrightarrow (\Delta'_1 {}_p\oplus \Delta'_\varepsilon)$, because no possible full distribution Δ'_ε can exist satisfying $\Delta'_\varepsilon \Longrightarrow \varepsilon$. □

Corollary 6.18 and Lemma 6.19 are not affected by infinite branching, because they are restricted to the deterministic case (i.e. the case of no branching at all). What fails is the combination of a number of deterministic distillations to make a non-deterministic one, in Theorem 6.20: it depends on Theorem 6.8, which in turn requires finite branching.

Corollary 6.23 [Zero-one law, general case] If in a finitary pLTS we have Δ, Δ' with $\Delta \Longrightarrow \Delta'$ and $|\Delta|>|\Delta'|$ then there is some state s' reachable with non-zero probability from Δ such that $\overline{s'} \Longrightarrow \varepsilon$. That is, the pLTS based on Δ must have a wholly diverging state somewhere.

Proof: Assume at first that $|\Delta|=1$; then the result is immediate from Theorem 6.20 since any $s' \in \lceil \Delta'_\varepsilon \rceil$ will do. The general result is obtained by dividing the given derivation by $|\Delta|$. □

7 The failure simulation preorder

This section is divided in four: the first subsection presents the definition of the *failure simulation preorder* in an arbitrary pLTS, together with some explanatory examples. It gives two equivalent characterisations of this preorder: a co-inductive one as a largest relation between subdistributions satisfying certain transfer properties, and one that is obtained through lifting and an additional closure property from a relation between states and subdistributions that we call *failure similarity*. It also investigates some elementary properties of the failure simulation preorder and of failure similarity. In the second subsection we restrict attention to finitary processes, and on this realm characterise the failure simulation preorder in terms of *simple failure similarity*. All further results on the failure simulation preorder, in particular precongruence for the operators of pCSP and soundness and completeness with respect to the must testing preorder, are in terms of this characterisation, and hence pertain to finitary processes only. The third subsection establishes monotonicity of the operators of pCSP with respect to the failure simulation preorder — in other words: shows that the failure simulation preorder is a precongruence with respect to these operators — and the last subsection is devoted to showing soundness with respect to must testing. Completeness is the subject of Section 8.

7.1 Two equivalent definitions and their rationale

We start with defining the weak action relations $\stackrel{\alpha}{\Longrightarrow}$ for $\alpha \in \mathsf{Act}_\tau$ and the refusal relations $\stackrel{A}{\not\longrightarrow}$ for $A \subseteq \mathsf{Act}$ that are the key ingredients in the definition of the failure simulation preorder [7, 2].

Definition 7.1 Let Δ and its variants be subdistributions in a pLTS $\langle S, \mathsf{Act}_\tau, \rightarrow \rangle$.

- For $a \in \mathsf{Act}$ write $\Delta \stackrel{a}{\Longrightarrow} \Delta'$ whenever $\Delta \Longrightarrow \Delta^{\text{pre}} \xrightarrow{a} \Delta^{\text{post}} \Longrightarrow \Delta'$. Extend this to Act_τ by allowing as a special case that $\stackrel{\tau}{\Longrightarrow}$ is simply $\Longrightarrow$, i.e. including identity (rather than requiring at least one $\xrightarrow{\tau}$).
- For $A \subseteq \mathsf{Act}$ and $s \in S$ write $s \stackrel{A}{\not\longrightarrow}$ if $s \stackrel{\alpha}{\not\longrightarrow}$ for every $\alpha \in A \cup \{\tau\}$; write $\Delta \stackrel{A}{\not\longrightarrow}$ if $s \stackrel{A}{\not\longrightarrow}$ for every $s \in \lceil \Delta \rceil$.
- More generally write $\Delta \Longrightarrow\stackrel{A}{\not\longrightarrow}$ if $\Delta \Longrightarrow \Delta^{\text{pre}}$ for some Δ^{pre} such that $\Delta^{\text{pre}} \stackrel{A}{\not\longrightarrow}$.

For example, referring to Example 3.16 we have $\llbracket Q_1 \rrbracket \stackrel{a}{\Longrightarrow} \llbracket \mathbf{0} \rrbracket$, while in Example 3.17 we have $\llbracket s_2 \rrbracket \stackrel{a}{\Longrightarrow} \frac{1}{2}\llbracket \mathbf{0} \rrbracket$ as well as $\llbracket s_2 \rrbracket \Longrightarrow\stackrel{B}{\not\longrightarrow}$ for any set B not containing a, because $s_2 \Longrightarrow \frac{1}{2}\llbracket a \rrbracket$.

Definition 7.2 (Failure Simulation Preorder) Define $\sqsupseteq_{FS}$ to be the largest relation in $\mathscr{D}(S) \times \mathscr{D}(S)$ such that if $\Delta \sqsupseteq_{FS} \Theta$ then

1. whenever $\Delta \stackrel{\alpha}{\Longrightarrow} (\sum_i p_i \Delta'_i)$, for $\alpha \in \mathsf{Act}_\tau$ and certain p_i with $(\sum_i p_i) \leq 1$, then there are $\Theta'_i \in \mathscr{D}(S)$ with $\Theta \stackrel{\alpha}{\Longrightarrow} (\sum_i p_i \Theta'_i)$ and $\Delta'_i \sqsupseteq_{FS} \Theta'_i$ for each i, and
2. whenever $\Delta \Longrightarrow \stackrel{A}{\not\longrightarrow}$ then also $\Theta \Longrightarrow \stackrel{A}{\not\longrightarrow}$.

Naturally $\Theta \sqsubseteq_{FS} \Delta$ just means $\Delta \sqsupseteq_{FS} \Theta$. We have chosen the orientation of the preorder symbol to match that of must testing, which goes back to the work of De Nicola & Hennessy [6]. This orientation also matches the one used in CSP [10] and related work, were we have SPECIFICATION $\sqsubseteq$ IMPLEMENTATION. At the same time, we like to stick to the convention popular in the CCS community of writing the simulated process to the left of the preorder symbol and the simulating process (that mimics moves of the simulated one) on the right. This helps when comparing with may testing and the simulation preorder in Section 9. We achieve this by writing IMPLEMENTATION $\sqsupseteq_{FS}$ SPECIFICATION.

In the first case of the above definition the summation is allowed to be empty, which has the following useful consequence.

Lemma 7.3 If Δ diverges and $\Delta \sqsupseteq_{FS} \Theta$, then also Θ diverges.

Proof: Divergence of Δ means that $\Delta \Longrightarrow \varepsilon$, whence with $\Delta \sqsupseteq_{FS} \Theta$ we can take the empty summation in Definition 7.2 to conclude that also $\Theta \Longrightarrow \varepsilon$. □

Although the regularity of Definition 7.2 is appealing — for example it is trivial to see that $\sqsubseteq_{FS}$ is reflexive and transitive, as it should be — in practice, for specific processes, it is easier to work with a characterisation of the failure simulation preorder in terms of a relation between between *states* and distributions.

Definition 7.4 (Failure Similarity) Define $\lhd_{FS}$ to be the largest relation in $S \times \mathscr{D}(S)$ such that if $s \lhd_{FS} \Theta$ then

1. whenever $\overline{s} \stackrel{\alpha}{\Longrightarrow} \Delta'$, for $\alpha \in \mathsf{Act}_\tau$, then there is a $\Theta' \in \mathscr{D}(S)$ with $\Theta \stackrel{\alpha}{\Longrightarrow} \Theta'$ and $\Delta' \; \overline{\lhd_{FS}} \; \Theta'$, and
2. whenever $\overline{s} \Longrightarrow \stackrel{A}{\not\longrightarrow}$ then $\Theta \Longrightarrow \stackrel{A}{\not\longrightarrow}$.

Any relation $\mathscr{R} \subseteq S \times \mathscr{D}(S)$ that satisfies the two clauses above is called a *failure simulation*.

Obviously, for any failure simulation $\mathscr{R}$ we have $\mathscr{R} \subseteq \lhd_{FS}$. The following two lemmas show that the lifted failure similarity relation $\overline{\lhd_{FS}} \subseteq \mathscr{D}(S) \times \mathscr{D}(S)$ has simulating properties analogous to 1 and 2 above.

Lemma 7.5 Suppose $\Delta \; \overline{\lhd_{FS}} \; \Theta$ and $\Delta \stackrel{\alpha}{\Longrightarrow} \Delta'$ for $\alpha \in \mathsf{Act}_\tau$. Then $\Theta \stackrel{\alpha}{\Longrightarrow} \Theta'$ for some Θ' such that $\Delta' \; \overline{\lhd_{FS}} \; \Theta'$.

Proof: $\Delta \; \overline{\lhd_{FS}} \; \Theta$ implies by Lemma 3.5 that $\qquad \Delta = \sum_{i \in I} p_i \cdot \overline{s_i}, \qquad s_i \lhd_{FS} \Theta_i, \qquad \Theta = \sum_{i \in I} p_i \cdot \Theta_i.$

By Corollary 6.16 and Proposition 3.10 we know from $\Delta \stackrel{\alpha}{\Longrightarrow} \Delta'$ that $\overline{s_i} \stackrel{\alpha}{\Longrightarrow} \Delta'_i$ for $\Delta'_i \in \mathscr{D}(S)$ such that $\Delta' = \sum_{i \in I} p_i \cdot \Delta'_i$. For each $i \in I$ we infer from $s_i \lhd_{FS} \Theta_i$ and $\overline{s_i} \stackrel{\alpha}{\Longrightarrow} \Delta'_i$ that there is a $\Theta'_i \in \mathscr{D}(S)$ with $\Theta_i \stackrel{\alpha}{\Longrightarrow} \Theta'_i$ and $\Delta'_i \; \overline{\lhd_{FS}} \; \Theta'$. Let $\Theta' := \sum_{i \in I} p_i \cdot \Theta'_i$. Then Definition 3.2(2) and Theorem 3.20(i) yield $\Delta' \; \overline{\lhd_{FS}} \; \Theta'$ and $\Theta \stackrel{\alpha}{\Longrightarrow} \Theta'$. □

Lemma 7.6 Suppose $\Delta \; \overline{\lhd_{FS}} \; \Theta$ and $\Delta \Longrightarrow \stackrel{A}{\not\longrightarrow}$. Then $\Theta \Longrightarrow \stackrel{A}{\not\longrightarrow}$.

Proof: Suppose $\Delta \; \overline{\lhd_{FS}} \; \Theta$ and $\Delta \Longrightarrow \Delta' \stackrel{A}{\not\longrightarrow}$. By Lemma 7.5 there exists some Θ' such that $\Theta \Longrightarrow \Theta'$ and $\Delta' \; \overline{\lhd_{FS}} \; \Theta'$. From Lemma 3.5 we know that $\Delta' = \sum_{i \in I} p_i \cdot \overline{s_i}, \qquad s_i \lhd_{FS} \Theta_i, \qquad \Theta' = \sum_{i \in I} p_i \cdot \Theta_i,$ with $s_i \in \lceil \Delta' \rceil$ for all $i \in I$. Since $\Delta' \stackrel{A}{\not\longrightarrow}$, we have that $s_i \stackrel{A}{\not\longrightarrow}$ for all $i \in I$. It follows from $s_i \lhd_{FS} \Theta_i$ that $\Theta_i \Longrightarrow \Theta'_i \stackrel{A}{\not\longrightarrow}$. By Theorem 3.20(i) we obtain that $\sum_{i \in I} p_i \cdot \Theta_i \Longrightarrow \sum_{i \in I} p_i \cdot \Theta'_i \stackrel{A}{\not\longrightarrow}$. By the transitivity of $\Longrightarrow$ we have that $\Theta \Longrightarrow \stackrel{A}{\not\longrightarrow}$. □

The next result shows how the failure simulation preorder can alternatively be defined in terms of failure similarity.

Proposition 7.7 For $\Delta, \Theta \in \mathscr{D}(S)$ we have $\Delta \sqsupseteq_{FS} \Theta$ just when there is a Θ^{match} with $\Theta \Longrightarrow \Theta^{\text{match}}$ and $\Delta \; \overline{\lhd_{FS}} \; \Theta^{\text{match}}$.

Proof: Let $\lhd'_{FS} \subseteq S \times \mathscr{D}(S)$ be the relation given by $s \lhd'_{FS} \Theta$ iff $\overline{s} \sqsupseteq_{FS} \Theta$. Then $\lhd'_{FS}$ is a failure simulation; hence $\lhd'_{FS} \subseteq \lhd_{FS}$. Now suppose $\Delta \sqsupseteq_{FS} \Theta$. Let $\Delta := \sum_i p_i \cdot \overline{s_i}$. Then there are Θ_i with $\Theta \Longrightarrow \sum_i p_i \cdot \Theta_i$ and $\overline{s_i} \sqsupseteq_{FS} \Theta_i$ for each i, whence $s_i \lhd'_{FS} \Theta_i$, and thus $s_i \lhd_{FS} \Theta_i$. Take $\Theta^{\text{match}} := \sum_i p_i \cdot \Theta_i$. Definition 3.2 yields $\Delta \overline{\lhd_{FS}} \Theta^{\text{match}}$.

For the other direction it suffices to show that $\overline{\lhd_{FS}}; \Longrightarrow^{-1}$ satisfies the two clauses of Definition 7.2, yielding $\overline{\lhd_{FS}}; \Longrightarrow^{-1} \subseteq \sqsupseteq_{FS}$. So suppose, for given $\Delta, \Theta \in \mathscr{D}(S)$, there is a Θ^{match} with $\Theta \Longrightarrow \Theta^{\text{match}}$ and $\Delta \overline{\lhd_{FS}} \Theta^{\text{match}}$.

Suppose $\Delta \overset{\alpha}{\Longrightarrow} \sum_{i\in I} p_i \cdot \Delta'_i$ for some $\alpha \in \mathsf{Act}_\tau$. By Lemma 7.5 there is some Θ' such that $\Theta^{\text{match}} \overset{\alpha}{\Longrightarrow} \Theta'$ and $(\sum_{i\in I} p_i \cdot \Delta'_i) \overline{\lhd_{FS}} \Theta'$. From Proposition 3.10 we know that $\Theta' = \sum_{i\in I} p_i \cdot \Theta'_i$ for subdistributions Θ'_i such that $\Delta'_i \overline{\lhd_{FS}} \Theta'_i$ for $i \in I$. Thus $\Theta \overset{\alpha}{\Longrightarrow} \sum_i p_i \cdot \Theta'_i$ by the transitivity of $\Longrightarrow$ (Theorem 3.22) and $\Delta'_i (\overline{\lhd_{FS}}; \Longrightarrow^{-1}) \Theta'_i$ for each $i \in I$ by the reflexivity of $\Longrightarrow$.

Suppose $\Delta \Longrightarrow\overset{A}{\not\longrightarrow}$. By Lemma 7.6 we have $\Theta^{\text{match}} \Longrightarrow\overset{A}{\not\longrightarrow}$. It follows that $\Theta \Longrightarrow\overset{A}{\not\longrightarrow}$ by the transitivity of $\Longrightarrow$. □

Note the appearance of the "anterior step" $\Theta \Longrightarrow \Theta^{\text{match}}$ in Proposition 7.7 immediately above; the following example shows it necessary in the sense that defining $\sqsupseteq_{FS}$ simply to be $\overline{\lhd_{FS}}$ (i.e. without anterior step) would not have been suitable.

Example 7.8 Compare the two processes $P := a \; {}_{\frac{1}{2}}\!\oplus \; b$ and $Q := \tau.P$. They are testing equivalent, and so for $\overline{\lhd_{FS}}$ to be complete we would have to have $[\![P]\!] \overline{\lhd_{FS}} [\![Q]\!]$. But we do not, for by Proposition 3.10 that would require $[\![a]\!] \overline{\lhd_{FS}} [\![Q]\!]$, which must fail since the former's move $\xrightarrow{a} [\![\,\mathbf{0}\,]\!]$ cannot be matched by the latter.

We do however have $P \sqsupseteq_{FS} Q$ because of the anterior step $Q \Longrightarrow P$ and of course $[\![P]\!] \overline{\lhd_{FS}} [\![P]\!]$. □

Remark 7.9 For $s \in S$ and $\Theta \in \mathscr{D}(S)$ we have $s \lhd_{FS} \Theta$ iff $\overline{s} \sqsupseteq_{FS} \Theta$; here no anterior step is needed. One direction of this statement has been obtained in the beginning of the proof of Proposition 7.7; for the other note that $s \lhd_{FS} \Theta$ implies $\overline{s} \overline{\lhd_{FS}} \Theta$ by Definition 3.2(1) which implies $\overline{s} \sqsupseteq_{FS} \Theta$ by Proposition 7.7 and the reflexivity of $\Longrightarrow$.

Example 7.8 shows that $\sqsupseteq_{FS}$ cannot be obtained as the lifting of any relation: it lacks the decomposition property of Proposition 3.10. Nevertheless, $\sqsupseteq_{FS}$ enjoys the property of linearity, as occurs in Definition 3.2:

Lemma 7.10 If $\Delta_i \sqsupseteq_{FS} \Theta_i$ for $i \in I$ then $\sum_{i\in I} p_i \cdot \Delta_i \sqsupseteq_{FS} \sum_{i\in I} p_i \cdot \Theta_i$ for any $p_i \in [0,1]$ $(i \in I)$ with $\sum_{i\in I} p_i \leq 1$.

Proof: This follows immediately from the linearity of $\overline{\lhd_{FS}}$ and $\Longrightarrow$ (cf. Theorem 3.20(i)), using Proposition 7.7. □

Example 7.11 (Divergence) From Example 3.15 we know that $[\![\mathsf{rec}\, x.\, x]\!] \Longrightarrow \varepsilon$. This, together with (1) in Section 3.1, and the fact that $\varepsilon \overset{A}{\not\longrightarrow}$ for any set of actions A, ensures that $s \lhd_{FS} [\![\mathsf{rec}\, x.\, x]\!]$ for any s, hence $\Theta \overline{\lhd_{FS}} [\![\mathsf{rec}\, x.\, x]\!]$ for any Θ, and thus that $\Theta \sqsupseteq_{FS} [\![\mathsf{rec}\, x.\, x]\!]$. Indeed similar reasoning applies to any Δ with $\Delta = \Delta_0 \xrightarrow{\tau} \Delta_1 \xrightarrow{\tau} \cdots \xrightarrow{\tau} \cdots$ because — as explained right before Example 3.15 — this also ensures that $\Delta \Longrightarrow \varepsilon$. In particular, we have $\varepsilon \Longrightarrow \varepsilon$ and hence $[\![\mathsf{rec}\, x.\, x]\!] \simeq_{FS} \varepsilon$.

Yet $[\![\mathsf{rec}\, x.\, x]\!] \not\sqsupseteq_{FS} \mathbf{0}$, because the move $[\![\mathsf{rec}\, x.\, x]\!] \Longrightarrow \varepsilon$ cannot be matched by a corresponding move from $[\![\,\mathbf{0}\,]\!]$ — see Lemma 7.3. □

Example 7.11 shows again that the anterior move in Proposition 7.7 is necessary: although $\varepsilon \sqsupseteq_{FS} [\![\mathsf{rec}\, x.\, x]\!]$ we do not have $\varepsilon \overline{\lhd_{FS}} [\![\mathsf{rec}\, x.\, x]\!]$, since by Lemma 3.6 any Θ with $\varepsilon \overline{\lhd_{FS}} \Theta$ must have $|\Theta| = 0$.

Example 7.12 Referring to the process Q_1 of Example 3.16, with Proposition 7.7 we easily see that $a \sqsupseteq_{FS} Q_1$ because we have $a \lhd_{FS} [\![Q_1]\!]$. Note that the move $[\![Q_1]\!] \Longrightarrow [\![a]\!]$ is crucial, since it enables us to match the move $[\![a]\!] \xrightarrow{a} [\![\,\mathbf{0}\,]\!]$ with $[\![Q_1]\!] \Longrightarrow [\![a]\!] \xrightarrow{a} [\![\,\mathbf{0}\,]\!]$. It also enables us to match refusals: if $[\![a]\!] \overset{B}{\not\longrightarrow}$ then B can not contain the action a, and therefore also $[\![Q_1]\!] \Longrightarrow\overset{B}{\not\longrightarrow}$.

The converse, that $a \sqsubseteq_{FS} Q_1$, is also true because it is straightforward to verify that the relation

$$\{(Q_1, [\![a]\!]), (\tau.Q_1, [\![a]\!]), (a, [\![a]\!]), (\mathbf{0}, [\![\,\mathbf{0}\,]\!])\}$$

is a failure simulation and thus is a subset of $\lhd_{FS}$. We therefore have $Q_1 \simeq_{FS} a$. □

Example 7.13 Let P be the process $a\,{}_{\frac{1}{2}}\!\oplus\, \mathsf{rec}\,x.\,x$ and consider the state s_2 introduced in Example 3.17. First note that $[\![P]\!]\ \overline{\lhd_{FS}}\ \frac{1}{2}\cdot[\![a]\!]$, since $\mathsf{rec}\,x.\,x \lhd_{FS} \varepsilon$. Then because $s_2 \Longrightarrow \frac{1}{2}\cdot[\![a]\!]$ we have $[\![P]\!] \sqsupseteq_{FS} s_2$. The converse, that $s_2 \sqsupseteq_{FS} [\![P]\!]$ holds, is true because $s_2 \lhd_{FS} [\![P]\!]$ follows from the fact that the relation

$$\{(s_k, [\![a]\!]\,{}_{1/k}\!\oplus [\![\mathsf{rec}\,x.\,x]\!]) \mid k \geq 2\} \cup \{(a, [\![a]\!]), (\mathbf{0}, [\![\,\mathbf{0}\,]\!])\}$$

is a failure simulation that contains the pair $(s_2, [\![P]\!])$. □

Our final examples pursue the consequences of the fact that the empty distribution ε is behaviourally indistinguishable from divergent processes like $[\![\mathsf{rec}\,x.\,x]\!]$.

Example 7.14 (Subdistributions formally unnecessary) For any subdistribution Δ, let Δ^e denote the (full) distribution defined by

$$\Delta^e \quad := \quad \Delta \ + \ (1-|\Delta|)\cdot\overline{[\![\mathsf{rec}\,x.\,x]\!]}\ .$$

Intuitively it is obtained from Δ by padding the missing support with the divergent state $[\![\mathsf{rec}\,x.\,x]\!]$.

Then $\Delta \simeq_{FS} \Delta^e$. This follows because $\Delta^e \Longrightarrow \Delta$, which is sufficient to establish $\Delta \sqsupseteq_{FS} \Delta^e$; but also $\Delta^e\ \overline{\lhd_{FS}}\ \Delta$ because $[\![\mathsf{rec}\,x.\,x]\!] \lhd_{FS} \varepsilon$, and that implies the converse $\Delta^e \sqsupseteq_{FS} \Delta$. The equivalence shows that formally we have no need for subdistributions, and that our technical development could be carried out using (full) distributions only. □

But abandoning subdistributions comes at a cost: the definition of weak transition, Definition 3.13, would be much more complex if expressed with full distributions, as would syntactic manipulations such as those used in the proof of Theorem 3.22.

More significant, however, is that diverging processes have a special character in failure simulation semantics. Placing them at the bottom of the $\sqsubseteq_{FS}$ preorder — as we do – requires that they failure-simulate every processes, thus allowing all visible actions and all refusals and so behaving in a sense "chaotically"; yet applying the operational semantics of Figure 2 to $\mathsf{rec}\,x.\,x$ literally would suggest exactly the opposite, since $\mathsf{rec}\,x.\,x$ allows no visible actions (all its derivatives enable only τ) and no refusals (all its derivatives have τ enabled). The case analyses that discrepancy would require are entirely escaped by allowing subdistributions, as the chaotic behaviour of the diverging ε follows naturally from the definitions, as we saw in Example 7.11.

We conclude with an example involving divergence and subdistributions.

Example 7.15 For $0 \leq c \leq 1$ let P_c be the process $\mathbf{0}\,{}_c\!\oplus\ \mathsf{rec}\,x.\,x$. We show that $[\![P_c]\!] \sqsubseteq_{FS} [\![P_{c'}]\!]$ just when $c \leq c'$. (Refusals can be ignored, since P_c refuses every set of actions, for all c.)

Suppose first that $c \leq c'$, and split the two processes as follows:

$$\begin{array}{rl} [\![P_c]\!] = & c\cdot[\![\,\mathbf{0}\,]\!] + (c'-c)\cdot[\![\mathsf{rec}\,x.\,x]\!] + (1-c')\cdot[\![\mathsf{rec}\,x.\,x]\!] \\ [\![P_{c'}]\!] = & c\cdot[\![\,\mathbf{0}\,]\!] + (c'-c)\cdot[\![\,\mathbf{0}\,]\!] \qquad\quad + (1-c')\cdot[\![\mathsf{rec}\,x.\,x]\!] \end{array} \quad .$$

Because $\mathbf{0} \lhd_{FS} [\![\mathsf{rec}\,x.\,x]\!]$ (the middle terms), we have immediately $[\![P_{c'}]\!]\ \overline{\lhd_{FS}}\ [\![P_c]\!]$, whence $[\![P_c]\!] \sqsubseteq_{FS} [\![P_{c'}]\!]$.

For the other direction, note that $[\![P_{c'}]\!] \Longrightarrow c'\cdot[\![\,\mathbf{0}\,]\!]$. If $[\![P_c]\!] \sqsubseteq_{FS} [\![P_{c'}]\!]$ then from Definition 7.2 we would have to have $[\![P_c]\!] \Longrightarrow c'\cdot\Theta'$ for some subdistribution Θ', a derivative of weight no more than c'. But the smallest weight P_c can reach via $\Longrightarrow$ is just c, so that we must have in fact $c \leq c'$. □

We end this subsection with two properties of failure similarity that will be useful later on.

Proposition 7.16 The relation $\lhd_{FS}$ is convex.

Proof: Suppose $s \lhd_{FS} \Theta_i$ and $p_i \in [0,1]$ for $i \in I$, with $\sum_{i\in I} p_i = 1$. We need to show that $s \lhd_{FS} \sum_{i\in I} p_i\cdot\Theta_i$.

If $\overline{s} \stackrel{\alpha}{\Longrightarrow} \Delta'$, then there exist Θ_i' for $i \in I$ such that $\Theta_i \stackrel{\alpha}{\Longrightarrow} \Theta_i'$ and $\Delta'\ \overline{\lhd_{FS}}\ \Theta_i'$. By Corollary 6.16 and Definition 3.2(2), we obtain that $\sum_{i\in I} p_i\cdot\Theta_i \stackrel{\alpha}{\Longrightarrow} \sum_{i\in I} p_i\cdot\Theta_i'$ and $\Delta'\ \overline{\lhd_{FS}}\ \sum_{i\in I} p_i\cdot\Theta_i'$.

If $\overline{s} \Longrightarrow\stackrel{A}{\not\longrightarrow}$ for some $A \subseteq \mathsf{Act}$, then $\Theta_i \Longrightarrow \Theta_i' \stackrel{A}{\not\longrightarrow}$ for all $i \in I$. By definition we have $\sum_{i\in I} p_i\cdot\Theta_i' \stackrel{A}{\not\longrightarrow}$. Theorem 3.20(i) yields $\sum_{i\in I} p_i\cdot\Theta_i \Longrightarrow \sum_{i\in I} p_i\cdot\Theta_i'$.

So we have checked that $s \lhd_{FS} \sum_{i\in I} p_i\cdot\Theta_i$. It follows that $\lhd_{FS}$ is convex. □

Proposition 7.17 The relation $\overline{\lhd_{FS}} \subseteq \mathscr{D}(S) \times \mathscr{D}(S)$ is reflexive and transitive.

Proof: Reflexivity is easy; it relies on the fact that $s \lhd_{FS} \overline{s}$ for every state s.

For transitivity, we first show that $\lhd_{FS};\overline{\lhd_{FS}}$ is a failure simulation. Suppose $s \lhd_{FS} \Theta \ \overline{\lhd_{FS}}\ \Phi$. If $\overline{s} \stackrel{\alpha}{\Longrightarrow} \Delta'$ then there is a Θ' such that $\Theta \stackrel{\alpha}{\Longrightarrow} \Theta'$ and $\Delta' \ \overline{\lhd_{FS}}\ \Theta'$. By Lemma 7.5, there exists a Φ' such that $\Phi \stackrel{\alpha}{\Longrightarrow} \Phi'$ and $\Theta' \ \overline{\lhd_{FS}}\ \Phi'$. Hence, $\Delta' \ \overline{\lhd_{FS}};\overline{\lhd_{FS}}\ \Phi'$. By Lemma 3.12 we know that

$$\overline{\lhd_{FS}};\overline{\lhd_{FS}} = \overline{\lhd_{FS};\overline{\lhd_{FS}}} \tag{20}$$

Therefore, we obtain $\Delta' \ \overline{\lhd_{FS};\overline{\lhd_{FS}}}\ \Phi'$.

If $s \Longrightarrow\stackrel{A}{\not\longrightarrow}$ for some $A \subseteq \mathsf{Act}$, then $\Theta \Longrightarrow\stackrel{A}{\not\longrightarrow}$ and hence $\Phi \Longrightarrow\stackrel{A}{\not\longrightarrow}$ by Lemma 7.6.

So we established that $\lhd_{FS};\overline{\lhd_{FS}} \subseteq \lhd_{FS}$. It now follows from Remark 3.3 and (20) that $\overline{\lhd_{FS}};\overline{\lhd_{FS}} \subseteq \overline{\lhd_{FS}}$. □

7.2 A simpler characterisation of failure similarity for finitary processes

Here we present a simpler characterisation of failure similarity, valid when considering finitary processes only. It is in terms of this characterisation that we will establish soundness and completeness of the failure simulation preorder with respect to the must testing preorder; consequently we have these results for finitary processes only.

Definition 7.18 (Simple Failure Similarity) Let $\lhd_{FS}^{s}$ be the largest relation in $S \times \mathscr{D}(S)$ such that if $s \lhd_{FS}^{s} \Theta$ then

1. whenever $\overline{s} \Longrightarrow \varepsilon$ then also $\Theta \Longrightarrow \varepsilon$, otherwise
2. whenever $s \stackrel{\alpha}{\longrightarrow} \Delta'$, for $\alpha \in \mathsf{Act}_\tau$, then there is a Θ' with $\Theta \stackrel{\alpha}{\Longrightarrow} \Theta'$ and $\Delta' \ \overline{\lhd_{FS}^{s}}\ \Theta'$, and
3. whenever $s \stackrel{A}{\not\longrightarrow}$ then $\Theta \Longrightarrow\stackrel{A}{\not\longrightarrow}$.

We first note that the relation $\lhd_{FS}^{s}$ is not interesting for infinitary processes since its lifted form $\overline{\lhd_{FS}^{s}}$ is not a transitive relation for those processes.

Example 7.19 Consider the process defined by the following two transitions: $t_0 \stackrel{\tau}{\longrightarrow} (\overline{\mathbf{0}}\ {}_{1/2}\!\oplus \overline{t_1})$ and $t_1 \stackrel{\tau}{\longrightarrow} \overline{t_1}$. We compare state t_0 with state s in Example 6.22 and have that $t_0 \lhd_{FS}^{s} \overline{s}$. The transition $t_0 \stackrel{\tau}{\longrightarrow} (\overline{\mathbf{0}}\ {}_{1/2}\!\oplus \overline{t_1})$ can be matched up by $s \Longrightarrow \frac{1}{2}\overline{\mathbf{0}}$ because $(\overline{\mathbf{0}}\ {}_{1/2}\!\oplus \overline{t_1})\ \overline{\lhd_{FS}^{s}}\ \frac{1}{2}\overline{\mathbf{0}}$ by noticing that $t_1 \lhd_{FS}^{s} \varepsilon$.

It also holds that $s \lhd_{FS}^{s} \overline{\mathbf{0}}$ because the relation $\{(s,\overline{\mathbf{0}}),(\mathbf{0},\overline{\mathbf{0}})\}$ is a simple failure simulation. The transition $s \stackrel{\tau}{\longrightarrow}_k (\overline{\mathbf{0}}\ {}_{\frac{1}{2}}\!\oplus \overline{s})$ for any $k \geq 2$ is matched up by $\mathbf{0} \Longrightarrow \overline{\mathbf{0}}$.

However, we do not have $t_0 \lhd_{FS}^{s} \overline{\mathbf{0}}$. The only candidate to simulate the transition $t_0 \stackrel{\tau}{\longrightarrow} (\overline{\mathbf{0}}\ {}_{1/2}\!\oplus \overline{t_1})$ is $\mathbf{0} \Longrightarrow \overline{\mathbf{0}}$, but $(\overline{\mathbf{0}}\ {}_{\frac{1}{2}}\!\oplus \overline{t_1})\ \overline{\not\lhd_{FS}^{s}}\ \overline{\mathbf{0}}$ because the divergent state t_1 cannot be simulated by $\mathbf{0}$.

Therefore, we have $\overline{t_0}\ \overline{\lhd_{FS}^{s}}\ \overline{s}\ \overline{\lhd_{FS}^{s}}\ \overline{\mathbf{0}}$ but $\overline{t_0}\ \overline{\not\lhd_{FS}^{s}}\ \overline{\mathbf{0}}$, thus transitivity of the relation $\overline{\lhd_{FS}^{s}}$ fails. Here the state s is not finitely branching. As a matter of fact, transitivity of $\overline{\lhd_{FS}^{s}}$ also fails for finitely branching but infinite state processes.

Consider an infinite state pLTS consisting of a collection of states s_k for $k \geq 2$ such that

$$s_k \stackrel{\tau}{\longrightarrow} \overline{\mathbf{0}}\ {}_{\frac{1}{k^2}}\!\oplus \overline{s_{k+1}}. \tag{21}$$

This pLTS is obtained from that in Example 3.17 by replacing a with $\mathbf{0}$. One can check that $\overline{t_0}\ \overline{\lhd_{FS}^{s}}\ \overline{s_2}\ \overline{\lhd_{FS}^{s}}\ \overline{\mathbf{0}}$ but we already know that $\overline{t_0}\ \overline{\not\lhd_{FS}^{s}}\ \overline{\mathbf{0}}$. Again, we loose the transitivity of $\overline{\lhd_{FS}^{s}}$. □

If we restrict our attention to finitary processes, then $\lhd_{FS}^{s}$ provides a simpler characterisation of failure similarity.

Theorem 7.20 (Equivalence of failure- and simple failure similarity) For finitary distributions $\Delta, \Theta \in \mathscr{D}(S)$ in a pLTS $\langle S, \mathsf{Act}_\tau, \rightarrow\rangle$ we have $\Delta \ \overline{\lhd_{FS}^{s}}\ \Theta$ iff $\Delta \ \overline{\lhd_{FS}}\ \Theta$.

Proof: Because $s \stackrel{\alpha}{\longrightarrow} \Delta'$ implies $\overline{s} \stackrel{\alpha}{\longrightarrow} \Delta'$ and $s \stackrel{A}{\not\longrightarrow}$ implies $\overline{s} \Longrightarrow\stackrel{A}{\not\longrightarrow}$ it is trivial that $\lhd_{FS}$ satisfies the conditions of Definition 7.18, so that $\lhd_{FS} \subseteq \lhd_{FS}^{s}$.

For the other direction we need to show that $\lhd_{FS}^{s}$ satisfies Clause 1 of Definition 7.4 with $\alpha = \tau$, that is

$$\text{if } s \lhd_{FS}^{s} \Theta \text{ and } \overline{s} \Longrightarrow \Delta' \text{ then there is some } \Theta' \in \mathscr{D}(S) \text{ with } \Theta \Longrightarrow \Theta' \text{ and } \Delta' \ \overline{\lhd_{FS}^{s}}\ \Theta'.$$

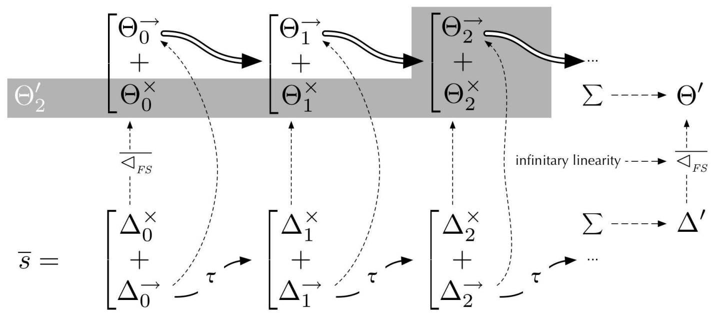


Figure 8: Illustration of Theorem 7.20

Once we have this, the relation $\lhd^s_{FS}$ clearly satisfies both clauses of Definition 7.4, so that we have $\lhd^s_{FS} \subseteq \lhd_{FS}$.

So suppose that $s \lhd^s_{FS} \Theta$ and that $\overline{s} \Longrightarrow \Delta'$ where — for the moment — we assume $|\Delta'| = 1$. Referring to Definition 3.13, there must be Δ_k, $\Delta^{\rightarrow}_k$ and $\Delta^{\times}_k$ for $k \geq 0$ such that $\overline{s} = \Delta_0$, $\Delta_k = \Delta^{\rightarrow}_k + \Delta^{\times}_k$, $\Delta^{\rightarrow}_k \xrightarrow{\tau} \Delta_{k+1}$ and $\Delta' = \sum_{k=1}^{\infty} \Delta^{\times}_k$. Since $\Delta^{\times}_0 + \Delta^{\rightarrow}_0 = \overline{s} \ \overline{\lhd^s_{FS}}\ \Theta$, using Proposition 3.10 we can define $\Theta =: \Theta^{\times}_0 + \Theta^{\rightarrow}_0$ so that $\Delta^{\times}_0 \ \overline{\lhd^s_{FS}}\ \Theta^{\times}_0$ and $\Delta^{\rightarrow}_0 \ \overline{\lhd^s_{FS}}\ \Theta^{\rightarrow}_0$. Since $\Delta^{\rightarrow}_0 \xrightarrow{\tau} \Delta_1$ and $\Delta^{\rightarrow}_0 \ \overline{\lhd^s_{FS}}\ \Theta^{\rightarrow}_0$ we have $\Theta^{\rightarrow}_0 \Longrightarrow \Theta_1$ with $\Delta_1 \ \overline{\lhd^s_{FS}}\ \Theta_1$.

Repeating the above procedure gives us inductively a series $\Theta_k, \Theta^{\rightarrow}_k, \Theta^{\times}_k$ of subdistributions, for $k \geq 0$, such that $\Theta_0 = \Theta$, $\Delta_k \ \overline{\lhd^s_{FS}}\ \Theta_k$, $\Theta_k = \Theta^{\rightarrow}_k + \Theta^{\times}_k$, $\Delta^{\times}_k \ \overline{\lhd^s_{FS}}\ \Theta^{\times}_k$, $\Delta^{\rightarrow}_k \ \overline{\lhd^s_{FS}}\ \Theta^{\rightarrow}_k$ and $\Theta^{\rightarrow}_k \stackrel{\tau}{\Longrightarrow} \Theta_k$. We define $\Theta' := \sum_i \Theta^{\times}_i$. By Additivity (Remark 3.7) we have $\Delta' \ \overline{\lhd^s_{FS}}\ \Theta'$. It remains to be shown that $\Theta \Longrightarrow \Theta'$.

For that final step, since $(\Theta \Longrightarrow)$ is closed according to Lemma 6.12, we can establish $\Theta \Longrightarrow \Theta'$ by exhibiting a sequence Θ'_i with $\Theta \Longrightarrow \Theta'_i$ for each i and with the Θ'_i's being arbitrarily close to Θ'. Induction establishes for each i that $\Theta \Longrightarrow \Theta'_i := (\Theta^{\rightarrow}_i + \sum_{k \leq i} \Theta^{\times}_k)$. Since $|\Delta'| = 1$, we must have $\lim_{i\rightarrow\infty} |\sum_{k=i}^{\infty} \Delta^{\rightarrow}_i| = 0$ and $\lim_{i\rightarrow\infty} |\Delta^{\rightarrow}_i| = 0$, whence by Lemma 3.6, using that $\Delta^{\rightarrow}_i \ \overline{\lhd^s_{FS}}\ \Theta^{\rightarrow}_i$, also $\lim_{i\rightarrow\infty} |\sum_{k=i}^{\infty} \Theta^{\rightarrow}_i| = 0$ and $\lim_{i\rightarrow\infty} |\Theta^{\rightarrow}_i| = 0$. Thus these Θ'_i's form the sequence we needed.

That concludes the case for $|\Delta'| = 1$. If on the other hand $\Delta' = \varepsilon$, i.e. we have $|\Delta'| = 0$, then $\Theta \Longrightarrow \varepsilon$ follows immediately from $s \lhd^s_{FS} \Theta$, and $\varepsilon \ \overline{\lhd^s_{FS}}\ \varepsilon$ trivially.

In the general case, if $s \Longrightarrow \Delta'$ then by Theorem 6.20 we have $s \Longrightarrow \Delta'_1 \ {}_p\!\oplus\ \Delta'_\varepsilon$ for some probability p and full distributions $\Delta'_1, \Delta'_\varepsilon$, with $\Delta' = p\cdot\Delta'_1$ and $\Delta'_\varepsilon \Longrightarrow \varepsilon$. From the mass-1 case above we have $\Theta \Longrightarrow \Theta'_1 \ {}_p\!\oplus\ \Theta'_\varepsilon$ with $\Delta'_1 \ \overline{\lhd^s_{FS}}\ \Theta'_1$ and $\Delta'_\varepsilon \ \overline{\lhd^s_{FS}}\ \Theta'_\varepsilon$; from the mass-0 case we have $\Theta'_\varepsilon \Longrightarrow \varepsilon$ and hence $\Theta'_1 \ {}_p\!\oplus\ \Theta'_\varepsilon \Longrightarrow p\cdot\Theta'_1$ by Theorem 3.20(i); thus transitivity yields $\Theta \Longrightarrow p\cdot\Theta'_1$, with $\Delta' = p\cdot\Delta'_1 \ \overline{\lhd^s_{FS}}\ p\cdot\Theta'_1$ as required, using Definition 3.2(2). □

The proof of Theorem 7.20 refers to Theorem 6.20 where the underlying pLTS is assumed to be finitary. As we would expect, Theorem 7.20 fails for infinitary pLTSs.

Example 7.21 We have seen in Example 7.19 that the state s from (19) is related to $\mathbf{0}$ via the relation $\lhd^s_{FS}$. We now compare s with $\mathbf{0}$ according to $\lhd_{FS}$. From state s we have the weak transition $\overline{s} \Longrightarrow \mathbf{0}\ {}_{1/2}\!\oplus\ \varepsilon$, which cannot be matched by any transition from $\mathbf{0}$, thus $s \not\lhd_{FS} \mathbf{0}$. This means that Theorem 7.20 fails for infinitely branching processes.

If we replace state s by the state s_2 from (21), similar phenomenon happens. Therefore, Theorem 7.20 also fails for finitely branching but infinite-state processes. □

7.3 Precongruence

The purpose of this section is to show that the semantic relation $\sqsupseteq_{FS}$ is preserved by the constructs of pCSP. The proofs follow closely the corresponding proofs in Section 4 of [2], but here there is a significant extra proof obligation: in order to relate two processes we have to demonstrate that if the first diverges then so does the second.

Here, in order to avoid such complications, we introduce yet another version of failure simulation; it modifies Definition 7.18 by checking divergence co-inductively instead of using a predicate.

Definition 7.22 Define $\lhd^c_{FS}$ to be the largest relation in $S \times \mathscr{D}(S)$ such that if $s \lhd^c_{FS} \Theta$ then

1. whenever $\overline{s} \Longrightarrow \varepsilon$, there are some Δ', Θ' such that $\overline{s} \Longrightarrow\xrightarrow{\tau}\Longrightarrow \Delta' \Longrightarrow \varepsilon$, $\Theta \Longrightarrow\xrightarrow{\tau}\Longrightarrow \Theta'$ and $\Delta' \; \overline{\lhd^c_{FS}} \; \Theta'$; otherwise
2. whenever $s \xrightarrow{\alpha} \Delta'$, for $\alpha \in \mathsf{Act}_\tau$, then there is a Θ' with $\Theta \stackrel{\alpha}{\Longrightarrow} \Theta'$ and $\Delta' \; \overline{\lhd^c_{FS}} \; \Theta'$, and
3. whenever $s \stackrel{A}{\not\longrightarrow}$ then $\Theta \Longrightarrow\stackrel{A}{\not\longrightarrow}$.

Lemma 7.23 The following statements about divergence are equivalent.

(1) $\Delta \Longrightarrow \varepsilon$.

(2) There is an infinite sequence $\Delta \xrightarrow{\tau} \Delta_1 \xrightarrow{\tau} \Delta_2 \xrightarrow{\tau} \dots$.

(3) There is an infinite sequence $\Delta \Longrightarrow\xrightarrow{\tau}\Longrightarrow \Delta_1 \Longrightarrow\xrightarrow{\tau}\Longrightarrow \Delta_2 \Longrightarrow\xrightarrow{\tau}\Longrightarrow \dots$.

Proof: By the definition of weak transition, it is immediate that $(1) \Leftrightarrow (2)$. Clearly we have $(2) \Rightarrow (3)$. To show that $(3) \Rightarrow (2)$, we introduce another characterisation of divergence. Let Δ be a subdistribution in a pLTS L. A *pLTS induced by* Δ is a pLTS whose states and transitions are subsets of those in L and all states are reachable from Δ.

(4) There is a pLTS induced by Δ where all states have outgoing τ transitions.

It holds that $(3) \Rightarrow (4)$ because we can construct a pLTS whose states and transitions are just those used in deriving the infinite sequence in (3). For this pLTS, each state has an outgoing τ transition, which gives $(4) \Rightarrow (2)$. □

The next lemma shows the usefulness of the relation $\lhd^c_{FS}$ by checking divergence in a co-inductive way.

Lemma 7.24 Suppose $\Delta \; \overline{\lhd^c_{FS}} \; \Theta$ and $\Delta \Longrightarrow \varepsilon$. Then there exist Δ', Θ' such that $\Delta \Longrightarrow\xrightarrow{\tau}\Longrightarrow \Delta' \Longrightarrow \varepsilon$, $\Theta \Longrightarrow\xrightarrow{\tau}\Longrightarrow \Theta'$, and $\Delta' \; \overline{\lhd^c_{FS}} \; \Theta'$.

Proof: Suppose $\Delta \; \overline{\lhd^c_{FS}} \; \Theta$ and $\Delta \Longrightarrow \varepsilon$. In analogy with Proposition 7.16 we can show that $\lhd^c_{FS}$ is convex. By Corollary 3.11, we can decompose Θ as $\sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot \Theta_s$ and $s \lhd^c_{FS} \Theta_s$ for each $s \in \lceil \Delta \rceil$. Now each s must also diverge. So there exist Δ'_s, Θ'_s such that $\overline{s} \Longrightarrow\xrightarrow{\tau}\Longrightarrow \Delta'_s \Longrightarrow \varepsilon$, $\Theta_s \Longrightarrow\xrightarrow{\tau}\Longrightarrow \Theta'_s$ and $\Delta'_s \; \overline{\lhd^c_{FS}} \; \Theta'_s$ for each $s \in \lceil \Delta \rceil$. Let $\Delta' = \sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot \Delta'_s$ and $\Theta' = \sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot \Theta'_s$. By Definition 3.2 and Theorem 3.20(i), we have $\Delta' \; \overline{\lhd^c_{FS}} \; \Theta'$, $\Delta \Longrightarrow\xrightarrow{\tau}\Longrightarrow \Delta'$, and $\Theta \Longrightarrow\xrightarrow{\tau}\Longrightarrow \Theta'$. We also have that $\Delta' \Longrightarrow \varepsilon$ because for each state s in Δ' it holds that $s \in \lceil \Delta'_s \rceil$ for some Δ'_s and $\Delta'_s \Longrightarrow \varepsilon$, which means $\overline{s} \Longrightarrow \varepsilon$. □

Lemma 7.25 $\lhd^c_{FS}$ coincides with $\lhd^s_{FS}$.

Proof: We only need to check that the first clause in Definition 7.18 is equivalent to the first clause in Definition 7.22. For one direction, we consider the relation

$$\mathscr{R} \;:=\; \{(s, \Theta) \mid \overline{s} \Longrightarrow \varepsilon \text{ and } \Theta \Longrightarrow \varepsilon\}$$

and show $\mathscr{R} \subseteq \lhd^c_{FS}$. Suppose $s \,\mathscr{R}\, \Theta$. By Lemma 7.23 there are two infinite sequences $s \xrightarrow{\tau} \Delta_1 \xrightarrow{\tau} \Delta_2 \xrightarrow{\tau} \dots$ and $\Theta \xrightarrow{\tau} \Theta_1 \xrightarrow{\tau} \dots$. Then we have both $\Delta_1 \Longrightarrow \varepsilon$ and $\Theta_1 \Longrightarrow \varepsilon$. Note that $\Delta_1 \Longrightarrow \varepsilon$ if and only if $\overline{t} \Longrightarrow \varepsilon$ for each $t \in \lceil \Delta_1 \rceil$. Therefore, $\Delta_1 \; \overline{\mathscr{R}} \; \Theta_1$ as we have $\Delta_1 = \sum_{t \in \lceil \Delta_1 \rceil} \Delta_1(t) \cdot \overline{t}$, $\Theta_1 = \sum_{t \in \lceil \Delta_1 \rceil} \Delta_1(t) \cdot \Theta_1$, and $t \,\mathscr{R}\, \Theta_1$. Here $|\Delta_1| = 1$ because Δ_1, like $\overline{s}$, is a distribution.

For the other direction, we show that $\Delta \; \overline{\lhd^c_{FS}} \; \Theta$ and $\Delta \Longrightarrow \varepsilon$ imply $\Theta \Longrightarrow \varepsilon$. Then as a special case, we get $s \lhd^c_{FS} \Theta$ and $s \Longrightarrow \varepsilon$ imply $\Theta \Longrightarrow \varepsilon$. By repeated application of Lemma 7.24, we can obtain two infinite sequences $\Delta \Longrightarrow\xrightarrow{\tau}\Longrightarrow \Delta_1 \Longrightarrow\xrightarrow{\tau}\Longrightarrow \dots$ and $\Theta \Longrightarrow\xrightarrow{\tau}\Longrightarrow \Theta_1 \Longrightarrow\xrightarrow{\tau}\Longrightarrow \dots$ such that $\Delta_i \; \overline{\lhd^c_{FS}} \; \Theta_i$ for all $i \geq 1$. By Lemma 7.23 this implies $\Theta \Longrightarrow \varepsilon$. □

The advantage of this new relation $\lhd^c_{FS}$ over $\lhd^s_{FS}$ is that in order to check $s \lhd^c_{FS} \Theta$ when s diverges it is sufficient to find a single matching move $\Theta \Longrightarrow\xrightarrow{\tau}\Longrightarrow \Theta'$, rather than an infinite sequence of moves. However to construct this matching move we cannot rely on clause 2 in Definition 7.22, as the move generated there might actually be empty, which we have seen in Example 3.14. Instead we need a method for generating weak moves which contain at least one occurrence of a τ-action.

Definition 7.26 [Productive moves] Let us write $s \mid_A t \xrightarrow{\tau}_p \Theta$ whenever we can infer $s \mid_A t \xrightarrow{\tau}_p \Theta$ from rule (PAR.R) or (PAR.I). In effect this means that t must contribute to the action.

These *productive* actions are extended to subdistributions in the standard manner, giving $\Delta \xrightarrow{\tau}_p \Theta$.

The following lemma appeared as Lemma 6.12 in [4]. It still holds in our current setting.

Lemma 7.27 (1) If $\Phi \Longrightarrow \Phi'$ then $\Phi \mid_A \Delta \Longrightarrow \Phi' \mid_A \Delta$ and $\Delta \mid_A \Phi \Longrightarrow \Delta \mid_A \Phi'$.

(2) If $\Phi \xrightarrow{a} \Phi'$ and $a \notin A$ then $\Phi \mid_A \Delta \xrightarrow{a} \Phi' \mid_A \Delta$ and $\Delta \mid_A \Phi \xrightarrow{a} \Delta \mid_A \Phi'$.

(3) If $\Phi \xrightarrow{a} \Phi'$, $\Delta \xrightarrow{a} \Delta'$ and $a \in A$ then $\Delta \mid_A \Phi \xrightarrow{\tau} \Delta' \mid_A \Phi'$.

(4) $(\sum_{j\in J} p_j \cdot \Phi_j) \mid_A (\sum_{k\in K} q_k \cdot \Delta_k) = \sum_{j\in J}\sum_{k\in K}(p_j \cdot q_k)\cdot(\Phi_j \mid_A \Delta_k)$.

(5) Given relations $\mathscr{R}, \mathscr{R}' \subseteq S \times \mathscr{D}(S)$ satisfying $u\mathscr{R}\Psi$ whenever $u = s \mid_A t$ and $\Psi = \Theta \mid_A t$ with $s\ \mathscr{R}'\ \Theta$ and $t \in S$. Then $\Delta\ \overline{\mathscr{R}'}\ \Theta$ and $\Phi \in \mathscr{D}(S)$ implies $(\Delta \mid_A \Phi)\ \overline{\mathscr{R}}\ (\Theta \mid_A \Phi)$. □

Proposition 7.28 Suppose $\Delta\ \overline{\lhd^c_{FS}}\ \Theta$ and $\Delta \mid_A t \xrightarrow{\tau}_p \Gamma$. Then $\Theta \mid_A t \Longrightarrow\xrightarrow{\tau}\Longrightarrow \Psi$ for some Ψ such that $\Gamma\ \overline{\mathscr{R}}\ \Psi$, where $\mathscr{R}$ is the relation given by $\{(s \mid_A t, \Theta \mid_A t) \mid s \lhd^c_{FS} \Theta\}$.

Proof: We first show a simplified version of the result. Suppose $s \lhd^c_{FS} \Theta$ and $s \mid_A t \xrightarrow{\tau}_p \Gamma$; we prove this entails $\Theta \mid_A t \Longrightarrow\xrightarrow{\tau}\Longrightarrow \Psi$ such that $\Gamma\ \overline{\mathscr{R}}\ \Psi$. There are only two possibilities for inferring the above productive move from $s \mid_A t$:

(i) $\Gamma = s \mid_A \Phi$ where $t \xrightarrow{\tau} \Phi$

(ii) or $\Gamma = \Delta \mid_A \Phi$ where for some $a \in A$, $s \xrightarrow{a} \Delta$ and $t \xrightarrow{a} \Phi$.

In the first case we have $\Theta \mid_A t \xrightarrow{\tau} \Theta \mid_A \Phi$ by Lemma 7.27(2) and $(s \mid_A \Phi)\ \overline{\mathscr{R}}\ (\Theta \mid_A \Phi)$ by Lemma 7.27(5), whereas in the second case $s \lhd^c_{FS} \Theta$ implies $\Theta \Longrightarrow\xrightarrow{a}\Longrightarrow \Theta'$ for some $\Theta' \in \mathscr{D}(S)$ with $\Delta\ \overline{\lhd^c_{FS}}\ \Theta'$, and we have $\Theta \mid_A t \Longrightarrow\xrightarrow{\tau}\Longrightarrow \Theta' \mid_A \Phi$ by Lemma 7.27(1) and (3), and $(\Delta \mid_A \Phi)\ \overline{\mathscr{R}}\ (\Theta' \mid_A \Phi)$ by Lemma 7.27(5).

The general case now follows using a standard decomposition/recomposition argument. Since $\Delta \mid_A t \xrightarrow{\tau}_p \Gamma$, Lemma 3.5 yields

$$\Delta = \sum_{i\in I} p_i \cdot \overline{s_i}, \qquad s_i \mid_A t \xrightarrow{\tau}_p \Gamma_i, \qquad \Gamma = \sum_{i\in I} p_i \cdot \Gamma_i,$$

for certain $s_i \in S$, $\Gamma_i \in \mathscr{D}(S)$ and $\sum_{i\in I} p_i \leq 1$. In analogy with Proposition 7.16 we can show that $\lhd^c_{FS}$ is convex. Hence, since $\Delta\ \overline{\lhd^c_{FS}}\ \Theta$, Corollary 3.11 yields that $\Theta = \sum_{i\in I} p_i \cdot \Theta_i$ for some $\Theta_i \in \mathscr{D}(S)$ such that $s_i \lhd^c_{FS} \Theta_i$ for $i \in I$. By the above argument we have $\Theta_i \mid_A t \Longrightarrow\xrightarrow{\tau}\Longrightarrow \Psi_i$ for some $\Psi_i \in \mathscr{D}(S)$ such that $\Gamma_i\ \overline{\mathscr{R}}\ \Psi_i$. The required Ψ can be taken to be $\sum_{i\in I} p_i \cdot \Psi_i$ as Definition 3.2(2) yields $\Gamma\ \overline{\mathscr{R}}\ \Psi$ and Theorem 3.20(i) and Definition 3.2(2) yield $\Theta \mid_A t \Longrightarrow\xrightarrow{\tau}\Longrightarrow \Psi$. □

Our next result shows that we can always factor out productive moves from an arbitrary action of a parallel process.

Lemma 7.29 Suppose $\Delta \mid_A t \xrightarrow{\tau} \Gamma$. Then there exists subdistributions $\Delta^{\rightarrow}$, $\Delta^{\times}$, Δ^{next}, $\Gamma^{\times}$ (possibly empty) such that

(i) $\Delta = \Delta^{\rightarrow} + \Delta^{\times}$

(ii) $\Delta^{\rightarrow} \xrightarrow{\tau} \Delta^{\text{next}}$

(iii) $\Delta^{\times} \mid_A t \xrightarrow{\tau}_p \Gamma^{\times}$

(iv) $\Gamma = \Delta^{\text{next}} \mid_A t + \Gamma^{\times}$

Proof: By Lemma 3.5 $\Delta \mid_A t \xrightarrow{\tau} \Gamma$ implies that

$$\Delta = \sum_{i\in I} p_i \cdot \overline{s_i}, \qquad s_i \mid_A t \xrightarrow{\tau} \Gamma_i, \qquad \Gamma = \sum_{i \in I} p_i \cdot \Gamma_i,$$

for certain $s_i \in S$, $\Gamma_i \in \mathscr{D}(S)$ and $\sum_{i\in I} p_i \leq 1$. Let $J = \{\, i \in I \mid s_i \mid_A t \xrightarrow{\tau}_p \Gamma_i \,\}$. Note that for each $i \in (I - J)$ Γ_i has the form $\Gamma'_i \mid_A t$, where $s_i \xrightarrow{\tau} \Gamma'_i$. Now let

$$\Delta^{\rightarrow} = \sum_{i\in(I-J)} p_i \cdot \overline{s_i}, \qquad \Delta^{\times} = \sum_{i\in J} p_i \cdot \overline{s_i}$$
$$\Delta^{\text{next}} = \sum_{i\in(I-J)} p_i \cdot \Gamma'_i, \qquad \Gamma^{\times} = \sum_{i\in J} p_i \cdot \Gamma_i$$

By construction (i) and (iv) are satisfied, and (ii) and (iii) follows by property (2) of Definition 3.2. □

Lemma 7.30 If $\Delta \mid_A t \Longrightarrow \varepsilon$ then there is a $\Delta' \in \mathscr{D}(S)$ such that $\Delta \Longrightarrow \Delta'$ and $\Delta' \mid_A t \xrightarrow{\tau}_p \Longrightarrow \varepsilon$.

Proof: Suppose $\Delta_0 \mid_A t \Longrightarrow \varepsilon$. By Lemma 7.23 there is an infinite sequence

$$\Delta_0 \mid_A t \xrightarrow{\tau} \Psi_1 \xrightarrow{\tau} \Psi_2 \xrightarrow{\tau} \dots \tag{22}$$

By induction on $k \geq 0$, we find distributions $\Gamma_{k+1},\ \Delta_k^{\rightarrow},\ \Delta_k^{\times},\ \Delta_{k+1},\ \Gamma_{k+1}^{\times}$ such that

(i) $\Delta_k \mid_A t \xrightarrow{\tau} \Gamma_{k+1}$

(ii) $\Gamma_{k+1} \leq \Psi_{k+1}$

(iii) $\Delta_k = \Delta_k^{\rightarrow} + \Delta_k^{\times}$

(iv) $\Delta_k^{\rightarrow} \xrightarrow{\tau} \Delta_{k+1}$

(v) $\Delta_k^{\times} \mid_A t \xrightarrow{\tau}_p \Gamma_{k+1}^{\times}$

(vi) $\Gamma_{k+1} = \Delta_{k+1} \mid_A t + \Gamma_{k+1}^{\times}$.

Induction Base: Take $\Gamma_1 := \Psi_1$ and apply Lemma 7.29.
Induction Step: Assume we already have Γ_k, Δ_k and $\Gamma_k^{\times}$. Since $\Delta_k \mid_A t \leq \Gamma_k \leq \Psi_k$ and $\Psi_k \xrightarrow{\tau} \Psi_{k+1}$, Proposition 3.10 gives us a Γ_{k+1} such that $\Delta_k \mid_A t \xrightarrow{\tau} \Gamma_{k+1}$ and $\Gamma_{k+1} \leq \Psi_{k+1}$. Now apply Lemma 7.29.

Let $\Delta' := \sum_{k=0}^{\infty} \Delta_k^{\times}$. By (iii) and (iv) above we obtain a weak τ move $\Delta_0 \Longrightarrow \Delta'$. Since $\Delta' \mid_A t = \sum_{k=0}^{\infty} (\Delta_k^{\times} \mid_A t)$, by (v) and Definition 3.2 we have $\Delta' \mid_A t \xrightarrow{\tau}_p \sum_{k=1}^{\infty} \Gamma_k^{\times}$. Note that here it does not matter if $\Delta' = \varepsilon$. Since $\Gamma_k^{\times} \leq \Gamma_k \leq \Psi_k$ and $\Psi_k \Longrightarrow \varepsilon$ it follows by Theorem 3.20(ii) that $\Gamma_k^{\times} \Longrightarrow \varepsilon$. Hence Theorem 3.20(i) yields $\sum_{k=1}^{\infty} \Gamma_k^{\times} \Longrightarrow \varepsilon$. □

We are now ready to prove the main result of this section, namely that $\sqsubseteq_{FS}$ is preserved by the parallel operator.

Proposition 7.31 *In a finitary pLTS, if* $\Delta \sqsupseteq_{FS} \Theta$ *then* $\Delta \mid_A \Phi \sqsupseteq_{FS} \Theta \mid_A \Phi$.

Proof: We first construct the following relation

$$\mathscr{R} \ := \ \{(s \mid_A t, \Theta \mid_A t) \mid s \lhd_{FS}^{c} \Theta\}$$

and check that $\mathscr{R} \subseteq \lhd_{FS}^{c}$. As in the proof of Proposition 4.6 in [2], one can check that each strong transition from $s \mid_A t$ can be matched up by a transition from $\Theta \mid_A t$, and the matching of failures can also be established. So we concentrate on the requirement involving divergence.

Suppose $s \lhd_{FS}^{c} \Theta$ and $\overline{s \mid_A t} \Longrightarrow \varepsilon$. We need to find some Γ, Ψ such that

(a) $\overline{s \mid_A t} \Longrightarrow \xrightarrow{\tau} \Longrightarrow \Gamma \Longrightarrow \varepsilon$,

(b) $\Theta \mid_A t \Longrightarrow \xrightarrow{\tau} \Longrightarrow \Psi$ and $\Gamma \ \overline{\mathscr{R}}\ \Psi$.

By Lemma 7.30 there are $\Delta', \Gamma \in \mathscr{D}(S)$ such that $\overline{s} \Longrightarrow \Delta'$ and $\Delta' \mid_A t \xrightarrow{\tau}_p \Gamma \Longrightarrow \varepsilon$. Since for finitary processes $\lhd^c_{FS}$ coincides with $\lhd^s_{FS}$ and $\lhd_{FS}$ by Lemma 7.25 and Theorem 7.20, there must be a $\Theta' \in \mathscr{D}(S)$ such that $\Theta \Longrightarrow \Theta'$ and $\Delta' \overline{\lhd^c_{FS}} \Theta'$. By Proposition 7.28 we have $\Theta' \mid_A t \Longrightarrow\xrightarrow{\tau}\Longrightarrow \Psi$ for some Ψ such that $\Gamma \overline{\mathscr{R}} \Psi$. Now $s \mid_A t \Longrightarrow \Delta' \mid_A t \xrightarrow{\tau} \Gamma \Longrightarrow \varepsilon$ and $\Theta \mid_A t \Longrightarrow \Theta' \mid_A t \Longrightarrow\xrightarrow{\tau}\Longrightarrow \Psi$ with $\Gamma \overline{\mathscr{R}} \Psi$, which had to be shown.

Therefore, we have shown that $\mathscr{R} \subseteq \lhd^c_{FS}$. Now let us focus our attention on the statement of the proposition, which involves $\sqsupseteq_{FS}$.

Suppose $\Delta \sqsupseteq_{FS} \Theta$. By Proposition 7.7 this means that there is some Θ^{match} such that $\Theta \Longrightarrow \Theta^{\text{match}}$ and $\Delta \overline{\lhd_{FS}} \Theta^{\text{match}}$. By Theorem 7.20 and Lemma 7.25 we have $\Delta \overline{\lhd^c_{FS}} \Theta^{\text{match}}$. Then Lemma 7.27(5) yields $(\Delta \mid_A \Phi) \overline{\mathscr{R}} (\Theta^{\text{match}} \mid_A \Phi)$. Therefore, we have $(\Delta \mid_A \Phi) \overline{\lhd^c_{FS}} (\Theta^{\text{match}} \mid_A \Phi)$, i.e. $(\Delta \mid_A \Phi) \overline{\lhd_{FS}} (\Theta^{\text{match}} \mid_A \Phi)$ by Lemma 7.25 and Theorem 7.20. By Lemma 7.27(1) we also have $(\Theta \mid_A \Phi) \Longrightarrow (\Theta^{\text{match}} \mid_A \Phi)$, which had to be established according to Proposition 7.7. □

In the proof of Proposition 7.31 we use the characterisation of $\lhd_{FS}$ as $\lhd^s_{FS}$, which assumes the pLTS to be finitary. In general, the relation $\lhd^s_{FS}$ is not closed under parallel composition.

Example 7.32 We use a modification of the infinite state pLTSs in Example 3.17 which as before has states s_k with $k \geq 2$, but we add an extra a-looping state s_a to give all together the system

$$\text{for } k \geq 2 \qquad s_k \xrightarrow{\tau} (\overline{s_a} \,{}_{\frac{1}{k^2}}\!\oplus \overline{s_{k+1}}) \qquad \text{and} \qquad s_a \xrightarrow{a} \overline{s_a}\ .$$

There is a failure simulation $s_k \lhd^s_{FS} (\overline{s_a} \,{}_{\frac{1}{k}}\!\oplus \overline{\mathbf{0}})$ because the transition $s_k \xrightarrow{\tau} (\overline{s_a} \,{}_{\frac{1}{k^2}}\!\oplus \overline{s_{k+1}})$ can be matched by a transition to $(\overline{s_a} \,{}_{\frac{1}{k^2}}\!\oplus (\overline{s_a} \,{}_{\frac{1}{k+1}}\!\oplus \overline{\mathbf{0}}))$ which simplifies to just $(\overline{s_a} \,{}_{\frac{1}{k}}\!\oplus \overline{\mathbf{0}})$ again — i.e. a sufficient simulating transition would be the identity instance of $\Longrightarrow$.

Now $s_2 \mid_{\{a\}} s_a$ wholly diverges even though s_2 itself does not, and (recall from above) we have $s_2 \lhd^s_{FS} (\overline{s_a} \,{}_{\frac{1}{2}}\!\oplus \mathbf{0})$. Yet $(\overline{s_a} \,{}_{\frac{1}{2}}\!\oplus \mathbf{0}) \mid_{\{a\}} s_a$ does not diverge, thus $s_2 \mid_{\{a\}} s_a \not\lhd^s_{FS} (\overline{s_a} \,{}_{\frac{1}{2}}\!\oplus \mathbf{0}) \mid_{\{a\}} s_a$.

Note that this counter-example does not go through if we use failure similarity $\lhd_{FS}$ instead of simple failure similarity $\lhd^s_{FS}$, since $s_2 \not\lhd_{FS} (\overline{s_a} \,{}_{\frac{1}{2}}\!\oplus \mathbf{0})$ — the former has the transition $s_2 \Longrightarrow s_a \,{}_{\frac{1}{2}}\!\oplus \varepsilon$, which cannot be matched by $s_a \,{}_{\frac{1}{2}}\!\oplus \mathbf{0}$. □

Proposition 7.33 (Precongruence) In a finitary pLTS, if $P \sqsupseteq_{FS} Q$ then $\alpha.P \sqsupseteq_{FS} \alpha.Q$ for $\alpha \in \mathsf{Act}$, and similarly if $P_1 \sqsupseteq_{FS} Q_1$ and $P_2 \sqsupseteq_{FS} Q_2$ then $P_1 \odot P_2 \ \sqsupseteq_{FS} \ Q_1 \odot Q_2$ for $\odot$ being any of the operators $\sqcap$, $\Box$, ${}_p\oplus$ and $\mid_A$.

Proof: The most difficult case is the closure of failure simulation under parallel composition, which is proved in Proposition 7.31. The other cases are simpler, thus omitted. □

Lemma 7.34 In a finitary pLTS, if $P \sqsubseteq_{FS} Q$ then for any test T it holds that $[P \mid_{\mathsf{Act}} T] \sqsubseteq_{FS} [Q \mid_{\mathsf{Act}} T]$.

Proof: We first construct the following relation

$$\mathscr{R} \ := \ \{(s \mid_{\mathsf{Act}} t, \Theta \mid_{\mathsf{Act}} t) \mid s \lhd^c_{FS} \Theta\}$$

where $s \mid_{\mathsf{Act}} t$ is a state in $[P \mid_{\mathsf{Act}} T]$ and $\Theta \mid_{\mathsf{Act}} t$ is a subdistribution in $[Q \mid_{\mathsf{Act}} T]$, and show that $\mathscr{R} \subseteq \lhd^c_{FS}$.

1. The matching of divergence between $s \mid_{\mathsf{Act}} t$ and $\Theta \mid_{\mathsf{Act}} t$ is almost the same as the proof of Proposition 7.31, besides that we need to check the requirements $t \stackrel{\omega}{\not\longrightarrow}$ and $\Gamma \stackrel{\omega}{\not\longrightarrow}$ are always met there.

2. We now consider the matching of transitions.

   - If $s \mid_{\mathsf{Act}} t \xrightarrow{\omega}$ then this action is actually performed by t. Suppose $t \xrightarrow{\omega} \Gamma$. Then $s \mid_{\mathsf{Act}} t \xrightarrow{\omega} s \mid_{\mathsf{Act}} \Gamma$ and $\Theta \mid_{\mathsf{Act}} t \xrightarrow{\omega} \Theta \mid_{\mathsf{Act}} \Gamma$. Obviously we have $(s \mid_{\mathsf{Act}} \Gamma, \Theta \mid_{\mathsf{Act}} \Gamma) \in \overline{\mathscr{R}}$.
   - If $s \mid_{\mathsf{Act}} t \xrightarrow{\tau}$ then we must have $s \mid_{\mathsf{Act}} t \stackrel{\omega}{\not\longrightarrow}$, otherwise the τ transition would be a "scooting" transition and the pLTS is not ω-respecting. It follows that $t \stackrel{\omega}{\not\longrightarrow}$. There are three subcases.
     – $t \xrightarrow{\tau} \Gamma$. So the transition $s \mid_{\mathsf{Act}} t \xrightarrow{\tau} s \mid_{\mathsf{Act}} \Gamma$ can simply be matched up by $\Theta \mid_{\mathsf{Act}} t \xrightarrow{\tau} \Theta \mid_{\mathsf{Act}} \Gamma$.

– $s \xrightarrow{\tau} \Delta$. Since $s \lhd_{FS}^c \Theta$, there exists some Θ' such that $\Theta \Longrightarrow \Theta'$ and $\Delta \; \overline{\lhd_{FS}^c} \; \Theta'$. Note that in this case $t \stackrel{\omega}{\not\longrightarrow}$. It follows that $\Theta \mid_{\mathsf{Act}} t \Longrightarrow \Theta' \mid_{\mathsf{Act}} t$ which can match up the transition $s \mid_{\mathsf{Act}} t \longrightarrow \Delta \mid_{\mathsf{Act}} t$ because $(\Delta \mid_{\mathsf{Act}} t, \Theta' \mid_{\mathsf{Act}} t) \in \overline{\mathscr{R}}$.

– $s \xrightarrow{a} \Delta$ and $t \xrightarrow{a} \Gamma$ for some action $a \in \mathsf{Act}$. Since $s \lhd_{FS}^c \Theta$, there exists some Θ' such that $\Theta \stackrel{a}{\Longrightarrow} \Theta'$ and $\Delta \; \overline{\lhd_{FS}^c} \; \Theta'$. Note that in this case $t \stackrel{\omega}{\not\longrightarrow}$. It follows that $\Theta \mid_{\mathsf{Act}} t \Longrightarrow \Theta' \mid_{\mathsf{Act}} \Gamma$ which can match up the transition $s \mid_{\mathsf{Act}} t \longrightarrow \Delta \mid_{\mathsf{Act}} \Gamma$ because $(\Delta \mid_{\mathsf{Act}} \Gamma, \Theta' \mid_{\mathsf{Act}} \Gamma) \in \overline{\mathscr{R}}$.

- Suppose $s \mid_{\mathsf{Act}} t \stackrel{A}{\not\longrightarrow}$ for any $A \subseteq \mathsf{Act} \cup \{\omega\}$. There are two possibilities.

  – If $s \mid_{\mathsf{Act}} t \stackrel{\omega}{\not\longrightarrow}$, then $t \stackrel{\omega}{\not\longrightarrow}$ and there are two subsets A_1, A_2 of A such that $s \stackrel{A_1}{\not\longrightarrow}$, $t \stackrel{A_2}{\not\longrightarrow}$ and $A = A_1 \cup A_2$. Since $s \lhd_{FS}^c \Theta$ there exists some Θ' such that $\Theta \Longrightarrow \Theta'$ and $\Theta' \stackrel{A_1}{\not\longrightarrow}$. Therefore, we have $\Theta \mid_{\mathsf{Act}} t \Longrightarrow \Theta' \mid_{\mathsf{Act}} t \stackrel{A}{\not\longrightarrow}$.

  – If $s \mid_{\mathsf{Act}} t \xrightarrow{\omega}$ then $t \xrightarrow{\omega}$ and $\omega \notin A$. Therefore, we have $\Theta \mid_{\mathsf{Act}} t \xrightarrow{\omega}$ and $\Theta \mid_{\mathsf{Act}} t \stackrel{\tau}{\not\longrightarrow}$ because there is no "scooting" transition in $\Theta \mid_{\mathsf{Act}} t$. It follows that $\Theta \mid_{\mathsf{Act}} t \stackrel{A}{\not\longrightarrow}$.

Therefore, we have shown that $\mathscr{R} \subseteq \lhd_{FS}^c$, from which our expected result can be establishing using similar arguments in the last part of the proof of Proposition 7.31. □

7.4 Soundness

In this section we prove that failure simulations are sound for showing that processes are related via the failure-based testing preorder. We assume initially that we are using only one success action ω, so that $|\Omega| = 1$.

Because we prune our pLTSs before extracting values from them, we will be concerned mainly with ω-respecting structures.

Definition 7.35 Let Δ be a subdistribution in a pLTS $\langle S, \{\omega, \tau\}, \rightarrow\rangle$. We write $\mathscr{V}(\Delta)$ for the set of testing outcomes $\{\$\Delta' \mid \Delta \Longrightarrow\!\!\!\succ \Delta'\}$.

Lemma 7.36 Let Δ and Θ be subdistributions in an ω-respecting pLTS $\langle S, \{\tau, \omega\}, \rightarrow\rangle$. If subdistribution Δ is stable and $\Delta \; \overline{\lhd_{FS}} \; \Theta$, then $\mathscr{V}(\Theta) \leq_{\mathrm{Sm}} \mathscr{V}(\Delta)$.

Proof: We first show that if s is stable and $s \lhd_{FS} \Theta$ then $\mathscr{V}(\Theta) \leq_{\mathrm{Sm}} \mathscr{V}(s)$. Since s is stable, we have only two cases:

(i) $s \not\longrightarrow$ Here $\mathscr{V}(s) = \{0\}$ and since $s \lhd_{FS} \Theta$ we have $\Theta \Longrightarrow \Theta'$ with $\Theta' \not\longrightarrow$, whence in fact $\Theta \Longrightarrow\!\!\!\succ \Theta'$ and $\$\Theta' = 0$. Thus $0 \in \mathscr{V}(\Theta)$ which means $\mathscr{V}(\Theta) \leq_{\mathrm{Sm}} \mathscr{V}(s)$.

(ii) $s \xrightarrow{\omega} \Delta'$ for some Δ' Here $\mathscr{V}(s){=}\{1\}$ and $\Theta \Longrightarrow \Theta' \xrightarrow{\omega}$ with $\$\Theta'{=}|\Theta'|$. Because the pLTS is ω-respecting, in fact $\Theta \Longrightarrow\!\!\!\succ \Theta'$ and so again $\mathscr{V}(\Theta) \leq_{\mathrm{Sm}} \mathscr{V}(s)$.

Now for the general case we suppose $\Delta \; \overline{\lhd_{FS}} \; \Theta$. Use Proposition 3.10 to decompose Θ into $\sum_{s \in \lceil\Delta\rceil} \Delta(s) \cdot \Theta_s$ such that $s \lhd_{FS} \Theta_s$ for each $s \in \lceil\Delta\rceil$, and recall each such state s is stable. From above we have that $\mathscr{V}(\Theta_s) \leq_{\mathrm{Sm}} \mathscr{V}(s)$ for those s, and so $\mathscr{V}(\Theta) = \sum_{s \in \lceil\Delta\rceil} \Delta(s) \cdot \mathscr{V}(\Theta_s) \leq_{\mathrm{Sm}} \sum_{\in \lceil\Delta\rceil} \Delta(s) \cdot \mathscr{V}(s) = \mathscr{V}(\Delta)$. □

Lemma 7.37 Let Δ be a subdistribution in an ω-respecting pLTS $\langle S, \{\tau, \omega\}, \rightarrow\rangle$. If $\Delta \Longrightarrow \Delta'$ then $\mathscr{V}(\Delta') \subseteq \mathscr{V}(\Delta)$.

Proof: Note that if $\Delta' \Longrightarrow\!\!\!\succ \Delta''$ then $\Delta \Longrightarrow \Delta' \Longrightarrow\!\!\!\succ \Delta''$, so that every extreme derivative of Δ' is also an extreme derivative of Δ. □

Lemma 7.38 Let Δ and Θ be subdistributions in an ω-respecting pLTS $\langle S, \{\tau, \omega\}, \rightarrow\rangle$. If $\Theta \sqsubseteq_{FS} \Delta$, then it holds that $\mathscr{V}(\Theta) \leq_{\mathrm{Sm}} \mathscr{V}(\Delta)$.

Proof: Let Δ and Θ be subdistributions in an ω-respecting pLTS $\langle S, \{\tau, \omega\}, \rightarrow\rangle$. We first claim that

$$\text{If } \Delta \; \overline{\lhd_{FS}} \; \Theta, \text{ then } \mathscr{V}(\Theta) \leq_{\mathrm{Sm}} \mathscr{V}(\Delta).$$

We assume $\Delta \; \overline{\lhd_{FS}} \; \Theta$. For any $\Delta \Longrightarrow\!\!\!\succ \Delta'$ we have the matching transition $\Theta \Longrightarrow \Theta'$ such that $\Delta' \; \overline{\lhd_{FS}} \; \Theta'$. It follows from Lemmas 7.36 and 7.37 that $\mathscr{V}(\Theta) \supseteq \mathscr{V}(\Theta') \leq_{\mathrm{Sm}} \mathscr{V}(\Delta')$. Consequently, we obtain $\mathscr{V}(\Theta) \leq_{\mathrm{Sm}} \mathscr{V}(\Delta)$.

Now suppose $\Theta \sqsubseteq_{FS} \Delta$. By Proposition 7.7, there exists some Θ' such that $\Theta \Longrightarrow \Theta'$ and $\Delta \; \overline{\lhd_{FS}} \; \Theta'$. By the above claim and Lemma 7.37 we obtain $\mathscr{V}(\Theta) \supseteq \mathscr{V}(\Theta') \leq_{\mathrm{Sm}} \mathscr{V}(\Delta)$, thus $\mathscr{V}(\Theta) \leq_{\mathrm{Sm}} \mathscr{V}(\Delta)$. □

Theorem 7.39 For any finitary processes P and Q, if $P \sqsubseteq_{FS} Q$ then $P \sqsubseteq_{\text{pmust}} Q$.

Proof: We reason as follows.

	$P \sqsubseteq_{FS} Q$	
implies	$[P \mid_{\mathsf{Act}} T] \sqsubseteq_{FS} [Q \mid_{\mathsf{Act}} T]$	Lemma 7.34, for any test T
implies	$\mathscr{V}([P \mid_{\mathsf{Act}} T]) \leq_{\text{Sm}} \mathscr{V}([Q \mid_{\mathsf{Act}} T])$	$[\cdot]$ is ω-respecting; Lemma 7.38
iff	$\mathscr{A}(T, P) \leq_{\text{Sm}} \mathscr{A}(T, Q)$	Definition 7.35
iff	$P \sqsubseteq_{\text{pmust}} Q$.	Definition 4.9

□

In the proof of the soundness result above we use Lemma 7.34, which holds for finitary processes only. For infinitary processes, a preorder induced by $\lhd^s_{FS}$ is not sound for must testing.

Example 7.40 We have seen in Example 7.19 that the state s from (19) is related to $\overline{\mathbf{0}}$ via the relation $\lhd^s_{FS}$. If we apply test $\tau.\omega$ to both s and $\mathbf{0}$, we obtain $(0, 1]$ as an outcome set for the former and $\{1\}$ for the latter. Although $\overline{s}\ \overline{\lhd^s_{FS}}\ \overline{\mathbf{0}}$, we have $\mathscr{A}(\tau.\omega, \mathbf{0}) \not\leq_{\text{Sm}} \mathscr{A}(\tau.\omega, s)$.

If we replace state s by the state s_2 from (21), similar phenomenon happens. Although $\overline{s_2}\ \overline{\lhd^s_{FS}}\ \overline{\mathbf{0}}$, we have $\mathscr{A}(\tau.\omega, \mathbf{0}) = \{1\} \not\leq_{\text{Sm}} \{\frac{1}{2}\} = \mathscr{A}(\tau.\omega, s_2)$. □

8 Failure simulation is complete for must testing

This section establishes the completeness of the failure simulation preorder w.r.t. the must testing preorder. It does so in three steps. First we provide a characterisation of the preorder relation $\sqsubseteq_{FS}$ by an inductively defined relation. Secondly, using this, we develop a modal logic which can be used to characterise the failure simulation preorder on finitary pLTSs. Finally, we adapt the results of [2] to show that the modal formulae can in turn be characterised by tests; again this result depends on the underlying pLTS being finitary. From this, completeness follows.

8.1 Inductive characterisation

The relation $\lhd^s_{FS}$ of Definition 7.18 is given co-inductively: it is the largest fixed point of an equation $\mathscr{R} = \mathscr{F}(\mathscr{R})$. An alternative approach therefore is to use that $\mathscr{F}(-)$ to define $\lhd^s_{FS}$ as a limit of approximants:

Definition 8.1 For every $k \geq 0$ we define the relations $\lhd^k_{FS} \subseteq S \times \mathscr{D}(S)$ as follows:

(i) $\lhd^0_{FS} \ := \ S \times \mathscr{D}(S)$

(ii) $\lhd^{k+1}_{FS} := \ \mathscr{F}(\lhd^k_{FS})$

Finally let $\lhd^\infty_{FS} := \bigcap_{k=0}^\infty \lhd^k_{FS}$.

A simple inductive argument ensures that $\lhd^s_{FS} \subseteq \lhd^k_{FS}$, for every $k \geq 0$, and therefore that $\lhd^s_{FS} \subseteq \lhd^\infty_{FS}$. The converse is however not true in general.

A (non-probabilistic) example is well-known in the literature: it makes essential use of an infinite branching. Let P be the process $\mathsf{rec}\, x.\, a.x$ and s a state in a pLTS which starts by making an infinitary choice, namely for each $k \geq 0$ it has the option to perform a sequence of a actions with length k in succession and then deadlock. This can be described by the infinitary CSP expression $\bigsqcap_{k=0}^\infty a^k$. Then $[\![P]\!] \not\lhd^s_{FS} s$, because the move $[\![P]\!] \xrightarrow{a} [\![P]\!]$ can not be matched by s. However an easy inductive argument shows that $[\![P]\!] \lhd^k_{FS} a^k$ for every k, and therefore that $[\![P]\!] \lhd^\infty_{FS} s$.

Once we restrict our non-probabilistic systems to be finitely branching, however, a simple counting argument will show that $\lhd^s_{FS}$ coincides with $\lhd^\infty_{FS}$; see [9, Theorem 2.1] for the argument applied to bisimulation equivalence. In the probabilistic case we restrict to both finite-state *and* finite-branching -systems, and the effect of that is captured by topological *compactness*. Finiteness is lost unavoidably when we remember that e.g. the process $a \sqcap b$ can move via $\Longrightarrow$ to a distribution $[\![a]\!]\ {}_p\!\oplus\ [\![b]\!]$ for any of the uncountably many probabilities $p \in [0, 1]$. Nevertheless, those uncountably many weak transitions can be generated by arbitrary interpolation of two transitions $[\![a \sqcap b]\!] \xrightarrow{\tau} [\![a]\!]$ and $[\![a \sqcap b]\!] \xrightarrow{\tau} [\![b]\!]$, and that is the key structural property that compactness captures.

Because compactness follows from closure and boundedness, we approach this topic via closure.

Note that the metric space $(\mathscr{D}(S), d_1)$ with $d_1(\Delta, \Theta) = max_{s\in S}|\Delta(s) - \Theta(s)|$ and $(S \rightarrow \mathscr{D}(S), d_2)$ with $d_2(f, g) = max_{s\in S} d_1(f(s), g(s))$ are complete. Let X be a subset of either $\mathscr{D}(S)$ or $S \rightarrow \mathscr{D}(S)$. Clearly, X is bounded. So if X is closed, it is also compact.

Definition 8.2 A relation $\mathscr{R} \subseteq S \times \mathscr{D}(S)$ is said to be *closed* if for every $s \in S$ the set $s\cdot \mathscr{R} = \{\, \Delta \mid s \mathrel{\mathscr{R}} \Delta \,\}$ is closed.

Two examples of closed relations are $\Longrightarrow$ and $\stackrel{a}{\Longrightarrow}$ for any a, as shown by Lemma 6.12 and Corollary 6.15.

Our next step is to show that each of the relations $\lhd^k_{FS}$ are closed. This requires some results to be first established.

Lemma 8.3 Let $\mathscr{R} \subseteq S \times \mathscr{D}(S)$ be closed. Then $\mathbf{Ch}(\mathscr{R})$ is also closed.

Proof: Straightforward. □

Corollary 8.4 Let $\mathscr{R} \subseteq S \times \mathscr{D}(S)$ be closed and convex. Then $\overline{\mathscr{R}}$ is also closed.

Proof: For any $\Delta \in \mathscr{D}(S)$, we know from Proposition 3.9 that $\Delta\cdot \overline{\mathscr{R}} = \{\mathrm{Exp}_\Delta(f) \mid f \in \mathbf{Ch}(\mathscr{R})\}$. The function $\mathrm{Exp}_\Delta(-)$ is continuous. By Lemma 8.3 the set of choice functions of $\mathscr{R}$ is closed, and it is also bounded, thus being compact. Its image is also compact, thus being closed. □

Lemma 8.5 Let $\mathscr{R} \subseteq S \times \mathscr{D}(S)$ be closed and convex, and $C \subseteq \mathscr{D}(S)$ be closed. Then the set $\{\, \Delta \mid \Delta\cdot \overline{\mathscr{R}} \cap C \neq \emptyset \,\}$ is also closed.

Proof: First define $\mathscr{E} : \mathscr{D}(S) \times (S \rightarrow \mathscr{D}(S)) \rightarrow \mathscr{D}(S)$ by $\mathscr{E}(\Theta, f) = \mathrm{Exp}_\Theta(f)$, which is obviously continuous. Then we know from the previous lemma that $\mathbf{Ch}(\mathscr{R})$ is closed. Finally let

$$Z = \pi_1(\mathscr{E}^{-1}(C) \cap (\mathscr{D}(S) \times \mathbf{Ch}(\mathscr{R})))$$

where π_1 is the projection onto the first component of a pair. We observe that the continuity of $\mathscr{E}$ ensures that the inverse image of the closed set C is closed. Furthermore, $\mathscr{E}^{-1}(C) \cap (\mathscr{D}(S) \times \mathbf{Ch}(\mathscr{R}))$ is compact because it is both closed and bounded. Its image under the continuous function π_1 is also compact. It follows that Z is closed. But $Z = \{\, \Delta \mid \Delta\cdot \overline{\mathscr{R}} \cap C \neq \emptyset \,\}$ because

$$\begin{aligned} \Delta \in Z \ &\text{iff } (\Delta, f) \in \mathscr{E}^{-1}(C) \text{ for some } f \in \mathbf{Ch}(\mathscr{R}) \\ &\text{iff } \mathscr{E}(\Delta, f) \in C \text{ for some } f \in \mathbf{Ch}(\mathscr{R}) \\ &\text{iff } \mathrm{Exp}_\Delta(f) \in C \text{ for some } f \in \mathbf{Ch}(\mathscr{R}) \\ &\text{iff } \Delta \ \overline{R}\ \Delta' \text{ for some } \Delta' \in C \end{aligned}$$

The last line is an application of Proposition 3.9, which requires convexity of $\mathscr{R}$. □

An immediate corollary of this last result is:

Corollary 8.6 In a finitary pLTS the following sets are closed:

(i) $\{\, \Delta \mid \Delta \Longrightarrow \varepsilon \,\}$

(ii) $\{\, \Delta \mid \Delta \Longrightarrow\stackrel{A}{\not\longrightarrow} \,\}$

Proof: By Lemma 6.12 we see that $\Longrightarrow$ is closed and convex. Therefore, we can apply the previous lemma with $C = \{\varepsilon\}$ to obtain the first result. To obtain the second we apply it with $C = \{\, \Theta \mid \Theta \stackrel{A}{\not\longrightarrow} \,\}$, which is easily seen to be closed. □

The result is also used in the proof of:

Proposition 8.7 *In a finitary pLTS, for every $k \geq 0$ the relation $\lhd^k_{FS}$ is closed and convex.*

Proof: By induction on k. For $k = 0$ it is obvious. So let us assume that $\lhd^k_{FS}$ is closed and convex. We have to show that

$$s \cdot \lhd^{(k+1)}_{FS} \text{ is closed and convex, for every state } s \tag{23}$$

If $s \Longrightarrow \varepsilon$ then this follows from the corollary above, since in this case $s \cdot \lhd^{(k+1)}_{FS}$ coincides with $\{ \Delta \mid \Delta \Longrightarrow \varepsilon \}$. So let us assume that this is not the case.

For every $A \subseteq \mathsf{Act}$ let $R_A = \{ \Delta \mid \Delta \Longrightarrow\xrightarrow{A}\!\!\!\!\not\ \}$, which we know by the corollary above to be closed and is obviously convex. Also for every $\Theta,\ \alpha$ let $G_{\Theta,\alpha} = \{ \Delta \mid (\Delta \cdot \stackrel{\alpha}{\Longrightarrow}) \cap (\Theta \cdot \overline{\lhd^k_{FS}}) \neq \emptyset \}$. By Corollary 6.16, the relation $\stackrel{\alpha}{\Longrightarrow}$ is lifted from a closed convex relation. By Corollary 8.4, the assumption that $\lhd^k_{FS}$ is closed and convex implies that $\overline{\lhd^k_{FS}}$ is also closed. So we can appeal to Lemma 8.5 and conclude that each $G_{\Theta,\alpha}$ is closed. By Definition 3.2(2) it is also easy to see that $G_{\Theta,\alpha}$ is convex. But it follows that $s \cdot \lhd^{(k+1)}_{FS}$ is also closed and convex as it can be written as

$$\cap \{ R_A \mid s \xrightarrow{A}\!\!\!\!\not\ \} \ \cap \ \cap \{ G_{\Theta,\alpha} \mid s \xrightarrow{\alpha} \Theta \}$$

□

Before the main result of this section we need one more technical result.

Lemma 8.8 Let S be a finite set of states. Suppose $\mathscr{R}^k \subseteq S \times \mathscr{D}(S)$ is a sequence of closed convex relations such that $\mathscr{R}^{(k+1)} \subseteq \mathscr{R}^k$. Then $(\cap_{k=0}^{\infty} \overline{\mathscr{R}^k}) \subseteq \overline{(\cap_{k=0}^{\infty} \mathscr{R}^k)}$.

Proof: Let $\mathscr{R}^\infty$ denote $(\cap_{k=0}^{\infty} \mathscr{R}^k)$, and suppose $\Delta \ \overline{\mathscr{R}^k}\ \Theta$ for every $k \geq 0$. We have to show that $\Delta \ \overline{\mathscr{R}^\infty}\ \Theta$.

Let $G = \{ f : S \to \mathscr{D}(S) \mid \Theta = \mathrm{Exp}_\Delta(f) \}$, which is easily seen to be a closed set. For each k let we know from Lemma 8.3 that the set $\mathbf{Ch}(\mathscr{R}^k)$ is closed. Finally consider the collection of closed sets $H^k = \mathbf{Ch}(\mathscr{R}^k) \cap G$; since $\Delta \ \overline{\mathscr{R}^k}\ \Theta$, Proposition 3.9 assures us that all of these are non-empty. Also $H^{(k+1)} \subseteq H^k$ and therefore by the finite-intersection property [16] $\cap_{k=0}^{\infty} H^k$ is also non-empty.

Let f be an arbitrary element of this intersection. For any state $s \in dom(\mathscr{R}^\infty)$, and for every $k \geq 0$ since $dom(\mathscr{R}^\infty) \subseteq dom(\mathscr{R}^k)$ we have $s \ \mathscr{R}^k \ f(s)$, that is $s \ \mathscr{R}^\infty \ f(s)$. So f is a choice function for $\mathscr{R}^\infty$, $f \in \mathbf{Ch}(\mathscr{R}^\infty)$. From convexity and Proposition 3.9 it follows that $\Delta \ \overline{\mathscr{R}^\infty}\ \mathrm{Exp}_\Delta(f)$. But from the definition of the G we know that $\Theta = \mathrm{Exp}_\Delta(f)$, and the required result follows. □

Theorem 8.9 In a finitary pLTS, $s \lhd^s_{FS} \Theta$ if and only if $s \lhd^\infty_{FS} \Theta$.

Proof: Since $\lhd^s_{FS} \subseteq \lhd^\infty_{FS}$ it is sufficient to show the opposite inclusion, which by definition holds if $\lhd^\infty_{FS}$ is a failure simulation, viz. if $\lhd^\infty_{FS} \subseteq \mathscr{F}(\lhd^\infty_{FS})$. Suppose $s \lhd^\infty_{FS} \Theta$, which means that $s \lhd^k_{FS} \Theta$ for every $k \geq 0$. According to Definition 7.18, in order to show $s \ \mathscr{F}(\lhd^\infty_{FS})\ \Theta$ we have to establish three properties, the first and last of which are trivial (for they are independent on the argument of $\mathscr{F}$).

So suppose $s \xrightarrow{\alpha} \Delta'$. We have to show that $\Theta \stackrel{\alpha}{\Longrightarrow} \Theta'$ for some Θ' such that $\Delta' \ \overline{\lhd^\infty_{FS}}\ \Theta'$.

For every $k \geq 0$ there exists some Θ'_k such that $\Theta \stackrel{\alpha}{\Longrightarrow} \Theta'_k$ and $\Delta' \ \overline{\lhd^k_{FS}}\ \Theta'_k$. Now construct the sets

$$D^k = \{ \Theta' \mid \Theta \stackrel{\alpha}{\Longrightarrow} \Theta' \text{ and } \Delta' \ \overline{\lhd^k_{FS}}\ \Theta' \}.$$

From Lemma 6.12 and Proposition 8.7 we know that these are closed. They are also non-empty and $D^{k+1} \subseteq D^k$. So by the finite-intersection property the set $\bigcap_{k=0}^{\infty} D^k$ is non-empty. For any Θ' in it we know $\Theta \stackrel{\alpha}{\Longrightarrow} \Theta'$ and $\Delta' \ \overline{\lhd^k_{FS}}\ \Theta'$ for every $k \geq 0$. By Proposition 8.7, the relations $\lhd^k_{FS}$ are all closed and convex. Therefore, Lemma 8.8 may be applied to them, which enables us to conclude $\Delta' \ \overline{\lhd^\infty_{FS}}\ \Theta'$. □

For Theorem 8.9 to hold, it is crucial that the pLTS is assumed to be finitary.

Example 8.10 Consider an infinitely branching pLTS with four states $s, t, u, v, \mathbf{0}$ and the transitions are

- $s \xrightarrow{a} \overline{\mathbf{0}}\ _{\frac{1}{2}}\oplus\ \overline{s}$
- $t \xrightarrow{a} \overline{\mathbf{0}}$, $t \xrightarrow{a} \overline{t}$
- $u \xrightarrow{a} \overline{u}$
- $v \xrightarrow{\tau} \overline{u}\ _{p}\oplus\ \overline{t}$ for all $p \in (0, 1)$.

This is a finite-state but not finitely branching system, due to the infinite branch in v. We have that $s \lhd_{FS}^{k} \overline{v}$ for all $k \geq 0$ but we do not have $s \lhd_{FS}^{s} \overline{v}$.

We first observe that $s \lhd_{FS}^{s} \overline{v}$ does not hold because s will eventually deadlock with probability 1, whereas a fraction of v will go to u and never deadlock.

We now show that $s \lhd_{FS}^{k} \overline{v}$ for all $k \geq 0$. For any k we start the simulation by choosing the move $v \xrightarrow{\tau} (\overline{u}\ _{\frac{1}{2^k}}\oplus\ \overline{t})$. By induction on k we show that

$$s \lhd_{FS}^{k} (\overline{u}\ _{\frac{1}{2^k}}\oplus\ \overline{t}). \tag{24}$$

The base case $k = 0$ is trivial. So suppose we already have (24). We now show that $s \lhd_{FS}^{(k+1)} (\overline{u}\ _{\frac{1}{2^{k+1}}}\oplus\ \overline{t})$. Neither s nor t nor u can diverge or refuse $\{a\}$, so the only relevant move is the a-move. We know that s can do the move $s \xrightarrow{a} \overline{0}\ _{\frac{1}{2}}\oplus\ \overline{s}$. This can be matched up by $(\overline{u}\ _{\frac{1}{2^{k+1}}}\oplus\ \overline{t}) \xrightarrow{a} (\overline{0}\ _{\frac{1}{2}}\oplus\ (\overline{u}\ _{\frac{1}{2^k}}\oplus\ \overline{t}))$. □

Analogously to what we did for $\lhd_{FS}^{s}$, we also give an inductive characterisation of $\sqsupseteq_{FS}$: For every $k \geq 0$ let $\Delta \sqsupseteq_{FS}^{k} \Theta$ if there exists a $\Theta \Longrightarrow \Theta^{\text{match}}$ such that $\Delta\ \overline{\lhd_{FS}^{k}}\ \Theta^{\text{match}}$, and let $\sqsupseteq_{FS}^{\infty}$ denote $\bigcap_{k=0}^{\infty} \sqsupseteq_{FS}^{k}$.

Corollary 8.11 In a finitary pLTS, $\Delta \sqsupseteq_{FS} \Theta$ if and only if $\Delta \sqsupseteq_{FS}^{\infty} \Theta$.

Proof: Since $\lhd_{FS}^{s} \subseteq \lhd_{FS}^{k}$ for every $k \geq 0$, it is straightforward to prove one direction: $\Delta \sqsupseteq_{FS} \Theta$ implies $\Delta \sqsupseteq_{FS}^{\infty} \Theta$. For the converse, $\Delta \sqsupseteq_{FS}^{\infty} \Theta$ means that for every k we have some Θ^k satisfying $\Theta \Longrightarrow \Theta^k$ and $\Delta\ \overline{\lhd_{FS}^{k}}\ \Theta^k$. By Proposition 7.7 we have to find some Θ^∞ such that $\Theta \Longrightarrow \Theta^\infty$ and $\Delta\ \overline{\lhd_{FS}^{k}}\ \Theta^\infty$. This can be done exactly as in the proof of Theorem 8.9. □

8.2 The modal logic

Let $\mathscr{F}$ be the set of modal formulae defined inductively as follows:

- $\mathbf{div}, \top \in \mathscr{F}$
- $\mathbf{ref}(A) \in \mathscr{F}$ when $A \subseteq \mathsf{Act}$,
- $\langle a \rangle \varphi \in \mathscr{F}$ when $\varphi \in \mathscr{F}$ and $a \in \mathsf{Act}$,
- $\varphi_1 \wedge \varphi_2 \in \mathscr{F}$ when $\varphi_1,\ \varphi_2 \in \mathscr{F}$,
- $\varphi_1\ _p\oplus\ \varphi_2 \in \mathscr{F}$ when $\varphi_1, \varphi_2 \in \mathscr{F}$ and $p \in [0, 1]$.

This generalises the modal language used in [2] by the addition of the new constant $\mathbf{div}$, representing the ability of a process to diverge. In [2] there is the probabilistic choice operator $\bigoplus_{i \in I} p_i \cdot \varphi_i$, where I is a non-empty finite index set, and $\sum_{i \in I} p_i = 1$. This can be simulated in our language by nested use of the binary probabilistic choice.

Relative to a given pLTS $\langle S, \mathsf{Act}_\tau, \rightarrow \rangle$ the *satisfaction relation* $\models\ \subseteq \mathscr{D}(S) \times \mathscr{F}$ is given by:

- $\Delta \models \top$ for any $\Delta \in \mathscr{D}(S)$,
- $\Delta \models \mathbf{div}$ iff $\Delta \Longrightarrow \varepsilon$,
- $\Delta \models \mathbf{ref}(A)$ iff $\Delta \Longrightarrow \stackrel{A}{\not\longrightarrow}$,
- $\Delta \models \langle a \rangle \varphi$ iff there is a Δ' with $\Delta \stackrel{a}{\Longrightarrow} \Delta'$ and $\Delta' \models \varphi$,
- $\Delta \models \varphi_1 \wedge \varphi_2$ iff $\Delta \models \varphi_1$ and $\Delta \models \varphi_2$,

- $\Delta \models \varphi_1 \, {}_p\!\oplus \, \varphi_2$ iff there are $\Delta_1, \Delta_2 \in \mathscr{D}(S)$ with $\Delta_1 \models \varphi_1$ and $\Delta_2 \models \varphi_2$, such that $\Delta \Longrightarrow p \cdot \Delta_1 + (1-p) \cdot \Delta_2$.

We write $\Delta \sqsupseteq^{\mathscr{F}} \Theta$ when $\Delta \models \varphi$ implies $\Theta \models \varphi$ for all $\varphi \in \mathscr{F}$ — note the opposing directions. This is because the modal formulae express "bad" properties of our processes, ultimately divergence and refusal: thus $\Theta \sqsubseteq^{\mathscr{F}} \Delta$ means that any bad thing implementation Δ does must have been allowed by the specification Θ.

For pCSP processes we use $P \sqsubseteq^{\mathscr{F}} Q$ to abbreviate $[\![P]\!] \sqsubseteq^{\mathscr{F}} [\![Q]\!]$ in the pLTS given in Section 2.

The set of formulae used here is obtained from that in Section 7 of [2] by adding one operator, $\mathbf{div}$, and relaxing the constraint on the construction of probabilistic choice formulae. But the interpretation is quite different, as it uses the new silent move relation $\Longrightarrow$. As a result our satisfaction relation no longer enjoys a natural, and expected, property. In the non-probabilistic setting if a recursive CCS process P satisfies a modal formula from HML, then there is a recursion-free finite unwinding of P which also satisfies it. Intuitively this reflects the fact that if a non-probabilistic process does a bad thing, then at some (finite) point it must actually do it. But this is not true in our new, probabilistic setting: for example Q_1 given in Example 4.10 can do an a and then refuse anything; but all finite unwindings of it achieve that with probability strictly less than one. That is, whereas $[\![Q_1]\!] \models \langle a \rangle \top$, no finite unwinding of Q_1 will satisfy $\langle a \rangle \top$.

Our first task is to show that the interpretation of the logic is consistent with the operational semantics of processes.

Theorem 8.12 If $\Delta \sqsupseteq_{FS} \Theta$ then $\Delta \sqsupseteq^{\mathscr{F}} \Theta$.

Proof: We must show that if $\Delta \sqsupseteq_{FS} \Theta$ then whenever $\Delta \models \varphi$ we have $\Theta \models \varphi$. The proof proceeds by induction on φ:

- The case when $\varphi = \top$ is trivial.
- Suppose φ is $\mathbf{div}$. Then $\Delta \models \mathbf{div}$ means that $\Delta \Longrightarrow \varepsilon$ and we have to show $\Theta \Longrightarrow \varepsilon$, which is immediate from Lemma 7.3.
- Suppose φ is $\langle a \rangle \varphi_a$. In this case we have $\Delta \stackrel{a}{\Longrightarrow} \Delta'$ for some Δ' satisfying $\Delta' \models \varphi_a$. The existence of a corresponding Θ' is immediate from Definition 7.2 Case 1 and the induction hypothesis.
- The case when φ is $\mathbf{ref}(A)$ follows by Definition 7.2 Clause 2, and the case $\varphi_1 \wedge \varphi_2$ by induction.
- When φ is $\varphi_1 \, {}_p\!\oplus \, \varphi_2$ we appeal again to Definition 7.2 Case 1, using $\alpha := \tau$ to infer the existence of suitable Θ'_1 and Θ'_2. □

We proceed to show that the converse to this theorem also holds, so that the failure simulation preorder $\sqsubseteq_{FS}$ coincides with the logical preorder $\sqsubseteq^{\mathscr{F}}$.

The idea is to mimic the development in Section 7 of [2], by designing *characteristic formulae* which capture the behaviour of states in a pLTS. But here the behaviour is not characterised relative to $\lhd^{s}_{FS}$, but rather to the sequence of approximating relations $\lhd^{k}_{FS}$.

Definition 8.13 In a finitary pLTS $\langle S, \mathsf{Act}_\tau, \rightarrow \rangle$, the k^{th} *characteristic formulae* φ^k_s, φ^k_Δ of states $s \in S$ and subdistributions $\Delta \in \mathscr{D}(S)$ are defined inductively as follows:

- $\varphi^0_s = \top$ and $\varphi^0_\Delta = \top$,
- $\varphi^{k+1}_s = \mathbf{div}$, provided $\overline{s} \Longrightarrow \varepsilon$,
- $\varphi^{k+1}_s = \mathbf{ref}(A) \wedge \bigwedge_{s \xrightarrow{a} \Delta} \langle a \rangle \varphi^k_\Delta$ where $A = \{a \in \mathsf{Act} \mid s \stackrel{a}{\not\rightarrow}\}$, provided $s \stackrel{\tau}{\not\rightarrow}$,
- $\varphi^{k+1}_s = \bigwedge_{s \xrightarrow{a} \Delta} \langle a \rangle \varphi^k_\Delta \wedge \bigwedge_{s \xrightarrow{\tau} \Delta} \varphi^k_\Delta$ otherwise,
- and $\varphi^{k+1}_\Delta = (\mathbf{div}) \ {}_{1-|\Delta|}\!\oplus \ \left(\bigoplus_{s \in \lceil \Delta \rceil} \frac{\Delta(s)}{|\Delta|} \cdot \varphi^{k+1}_s \right)$.

Lemma 8.14 For every $k \geq 0$, $s \in S$ and $\Delta \in \mathscr{D}(S)$ we have $\overline{s} \models \varphi^k_s$ and $\Delta \models \varphi^k_\Delta$.

Proof: By induction on k, with the case when $k = 0$ being trivial. The inductive case of the first statement proceeds by an analysis of the possible moves from s, from which that of the second statement follows immediately. □

Lemma 8.15 For $k \geq 0$,

(i) $\Theta \models \varphi_s^k$ implies $s \lhd_{FS}^k \Theta$,

(ii) $\Theta \models \varphi_\Delta^k$ implies $\Theta \Longrightarrow \Theta^{\text{match}}$ such that $\Delta \ \overline{\lhd_{FS}^k}\ \Theta^{\text{match}}$,

(iii) $\Theta \models \varphi_\Delta^k$ implies $\Delta \sqsupseteq_{FS}^k \Theta$.

Proof: For every k part (iii) follows trivially from (ii). We prove (i) and (ii) simultaneously, by induction on k, with the case $k = 0$ being trivial. The inductive case, for $k + 1$, follows the argument in the proof of Lemma 7.3 of [2].

(i) First suppose $\overline{s} \Longrightarrow \varepsilon$. Then $\varphi_s^{k+1} = \mathbf{div}$ and therefore $\Theta \models \mathbf{div}$, which gives the required $\Theta \Longrightarrow \varepsilon$.

Now suppose $s \xrightarrow{\tau} \Delta$. Here there are two cases; if in addition $\overline{s} \Longrightarrow \varepsilon$ we have already seen that $\Theta \Longrightarrow \varepsilon$ and this is the required matching move from Θ, since $\Delta \ \overline{\lhd_{FS}^k}\ \varepsilon$. So let us assume that $\overline{s} \not\Longrightarrow \varepsilon$. Then by the definition of φ_s^{k+1} we must have that $\Theta \models \varphi_\Delta^k$, and we obtain the required matching move from Θ from the inductive hypothesis: induction on part (ii) gives some Θ' such that $\Theta \Longrightarrow \Theta'$ and $\Delta \ \overline{\lhd_{FS}^k}\ \Theta'$.

The matching move for $s \xrightarrow{a} \Theta$ is obtained in a similar manner.

Finally suppose $s \stackrel{A}{\not\longrightarrow}$. Since this implies $s \stackrel{\tau}{\not\longrightarrow}$, by the definition of φ_s^{k+1} we must have that $\Theta \models \mathbf{ref}(A)$, which actually means that $\Theta \Longrightarrow\stackrel{A}{\not\longrightarrow}$.

(ii) By definition $\varphi_\Delta^{k+1} = (\mathbf{div}) \ {}_{1-|\Delta|}\!\oplus \ (\bigoplus_{s\in\lceil\Delta\rceil} \frac{\Delta(s)}{|\Delta|}\cdot\varphi_s^{k+1})$ and thus $\Theta \Longrightarrow (1-|\Delta|)\cdot\Theta_{\mathbf{div}} + \sum_{s\in\lceil\Delta\rceil} \Delta(s)\cdot\Theta_s$ such that $\Theta_{\mathbf{div}} \models \mathbf{div}$ and $\Theta_s \models \varphi_s^{k+1}$. By definition, $\Theta_{\mathbf{div}} \Longrightarrow \varepsilon$, so by Theorem 3.20(i) and the reflexivity and transitivity of $\Longrightarrow$ we obtain $\Theta \Longrightarrow \sum_{s\in\lceil\Delta\rceil} \Delta(s)\cdot\Theta_s$. By part (i) we know that $s \lhd_{FS}^{k+1} \Theta_s$ for every s in $\lceil\Delta\rceil$, which in turn means that $\Delta \ \overline{\lhd_{FS}^{k+1}}\ \sum_{s\in\lceil\Delta\rceil} \Delta(s)\cdot\Theta_s$. □

Theorem 8.16 In a finitary pLTS, $\Delta \sqsupseteq^{\mathscr{F}} \Theta$ if and only if $\Delta \sqsupseteq_{FS} \Theta$.

Proof: One direction follows immediately from Theorem 8.12. For the opposite direction suppose $\Delta \sqsupseteq^{\mathscr{F}} \Theta$. By Lemma 8.14 we have $\Delta \models \varphi_\Delta^k$, and hence $\Theta \models \varphi_\Delta^k$, for all $k \geq 0$. By part (iii) of the previous lemma we thus know that $\Delta \sqsupseteq_{FS}^\infty \Theta$. That $\Delta \sqsupseteq_{FS} \Theta$ now follows from Corollary 8.11. □

8.3 Characteristic tests for formulae

The import of Theorem 8.16 is that we can obtain completeness of the failure simulation preorder with respect to the must-testing preorder by designing for each formula φ a test which in some sense characterises the property of a process of satisfying φ. This has been achieved for the pLTS generated by the recursion free fragment of pCSP in Section 8 of [2]. Here we generalise this technique to the pLTS generated by the set of finitary pCSP terms.

As in [2], the generation of these tests depends on crucial characteristics of the testing function $\mathscr{A}(-,-)$, which are summarised in Lemmas 8.17 and 8.20 below, corresponding to Lemmas 6.7 and 6.8 in [2] respectively.

Lemma 8.17 Let Δ be a pCSP process, and T, T_i be tests.

1. $o \in \mathscr{A}(\omega, \Delta)$ iff $o = |\Delta|\cdot\vec{\omega}$.

2. $\vec{0} \in \mathscr{A}(\tau.\omega, \Delta)$ iff $\Delta \Longrightarrow \varepsilon$.

3. $\vec{0} \in \mathscr{A}(\Box_{a\in A}\, a.\omega, \Delta)$ iff $\Delta \Longrightarrow\stackrel{A}{\not\longrightarrow}$.

4. Suppose the action ω does not occur in the test T. Then $o \in \mathscr{A}(\tau.\omega \,\Box\, a.T, \Delta)$ with $o(\omega) = 0$ iff there is a $\Delta' \in \mathscr{D}(\mathsf{sCSP})$ with $\Delta \stackrel{a}{\Longrightarrow} \Delta'$ and $o \in \mathscr{A}(T, \Delta')$.

5. $o \in \mathscr{A}(T_1 \,{}_p\!\oplus T_2, \Delta)$ iff $o = p\cdot o_1 + (1-p)\cdot o_2$ for certain $o_i \in \mathscr{A}(T_i, \Delta)$.

6. $o \in \mathscr{A}(T_1 \sqcap T_2, \Delta)$ if there are a $q \in [0,1]$ and $\Delta_1, \Delta_2 \in \mathscr{D}(\mathsf{sCSP})$ such that $\Delta \Longrightarrow q\cdot\Delta_1 + (1-q)\cdot\Delta_2$ and $o = q\cdot o_1 + (1-q)\cdot o_2$ for certain $o_i \in \mathscr{A}(T_i, \Delta_i)$.

Here $\vec{0}, \vec{\omega} \in [0,1]^\Omega$, with $\vec{0}(\omega) = 0$ for all $\omega \in \Omega$, and $\vec{\omega}(\omega) = 1$ but $\vec{\omega}(\omega') = 0$ whenever $\omega' \neq \omega$.

Proof:

1. Since $\omega \mid_{\mathsf{Act}} \Delta \xrightarrow{\omega}$, the states in the support of $[\omega \mid_{\mathsf{Act}} \Delta]$ have no other outgoing transitions than ω. Therefore $[\omega \mid_{\mathsf{Act}} \Delta]$ is the unique extreme derivative of itself, and as $\$[\omega \mid_{\mathsf{Act}} \Delta] = |\Delta|\cdot\vec{\omega}$ we have $\mathscr{A}(\omega,\Delta) = \{|\Delta|\cdot\vec{\omega}\}$.

2. ($\Leftarrow$) Assume $\Delta \Longrightarrow \varepsilon$. By Lemma 7.27(1) we have $\tau.\omega \mid_{\mathsf{Act}} \Delta \Longrightarrow \tau.\omega \mid_{\mathsf{Act}} \varepsilon$. All states involved in this derivation (that is, all states u in the support of the intermediate distributions $\Delta_i^{\rightarrow}$ and $\Delta_i^{\times}$ of Definition 3.13) have the form $\tau.\omega \mid_{\mathsf{Act}} s$, and thus satisfy $u \stackrel{\omega}{\not\longrightarrow}$ for all $\omega \in \Omega$. Therefore we have $[\tau.\omega \mid_{\mathsf{Act}} \Delta] \Longrightarrow [\tau.\omega \mid_{\mathsf{Act}} \varepsilon]$. Trivially, $[\tau.\omega \mid_{\mathsf{Act}} \varepsilon] = \varepsilon$ is stable, and hence an extreme derivative of $[\tau.\omega \mid_{\mathsf{Act}} \Delta]$. Moreover, $\$\varepsilon = \vec{0}$, so $\vec{0} \in \mathscr{A}(\tau.\omega,\Delta)$.

   ($\Rightarrow$) Suppose $\vec{0} \in \mathscr{A}(\tau.\omega,\Delta)$, i.e., there is some extreme derivative Γ of $[\tau.\omega \mid_{\mathsf{Act}} \Delta]$ such that $\$\Gamma = \vec{0}$. Given the operational semantics of pCSP, all states $u \in \lceil\Gamma\rceil$ must have one of the forms $u = [\tau.\omega \mid_{\mathsf{Act}} t]$ or $u = [\omega \mid_{\mathsf{Act}} t]$. As $\$\Gamma = \vec{0}$, the latter possibility cannot occur. It follows that all transitions contributing to the derivation $[\tau.\omega \mid_{\mathsf{Act}} \Delta] \Longrightarrow\!\!\!\!\succ \Gamma$ are obtained by means of the rule (PAR.R), and in fact Γ has the form $[\tau.\omega \mid_{\mathsf{Act}} \Delta']$ for some distribution Δ' with $\Delta \Longrightarrow \Delta'$. As Γ must be stable, yet none of the states in its support are, it follows that $\lceil\Gamma\rceil = \emptyset$, i.e. $\Delta' = \varepsilon$.

3. Let $T := \Box_{a\in A}\, a.\omega$.

   ($\Leftarrow$) Assume $\Delta \Longrightarrow \Delta' \stackrel{A}{\not\longrightarrow}$ for some Δ'. Then $T \mid_{\mathsf{Act}} \Delta \Longrightarrow T \mid_{\mathsf{Act}} \Delta'$ by Lemma 7.27(1), and by the same argument as in the previous case, $[T \mid_{\mathsf{Act}} \Delta] \Longrightarrow [T \mid_{\mathsf{Act}} \Delta']$. All states in the support of $T \mid_{\mathsf{Act}} \Delta'$ are deadlocked. So $[T \mid_{\mathsf{Act}} \Delta] \Longrightarrow\!\!\!\!\succ [T \mid_{\mathsf{Act}} \Delta]$ and $\$(T \mid_{\mathsf{Act}} \Delta) = \vec{0}$. Thus we have $\vec{0} \in \mathscr{A}(T,\Delta)$.

   ($\Rightarrow$) Suppose $\vec{0} \in \mathscr{A}(T,\Delta)$. By the very same reasoning as in Case 2 we find that $\Delta \Longrightarrow \Delta'$ for some Δ' such that $T \mid_{\mathsf{Act}} \Delta'$ is stable. This implies $\Delta' \stackrel{A}{\not\longrightarrow}$.

4. Let T be a test in which the success action ω does not occur, and let $U := \tau.\omega \Box a.T$.

   ($\Leftarrow$) Assume there is a $\Delta' \in \mathscr{D}(\mathsf{sCSP})$ with $\Delta \stackrel{a}{\Longrightarrow} \Delta'$ and $o \in \mathscr{A}(T,\Delta')$. Without loss of generality we may assume that $\Delta \Longrightarrow \Delta^{\mathrm{pre}} \xrightarrow{a} \Delta^{\mathrm{post}} \Longrightarrow \Delta'$. Using Lemma 7.27(1) and (3), and the same reasoning as in the previous cases, $[U \mid_{\mathsf{Act}} \Delta] \Longrightarrow [U \mid_{\mathsf{Act}} \Delta^{\mathrm{pre}}] \xrightarrow{\tau} [T \mid_{\mathsf{Act}} \Delta^{\mathrm{post}}] \Longrightarrow [T \mid_{\mathsf{Act}} \Delta'] \Longrightarrow \Gamma$ for a stable subdistribution Γ with $\$\Gamma = o$. It follows that $o \in \mathscr{A}(U,\Delta)$.

   ($\Rightarrow$) Suppose $o \in \mathscr{A}(U,\Delta)$ with $o(\omega) = 0$. Then there is a stable subdistribution Γ such that $[U \mid_{\mathsf{Act}} \Delta] \Longrightarrow \Gamma$ and $\$\Gamma = o$. Since $o(\omega) = 0$ there is no state in the support of Γ of the form $\omega \mid_{\mathsf{Act}} t$. Hence there must be a $\Delta' \in \mathscr{D}(\mathsf{sCSP})$ such that $\Delta \Longrightarrow\xrightarrow{a} \Delta'$ and $[T \mid_{\mathsf{Act}} \Delta'] \Longrightarrow \Gamma$. It follows that $o \in \mathscr{A}(T,\Delta')$.

5. ($\Leftarrow$) Assume $o_i \in \mathscr{A}(T_i,\Delta)$ for $i = 1,2$. Then $[T_i \mid_{\mathsf{Act}} \Delta] \Longrightarrow \Gamma_i$ for some stable Γ_i with $\$\Gamma_i = o_i$. By Theorem 3.20(i) we have $[(T_1\, {}_p\!\oplus T_2) \mid_{\mathsf{Act}} \Delta] = p\cdot[T_1 \mid_{\mathsf{Act}} \Delta] + (1{-}p)\cdot[T_2 \mid_{\mathsf{Act}} \Delta] \Longrightarrow p\cdot\Gamma_1 + (1{-}p)\cdot\Gamma_2$, and $p\cdot\Gamma_1 + (1{-}p)\cdot\Gamma_2$ is stable. Moreover, $\$(p\cdot\Gamma_1 + (1{-}p)\cdot\Gamma_2) = p\cdot o_1 + (1{-}p)\cdot o_2$, so $o \in \mathscr{A}(T_1\, {}_p\!\oplus T_2,\Delta)$.

   ($\Rightarrow$) Suppose $o \in \mathscr{A}(T_1\, {}_p\!\oplus T_2,\Delta)$. Then there is a stable Γ with $\$\Gamma = o$ such that $[(T_1\, {}_p\!\oplus T_2) \mid_{\mathsf{Act}} \Delta] = p\cdot[T_1 \mid_{\mathsf{Act}} \Delta] + (1{-}p)\cdot[T_2 \mid_{\mathsf{Act}} \Delta] \Longrightarrow \Gamma$. By Theorem 3.20(ii) there are Γ_i for $i = 1,2$, such that $[T_i \mid_{\mathsf{Act}} \Delta] \Longrightarrow \Gamma_i$ and $\Gamma = p\cdot\Gamma_1 + (1{-}p)\cdot\Gamma_2$. As Γ_1 and Γ_2 are stable, we have $\$\Gamma_i \in \mathscr{A}(T_i,\Delta)$ for $i = 1,2$. Moreover, $o = \$\Gamma = p\cdot\$\Gamma_1 + (1{-}p)\cdot\$\Gamma_2$.

6. Suppose $q \in [0,1]$ and $\Delta_1,\Delta_2 \in \mathscr{D}(\mathsf{pCSP})$ with $\Delta \Longrightarrow q\cdot\Delta_1 + (1{-}q)\cdot\Delta_2$ and $o_i \in \mathscr{A}(T_i,\Delta_i)$. Then there are stable Γ_i with $[T_i \mid_{\mathsf{Act}} \Delta_i] \Longrightarrow \Gamma_i$ and $\$\Gamma_i = o_i$. Now $[(T_1 \sqcap T_2) \mid_{\mathsf{Act}} \Delta] \Longrightarrow q\cdot[(T_1 \sqcap T_2) \mid_{\mathsf{Act}} \Delta_1] + (1{-}q)\cdot[(T_1 \sqcap T_2) \mid_{\mathsf{Act}} \Delta_2] \xrightarrow{\tau} q\cdot[T_1 \mid_{\mathsf{Act}} \Delta_1] + (1{-}q)\cdot[T_2 \mid_{\mathsf{Act}} \Delta_2] \Longrightarrow q\cdot\Gamma_1 + (1{-}q)\cdot\Gamma_2$. The latter subdistribution is stable and satisfies $\$(q\cdot\Gamma_1 + (1{-}q)\cdot\Gamma_2) = q\cdot o_1 + (1{-}q)\cdot o_2$. Hence $q\cdot o_1 + (1{-}q)\cdot o_2 \in \mathscr{A}(T_1 \sqcap T_2,\Delta)$. □

We also have the converse to part (6) of this lemma, again mimicking Lemma 6.8 of [2]. For that purpose, we use two technical lemmas whose proofs are similar to those for Lemmas 7.29 and 7.30 respectively.

Lemma 8.18 Suppose $\Delta \mid_A (T_1 \sqcap T_2) \xrightarrow{\tau} \Gamma$. Then there exist subdistributions $\Delta^{\rightarrow}$, $\Delta_1^{\times}$, $\Delta_2^{\times}$, Δ^{next} (possibly empty) such that

(i) $\Delta = \Delta^{\rightarrow} + \Delta_1^{\times} + \Delta_2^{\times}$

(ii) $\Delta^{\rightarrow} \xrightarrow{\tau} \Delta^{\mathrm{next}}$

(iii) $\Gamma = \Delta^{\text{next}} \mid_A (T_1 \sqcap T_2) + \Delta^\times_1 \mid_A T_1 + \Delta^\times_2 \mid_A T_2$

Proof: By Lemma 3.5 $\Delta \mid_A (T_1 \sqcap T_2) \xrightarrow{\tau} \Gamma$ implies that

$$\Delta = \sum_{i\in I} p_i \cdot \overline{s_i}, \qquad s_i \mid_A (T_1 \sqcap T_2) \xrightarrow{\tau} \Gamma_i, \qquad \Gamma = \sum_{i \in I} p_i \cdot \Gamma_i,$$

for certain $s_i \in S$, $\Gamma_i \in \mathscr{D}(\mathsf{sCSP})$ and $\sum_{i\in I} p_i \leq 1$. Let $J_1 = \{\, i \in I \mid \Gamma_i = s_i \mid_A T_1 \,\}$ and $J_2 = \{\, i \in I \mid \Gamma_i = s_i \mid_A T_2 \,\}$. Note that for each $i \in (I - J_1 - J_2)$ we have Γ_i in the form $\Gamma'_i \mid_A (T_1 \sqcap T_2)$, where $s_i \xrightarrow{\tau} \Gamma'_i$. Now let

$$\Delta^{\rightarrow} = \sum_{i\in(I-J_1-J_2)} p_i \cdot \overline{s_i}, \qquad \Delta^\times_k = \sum_{i \in J_k} p_i \cdot \overline{s_i}, \qquad \Delta^{\text{next}} = \sum_{i\in(I-J_1-J_2)} p_i \cdot \Gamma'_i \; .$$

where $k = 1, 2$. By construction (i) and (iii) are satisfied, and (ii) follows by property (2) of Definition 3.2. □

Lemma 8.19 If $\Delta \mid_A (T_1 \sqcap T_2) \Longrightarrow\!\!\!\!\succ \Psi$ then there are Φ_1 and Φ_2 such that

(i) $\Delta \Longrightarrow \Phi_1 + \Phi_2$

(ii) $\Phi_1 \mid_A T_1 + \Phi_2 \mid_A T_2 \Longrightarrow\!\!\!\!\succ \Psi$

Proof: Suppose $\Delta_0 \mid_A (T_1 \sqcap T_2) \Longrightarrow\!\!\!\!\succ \Psi$. We know from Definition 3.13 that there is a collection of subdistributions $\Psi_k, \Psi^{\rightarrow}_k, \Psi^\times_k$, for $k \geq 0$, satisfying the properties

$$\begin{array}{llrl}
\Delta_0 \mid_A (T_1 \sqcap T_2) & = & \Psi_0 = & \Psi^{\rightarrow}_0 + \Psi^\times_0 \\
\Psi^{\rightarrow}_0 & \xrightarrow{\tau} & \Psi_1 = & \Psi^{\rightarrow}_1 + \Psi^\times_1 \\
\vdots & & & \vdots \\
\Psi^{\rightarrow}_k & \xrightarrow{\tau} & \Psi_{k+1} = & \Psi^{\rightarrow}_{k+1} + \Psi^\times_{k+1} \\
 & & & \vdots \\
\hline
 & & & \Psi = \sum_{k=0}^{\infty} \Psi^\times_k
\end{array}$$

and Ψ is stable.

Take $\Gamma_0 := \Psi_0$. By induction on $k \geq 0$, we find distributions $\Gamma_{k+1},\ \Delta^{\rightarrow}_k,\ \Delta^\times_{k1},\ \Delta^\times_{k2},\ \Delta_{k+1}$ such that

(i) $\Delta_k \mid_A (T_1 \sqcap T_2) \xrightarrow{\tau} \Gamma_{k+1}$

(ii) $\Gamma_{k+1} \leq \Psi_{k+1}$

(iii) $\Delta_k = \Delta^{\rightarrow}_k + \Delta^\times_{k1} + \Delta^\times_{k2}$

(iv) $\Delta^{\rightarrow}_k \xrightarrow{\tau} \Delta_{k+1}$

(v) $\Gamma_{k+1} = \Delta_{k+1} \mid_A (T_1 \sqcap T_2) + \Delta^\times_{k1} \mid_A T_1 + \Delta^\times_{k2} \mid_A T_2$

Induction Step: Assume we already have Γ_k and Δ_k. Note that $\Delta_k \mid_A (T_1 \sqcap T_2) \leq \Gamma_k \leq \Psi_k = \Psi^{\rightarrow}_k + \Psi^\times_k$ and $T_1 \sqcap T_2$ can make a τ move. Since Ψ is stable, we know that either $\Psi^\times_k = \varepsilon$ or $\Psi^\times_k \stackrel{\tau}{\not\longrightarrow}$. In both cases it holds that $\Delta_k \mid_A (T_1 \sqcap T_2) \leq \Psi^{\rightarrow}_k$. Proposition 3.10 gives a subdistribution $\Gamma_{k+1} \leq \Psi_{k+1}$ such that $\Delta_k \mid_A (T_1 \sqcap T_2) \xrightarrow{\tau} \Gamma_{k+1}$. Now apply Lemma 8.18.

Let $\Phi_1 = \sum_{k=0}^{\infty} \Delta^\times_{k1}$ and $\Phi_2 = \sum_{k=0}^{\infty} \Delta^\times_{k2}$. By (iii) and (iv) above we obtain a weak τ move $\Delta \Longrightarrow \Phi_1 + \Phi_2$. For $k \geq 0$, let $\Gamma^{\rightarrow}_k := \Delta_k \mid_A (T_1 \sqcap T_2)$, let $\Gamma^\times_0 := \varepsilon$ and let $\Gamma^\times_{k+1} := \Delta^\times_{k1} \mid_A T_1 + \Delta^\times_{k2} \mid_A T_2$. Moreover, $\Gamma := \Phi_1 \mid_A T_1 + \Phi_2 \mid_A T_2$. Now all conditions of Definition 3.23 are fulfilled, so $\Delta_0 \mid_A (T_1 \sqcap T_2) \Longrightarrow \Gamma$ is an initial segment of $\Delta_0 \mid_A (T_1 \sqcap T_2) \Longrightarrow \Psi$. By Proposition 3.24 we have $\Phi_1 \mid_A T_1 + \Phi_2 \mid_A T_2 \Longrightarrow\!\!\!\!\succ \Psi$. □

Lemma 8.20 If $o \in \mathscr{A}(T_1 \sqcap T_2, \Delta)$ then there are a $q \in [0, 1]$ and $\Delta_1, \Delta_2 \in \mathscr{D}(\mathsf{sCSP})$ such that $\Delta \Longrightarrow q \cdot \Delta_1 + (1-q) \cdot \Delta_2$ and $o = q \cdot o_1 + (1-q) \cdot o_2$ for certain $o_i \in \mathscr{A}(T_i, \Delta_i)$.

Proof: If $o \in \mathscr{A}(T_1 \sqcap T_2, \Delta)$ then there is an extreme derivative Ψ of $[(T_1 \sqcap T_2) \mid_{\mathsf{Act}} \Delta]$ such that $\$\Psi = o$. By Lemma 8.19 there are $\Phi_{1,2}$ such that

(i) $\Delta \Longrightarrow \Phi_1 + \Phi_2$

(ii) and $[T_1 \mid_{\mathsf{Act}} \Phi_1] + [T_2 \mid_{\mathsf{Act}} \Phi_2] \Longrightarrow\!\!\!\!\succ \Psi$.

By Theorem 3.20(ii) there are some subdistributions Ψ_1 and Ψ_2 such that $\Psi = \Psi_1 + \Psi_2$ and $T_i \mid_{\mathsf{Act}} \Phi_i \Longrightarrow\!\!\!\!\succ \Psi_i$ for $i = 1, 2$. Let $o'_i = \$\Psi_i$. As Ψ_i is stable we obtain $o'_i \in \mathscr{A}(T_i, \Psi_i)$. We also have $o = \$\Psi = \$\Psi_1 + \$\Psi_2 = o'_1 + o'_2$.

We now distinguish two cases:

- If $\Psi_1 = \varepsilon$, then we take $\Delta_i = \Phi_i$, $o_i = o'_i$ for $i = 1, 2$ and $q = 0$. Symmetrically, if $\Psi_2 = \varepsilon$, then we take $\Delta_i = \Phi_i$, $o_i = o'_i$ for $i = 1, 2$ and $q = 1$.
- If $\Psi_1 \neq \varepsilon$ and $\Psi_2 \neq \varepsilon$, then we let $q = \frac{|\Phi_1|}{|\Phi_1 + \Phi_2|}$, $\Delta_1 = \frac{1}{q}\Phi_1$, $\Delta_2 = \frac{1}{1-q}\Phi_2$, $o_1 = \frac{1}{q}o'_1$ and $o_2 = \frac{1}{1-q}o'_2$.

It is easy to check that $q \cdot \Delta_1 + (1-q) \cdot \Delta_2 = \Phi_1 + \Phi_2$, $q \cdot o_1 + (1-q) \cdot o_2 = o'_1 + o'_2$ and $o_i \in \mathscr{A}(T_i, \Delta_i)$ for $i = 1, 2$. □

Proposition 8.21 For every formula $\varphi \in \mathscr{F}$ there exists a pair (T_φ, v_φ) with T_φ an Ω-test and $v_\varphi \in [0,1]^\Omega$ such that

$$\Delta \models \varphi \text{ if and only if } \exists o \in \mathscr{A}(T_\varphi, \Delta) : o \leq v_\varphi. \tag{25}$$

T_φ is called a *characteristic test* of φ and v_φ its *target value*.

Proof: The proof is adapted from that of Lemma 8.1 in [2], from where we take the following remarks: As in vector-based testing Ω is assumed to be countable (cf. page 17) and Ω-tests are finite expressions, for every Ω-test there is an $\omega \in \Omega$ not occurring in it. Furthermore, if a pair (T_φ, v_φ) satisfies requirement (25), then any pair obtained from (T_φ, v_φ) by bijectively renaming the elements of Ω also satisfies that requirement. Hence two given characteristic tests can be assumed to be Ω-*disjoint*, meaning that no $\omega \in \Omega$ occurs in both of them.

Our modal logic $\mathscr{F}$ is identical to that used in [2], with the addition of one extra constant $\mathbf{div}$. So we need a new characteristic test and target value for this latter formula, and reuse those from [2] for the rest of the language:[4]

- Let $\varphi = \top$. Take $T_\varphi := \omega$ for some $\omega \in \Omega$, and $v_\varphi := \vec{\omega}$.
- Let $\varphi = \mathbf{div}$. Take $T_\varphi := \tau.\omega$ for some $\omega \in \Omega$, and $v_\varphi := \vec{0}$.
- Let $\varphi = \mathbf{ref}(A)$ with $A \subseteq \mathsf{Act}$. Take $T_\varphi := \Box_{a \in A}\, a.\omega$ for some $\omega \in \Omega$, and $v_\varphi := \vec{0}$.
- Let $\varphi = \langle a \rangle \psi$. By induction, ψ has a characteristic test T_ψ with target value v_ψ. Take $T_\varphi := \tau.\omega \,\Box\, a.T_\psi$ where $\omega \in \Omega$ does not occur in T_ψ, and $v_\varphi := v_\psi$.
- Let $\varphi = \varphi_1 \wedge \varphi_2$. Choose an Ω-disjoint pair (T_i, v_i) of characteristic tests T_i with target values v_i, for $i = 1, 2$. Furthermore, let $p \in (0, 1]$ be chosen arbitrarily, and take $T_\varphi := T_1 \,{}_p\!\oplus T_2$ and $v_\varphi := p \cdot v_1 + (1-p) \cdot v_2$.
- Let $\varphi = \varphi_1 \,{}_p\!\oplus \varphi_2$. Again choose an Ω-disjoint pair (T_i, v_i) of characteristic tests T_i with target values v_i, $i = 1, 2$, this time ensuring that there are two distinct success actions $\omega_1,\ \omega_2$ that do not occur in any of these tests. Let $T'_i := T_i \,{}_{\frac{1}{2}}\!\oplus \omega_i$ and $v'_i := \frac{1}{2}v_i + \frac{1}{2}\vec{\omega}_i$. Note that for $i = 1, 2$ we have that T'_i is also a characteristic test of φ_i with target value v'_i. Take $T_\varphi := T'_1 \sqcap T'_2$ and $v_\varphi := p \cdot v'_1 + (1-p) \cdot v'_2$.

Note that $v_\varphi(\omega) = 0$ whenever $\omega \in \Omega$ does not occur in T_φ.

As in the proof of Lemma 8.1 of [2] we now check by induction on φ that (25) above holds; the proof relies on Lemmas 8.17 and 8.20.

- Let $\varphi = \top$. For all $\Delta \in \mathscr{D}(\mathsf{sCSP})$ we have $\Delta \models \varphi$ as well as $\exists o \in \mathscr{A}(T_\varphi, \Delta) : o \leq v_\varphi$, using Lemma 8.17(1).
- Let $\varphi = \mathbf{div}$. Suppose $\Delta \models \varphi$. Then we have that $\Delta \Longrightarrow \varepsilon$. By Lemma 8.17(2), $\vec{0} \in \mathscr{A}(T_\varphi, \Delta)$.
  Now suppose $\exists o \in \mathscr{A}(T_\varphi, \Delta) : o \leq v_\varphi$. This implies $o = \vec{0}$, so by Lemma 8.17(2), $\Delta \Longrightarrow \varepsilon$. Hence $\Delta \models \varphi$.
- Let $\varphi = \mathbf{ref}(A)$ with $A \subseteq \mathsf{Act}$. Suppose $\Delta \models \varphi$. Then $\Delta \Longrightarrow \xrightarrow{A}\!\!\!\!\!\not\,$. By Lemma 8.17(3), $\vec{0} \in \mathscr{A}(T_\varphi, \Delta)$.
  Now suppose $\exists o \in \mathscr{A}(T_\varphi, \Delta) : o \leq v_\varphi$. This implies $o = \vec{0}$, so $\Delta \Longrightarrow \xrightarrow{A}\!\!\!\!\!\not\,$ by Lemma 8.17(3). Hence $\Delta \models \varphi$.
- Let $\varphi = \langle a \rangle \psi$ with $a \in \mathsf{Act}$. Suppose $\Delta \models \varphi$. Then there is a Δ' with $\Delta \stackrel{a}{\Longrightarrow} \Delta'$ and $\Delta' \models \psi$. By induction, $\exists o \in \mathscr{A}(T_\psi, \Delta') : o \leq v_\psi$. By Lemma 8.17(4), $o \in \mathscr{A}(T_\varphi, \Delta)$.

[4]However, because we employ state-based testing here, as opposed to action-based testing in [2], we translate the action-based test $\omega \,\Box\, a.T_\psi$ for the action modality $\langle a \rangle \psi$ into the state-based test $\tau.\omega \,\Box\, a.T_\psi$.

Now suppose $\exists o \in \mathscr{A}(T_\varphi, \Delta) : o \le v_\varphi$. This implies $o(\omega) = 0$, so by Lemma 8.17(4) there is a Δ' with $\Delta \stackrel{a}{\Longrightarrow} \Delta'$ and $o \in \mathscr{A}(T_\psi, \Delta')$. By induction, $\Delta' \models \psi$, so $\Delta \models \varphi$.

- Let $\varphi = \varphi_1 \wedge \varphi_2$ and suppose $\Delta \models \varphi$. Then $\Delta \models \varphi_i$ for $i{=}1,2$ and hence, by induction, $\exists o_i \in \mathscr{A}(T_i, \Delta) : o_i \le v_i$. Thus $o := p \cdot o_1 + (1{-}p) \cdot o_2 \in \mathscr{A}(T_\varphi, \Delta)$ by Lemma 8.17(5), and $o \le v_\varphi$.

  Now suppose $\exists o \in \mathscr{A}(T_\varphi, \Delta) : o \le v_\varphi$. Then, using Lemma 8.17(5), $o = p \cdot o_1 + (1{-}p) \cdot o_2$ for certain $o_i \in \mathscr{A}(T_i, \Delta)$. Recall that T_1, T_2 are Ω-disjoint tests. One has $o_i \le v_i$ for both $i = 1, 2$, for if $o_i(\omega) > v_i(\omega)$ for some $i = 1$ or 2 and $\omega \in \Omega$, then ω must occur in T_i and hence cannot occur in T_{3-i}. This implies $v_{3-i}(\omega) = 0$ and thus $o(\omega) > v_\varphi(\omega)$, in contradiction with the assumption. By induction, $\Delta \models \varphi_i$ for $i = 1, 2$, and hence $\Delta \models \varphi$.

- Let $\varphi = \varphi_1 \,{}_p\!\oplus\, \varphi_2$. Suppose $\Delta \models \varphi$. Then there are $\Delta_1, \Delta_2 \in \mathscr{D}(\mathsf{sCSP})$ with $\Delta_1 \models \varphi_1$ and $\Delta_2 \models \varphi_2$ such that $\Delta \Longrightarrow p \cdot \Delta_1 + (1{-}p) \cdot \Delta_2$. By induction, for $i = 1, 2$ there are $o_i \in \mathscr{A}(T_i, \Delta_i)$ with $o_i \le v_i$. Hence, there are $o'_i \in \mathscr{A}(T'_i, \Delta_i)$ with $o'_i \le v'_i$. Thus $o := p \cdot o'_1 + (1{-}p) \cdot o'_2 \in \mathscr{A}(T_\varphi, \Delta)$ by Lemma 8.17(6), and $o \le v_\varphi$.

  Now suppose $\exists o \in \mathscr{A}(T_\varphi, \Delta) : o \le v_\varphi$. Then, by Lemma 8.20, there are $q \in [0, 1]$ and $\Delta_1, \Delta_2 \in \mathscr{D}(\mathsf{sCSP})$ such that $\Delta \Longrightarrow q \cdot \Delta_1 + (1{-}q) \cdot \Delta_2$ and $o = q \cdot o'_1 + (1{-}q) \cdot o'_2$ for certain $o'_i \in \mathscr{A}(T'_i, \Delta_i)$. Now $\forall i : o'_i(\omega_i) = v'_i(\omega_i) = \frac{1}{2}$, so, using that T_1, T_2 are Ω-disjoint tests, $\frac{1}{2}q = q \cdot o'_1(\omega_1) = o(\omega_1) \le v_\varphi(\omega_1) = p \cdot v'_1(\omega_1) = \frac{1}{2}p$ and likewise $\frac{1}{2}(1{-}q) = (1{-}q) \cdot o'_2(\omega_2) = o(\omega_2) \le v_\varphi(\omega_2) = (1{-}p) \cdot v'_2(\omega_2) = \frac{1}{2}(1{-}p)$. Together, these inequalities say that $q = p$. Exactly as in the previous case one obtains $o'_i \le v'_i$ for both $i = 1, 2$. Given that $T'_i = T_i \,{}_{\frac{1}{2}}\!\oplus\, \omega_i$, using Lemma 8.17(5), it must be that $o'_i = \frac{1}{2}o_i + \frac{1}{2}\vec{\omega}_i$ for some $o_i \in \mathscr{A}(T_i, \Delta_i)$ with $o_i \le v_i$. By induction, $\Delta_i \models \varphi_i$ for $i = 1, 2$, and hence $\Delta \models \varphi$. □

Theorem 8.22 If $\Delta \sqsupseteq^{\Omega}_{\text{pmust}} \Theta$ then $\Delta \sqsupseteq^{\mathscr{F}} \Theta$.

Proof: Suppose $\Delta \sqsupseteq^{\Omega}_{\text{pmust}} \Theta$ and $\Delta \models \varphi$ for some $\varphi \in \mathscr{F}$. Let T_φ be a characteristic test of φ with target value v_φ. Then Proposition 8.21 yields $\exists o \in \mathscr{A}(T_\varphi, \Delta) : o \le v_\varphi$, and hence, given that $\Delta \sqsupseteq^{\Omega}_{\text{pmust}} \Theta$, by the Smyth preorder we have $\exists o' \in \mathscr{A}(T_\varphi, \Theta) : o' \le v_\varphi$. Thus $\Theta \models \varphi$. □

Corollary 8.23 For any finitary processes P and Q, if $P \sqsubseteq_{\text{must}} Q$ then $P \sqsubseteq_{FS} Q$.

Proof: From Theorems 8.22 and 8.16 we know that if $P \sqsubseteq^{\Omega}_{\text{pmust}} Q$ then $P \sqsubseteq_{FS} Q$. Theorem B.4 from Section B.1 tells us that Ω-testing is reducible to scalar testing. So the required result follows. □

9 Simulations and may testing

In this section we follow the same strategy as for failure simulations and testing (Section 7) except that we restrict our treatment to full distributions: this is possible because partial distributions are not necessary for this case; and it is desirable because the approach becomes simpler as a result.

Definition 9.1 [Simulation Preorder] Define $\sqsubseteq_S$ to be the largest relation in $\mathscr{D}_1(S) \times \mathscr{D}_1(S)$ such that if $\Delta \sqsubseteq_S \Theta$ then

> whenever $\Delta \stackrel{\alpha}{\Longrightarrow} (\sum_i p_i \Delta'_i)$, for finitely many p_i with $\sum_i p_i = 1$, there are Θ'_i with $\Theta \stackrel{\alpha}{\Longrightarrow} (\sum_i p_i \Theta'_i)$ and $\Delta'_i \sqsubseteq_S \Theta'_i$ for each i.

Note that, unlike for Definition 9.1, this summation cannot be empty.

Again it is trivial to see that $\sqsubseteq_S$ is reflexive and transitive; and again it is sometimes easier to work with an equivalent formulation based on a state-level "simulation" defined as follows.

Definition 9.2 [Simulation] Define $\lhd_s$ to be the largest relation in $S \times \mathscr{D}_1(S)$ such that if $s \lhd_s \Theta$ then whenever $s \stackrel{\alpha}{\longrightarrow} \Delta'$ there is a Θ' with $\Theta \stackrel{\alpha}{\Longrightarrow} \Theta'$ and $\Delta' \;\overline{\lhd_s}\; \Theta'$.

Definition 9.2 differs from the analogous Definition 7.18 in three ways: it is missing the clause for divergence, and for refusal; and it is (implicitly) limited to $\stackrel{\alpha}{\Longrightarrow}$-transitions that simulate by producing full distributions only[5]. Without that latter limitation, any simulation relation could be scaled down uniformly without losing its simulation properties, for example allowing counter-intuitively a to be simulated by $a\ _{\frac{1}{2}}\oplus\ \varepsilon$.

Lemma 9.3 The above preorder and simulation are equivalent in the following sense: for distributions Δ, Θ we have $\Delta \sqsubseteq_S \Theta$ just when there is a Θ^{match} with $\Theta \Longrightarrow \Theta^{\text{match}}$ and $\Delta\ \overline{\lhd_s}\ \Theta^{\text{match}}$.

Proof: The proof is as for the failure case, except that in Theorem 7.20 we can assume total distributions, and so do not need the second part of its proof where divergence is treated. □

9.1 Soundness

In this section we prove that simulations are sound for showing that processes are related via the may-testing preorder. We assume initially that we are using only one success action ω, so that $|\Omega| = 1$.

Because we prune our pLTSs before extracting values from them, we will be concerned mainly with ω-respecting structures, and for those we have the following.

Lemma 9.4 Let Δ and Θ be two distributions. If Δ is stable and $\Delta\ \overline{\lhd_s}\ \Theta$, then $\mathscr{V}(\Delta) \leq_{\text{Ho}} \mathscr{V}(\Theta)$.

Proof: We first show that if s is stable and $s \lhd_s \Theta$ then $\mathscr{V}(s) \leq_{\text{Ho}} \mathscr{V}(\Theta)$. Since s is stable, we have only two cases:

(i) $s \not\longrightarrow$ Here $\mathscr{V}(s)=\{0\}$ and since $\mathscr{V}(\Theta)$ is not empty we have $\mathscr{V}(s) \leq_{\text{Ho}} \mathscr{V}(\Theta)$.

(ii) $s \stackrel{\omega}{\longrightarrow} \Delta'$ for some Δ' Here $\mathscr{V}(s)=\{1\}$ and $\Theta \Longrightarrow \Theta' \stackrel{\omega}{\longrightarrow}$ with $\mathscr{V}(\Theta')=\{1\}$. By Lemma 7.37 specialised to full distributions, we have $1 \in \mathscr{V}(\Theta)$. Therefore, $\mathscr{V}(s) \leq_{\text{Ho}} \mathscr{V}(\Theta)$.

Now for the general case we suppose $\Delta\ \overline{\lhd_s}\ \Theta$. Use Proposition 3.10 to decompose Θ into $\sum_{s\in\lceil\Delta\rceil} \Delta(s)\cdot\Theta_s$ such that $s \lhd_s \Theta_s$ for each $s \in \lceil\Delta\rceil$, and recall each such state s is stable. From above we have that $\mathscr{V}(s) \leq_{\text{Ho}} \mathscr{V}(\Theta_s)$ for those s, and so $\mathscr{V}(\Delta) = \sum_{\in\lceil\Delta\rceil} \Delta(s)\cdot\mathscr{V}(s) \leq_{\text{Ho}} \sum_{s\in\lceil\Delta\rceil} \Delta(s)\cdot\mathscr{V}(\Theta_s) = \mathscr{V}(\Theta)$. □

Lemma 9.5 Let Δ and Θ be distributions in an ω-respecting finitary pLTS $\langle S, \{\tau,\omega\}, \rightarrow\rangle$. If $\Delta\ \overline{\lhd_s}\ \Theta$, then we have $\mathscr{V}(\Delta) \leq_{\text{Ho}} \mathscr{V}(\Theta)$.

Proof: Since $\Delta\ \overline{\lhd_s}\ \Theta$, we consider subdistributions Δ'' with $\Delta \Longrightarrow\!\!\!\succ \Delta''$; by distillation of divergence (Theorem 6.20) we have full distributions Δ', Δ'_1 and Δ'_2 and probability p such that $s \Longrightarrow \Delta' = (\Delta'_1\ {}_p\oplus\ \Delta'_2)$ and $\Delta'' = p\cdot\Delta'_1$ and $\Delta'_2 \Longrightarrow \varepsilon$. There is thus a matching transition $\Theta \Longrightarrow \Theta'$ such that $\Delta'\ \overline{\lhd_s}\ \Theta'$. By Proposition 3.10, we can find distributions Θ'_1, Θ'_2 such that $\Theta' = \Theta'_1\ {}_p\oplus\ \Theta'_2$, $\Delta'_1\ \overline{\lhd_s}\ \Theta'_1$ and $\Delta'_2\ \overline{\lhd_s}\ \Theta'_2$.

Since $\lceil\Delta'_1\rceil = \lceil\Delta''\rceil$ we have that Δ'_1 is stable. It follows from Lemma 9.4 that $\mathscr{V}(\Delta'_1) \leq_{\text{Ho}} \mathscr{V}(\Theta'_1)$. Thus we finish off with

$$
\begin{array}{lll}
 & \mathscr{V}(\Delta'') & \\
= & \mathscr{V}(p\cdot\Delta'_1) & \Delta'' = p\cdot\Delta'_1 \\
= & p\cdot\mathscr{V}(\Delta'_1) & \text{linearity of } \mathscr{V} \\
\leq_{\text{Ho}} & p\cdot\mathscr{V}(\Theta'_1) & \text{above argument based on distillation} \\
= & \mathscr{V}(p\cdot\Theta'_1) & \text{linearity of } \mathscr{V} \\
\leq_{\text{Ho}} & \mathscr{V}(\Theta') & \Theta' = \Theta'_1\ {}_p\oplus\ \Theta'_2 \\
\leq_{\text{Ho}} & \mathscr{V}(\Theta)\ . & \text{Lemma 7.37 specialised to full distributions}
\end{array}
$$

Since Δ'' was arbitrary, we have our result. □

Lemma 9.6 Let Δ and Θ be distributions in an ω-respecting finitary pLTS $\langle S, \{\tau,\omega\}, \rightarrow\rangle$. If $\Delta \sqsubseteq_S \Theta$, then it holds that $\mathscr{V}(\Delta) \leq_{\text{Ho}} \mathscr{V}(\Theta)$.

[5] Even though for simplicity of presentation in Definition 3.2 the relation $\Longrightarrow$ was defined by using subdistributions, it can be equivalently defined by using full distributions.

Proof: Suppose $\Delta \sqsubseteq_S \Theta$. By Lemma 9.3, there exists some Θ^{match} such that $\Theta \Longrightarrow \Theta^{\text{match}}$ and $\Delta \mathrel{\overline{\lhd_s}} \Theta^{\text{match}}$. By Lemmas 9.5 and 7.37 we obtain $\mathscr{V}(\Delta) \leq_{\text{Ho}} \mathscr{V}(\Theta') \subseteq \mathscr{V}(\Theta)$. □

Theorem 9.7 For any finitary processes P and Q, if $P \sqsubseteq_S Q$ then $P \sqsubseteq_{\text{pmay}} Q$.

Proof: We reason as follows.

	$P \sqsubseteq_S Q$	
implies	$[P \mid_{\mathsf{Act}} T] \sqsubseteq_S [Q \mid_{\mathsf{Act}} T]$	the counterpart of Lemma 7.34 for simulation, for any test T
implies	$\mathscr{V}([P \mid_{\mathsf{Act}} T]) \leq_{\text{Ho}} \mathscr{V}([Q \mid_{\mathsf{Act}} T])$	$[\cdot]$ is ω-respecting; Lemma 9.6
iff	$\mathscr{A}(T,P) \leq_{\text{Ho}} \mathscr{A}(T,Q)$	Definition 7.35
iff	$P \sqsubseteq_{\text{pmay}} Q$.	Definition 4.9

□

9.2 Completeness

Let $\mathscr{L}$ be the subclass of $\mathscr{F}$ by skipping the $\mathbf{div}$ and $\mathbf{ref}(A)$ clauses. We write $P \sqsubseteq^{\mathscr{L}} Q$ just when $[\![P]\!] \models \varphi$ implies $[\![Q]\!] \models \varphi$. We have the counterparts of Theorems 8.16 and 8.22, with similar proofs.

Theorem 9.8 In a finitary pLTS $\Delta \sqsubseteq^{\mathscr{L}} \Theta$ if and only if $\Delta \sqsubseteq_S \Theta$.

Theorem 9.9 For any pCSP processes P and Q, if $P \sqsubseteq^{\Omega}_{\text{pmay}} Q$ then $P \sqsubseteq^{\mathscr{L}} Q$.

Corollary 9.10 Suppose P and Q are finitary pCSP processes. If $P \sqsubseteq_{\text{pmay}} Q$ then $P \sqsubseteq_S Q$.

Proof: From Theorems 9.8 and 9.9 we know that if $P \sqsubseteq^{\Omega}_{\text{pmay}} Q$ then $P \sqsubseteq_S Q$. Theorem B.4 from Section B.1 says that Ω-testing is reducible to scalar testing. So the required result follows. □

As one would expect, the completeness result in Corollary 9.10 would fail for infinitary processes.

Example 9.11 Consider the state s_2 which we saw in Example 3.17. It turns out that

$$\tau.(\mathbf{0} \;{}_{\frac{1}{2}}\!\oplus a) \sqsubseteq_{\text{pmay}} s_2$$

However, we do not have

$$\tau.(\mathbf{0} \;{}_{\frac{1}{2}}\!\oplus a) \lhd_s \overline{s_2}$$

because the transition

$$\tau.(\mathbf{0} \;{}_{\frac{1}{2}}\!\oplus a) \xrightarrow{\tau} (\mathbf{0} \;{}_{\frac{1}{2}}\!\oplus a)$$

cannot be matched by a transition from s_2 as there is no *full distribution* Δ such that $s_2 \Longrightarrow \Delta$ and $(\mathbf{0} \;{}_{\frac{1}{2}}\!\oplus a) \mathrel{\overline{\lhd_s}} \Delta$.
□

10 Conclusion and related work

In this paper we continued our previous work [4, 5, 2] in our quest for a testing theory for processes which exhibit both nondeterministic and probabilistic behaviour. We have generalised our results in [2] of characterising the may preorder as a simulation relation and the must preorder as a failure-simulation relation, from finite processes to finitary processes. Although the general proof schema is inherited from [2], the details here are much more complicated. One important reason is the inapplicability of structural induction, an important proof principle used in proving some fundamental properties for finite processes, when we shift to finitary processes. So we have to make use of more advanced mathematical tools such as fixed points on complete lattices, compact sets in topological spaces, especially in complete metric spaces, etc. Technically, we develop weak transitions between probabilistic processes, elaborate their topological properties, and capture divergence in terms of partial distributions. In order to obtain the characterisation results of testing preorders as simulation relations, we found it necessary to investigate fundamental structural properties of derivation sets (finite generability) and similarities (infinite approximations), which are of independent interest. The use of Markov Decision Processes and Zero-One laws was essential in obtaining our results.

There is a great amount of work about probabilistic testing semantics and simulation semantics. Here we mention the closely related work [24], where Segala defined two preorders called trace distribution precongruence ($\sqsubseteq_{TD}$) and failure distribution precongruence ($\sqsubseteq_{FD}$). He proved that the former coincides with an action-based version of $\sqsubseteq^{\Omega}_{\text{pmay}}$ and that for "probabilistically convergent" systems the latter coincides with an action-based version of $\sqsubseteq^{\Omega}_{\text{pmust}}$. The condition of probabilistic convergence amounts in our framework to the requirement that for $\Delta \in \mathscr{D}_1(S)$ and $\Delta \Longrightarrow \Delta'$ we have $|\Delta'| = 1$. In [17] it has been shown that $\sqsubseteq_{TD}$ coincides with a notion of simulation akin to $\sqsubseteq_S$. Other probabilistic extensions of simulation and testing approaches occurring in the literature are reviewed in [4, 2].

References

[1] S.D. Brookes, C.A.R. Hoare & A.W. Roscoe (1984): *A theory of communicating sequential processes. Journal of the ACM* 31(3), pp. 560–599.

[2] Y. Deng, R.J. van Glabbeek, M. Hennessy & C.C. Morgan (2008): *Characterising testing preorders for finite probabilistic processes. Logical Methods in Computer Science* 4(4:4).

[3] Y. Deng, R.J. van Glabbeek, M. Hennessy & C.C. Morgan (2011): *Real-reward testing for probabilistic processes (extended abstract).* Proc. *QAPL'11.*

[4] Y. Deng, R.J. van Glabbeek, M. Hennessy, C.C. Morgan & C. Zhang (2007): *Remarks on testing probabilistic processes. ENTCS* 172, pp. 359–397.

[5] Y. Deng, R.J. van Glabbeek, C.C. Morgan & C. Zhang (2007): *Scalar outcomes suffice for finitary probabilistic testing.* In Proc. *ESOP'07*, LNCS 4421, Springer, pp. 363–378.

[6] R. De Nicola & M. Hennessy (1984): *Testing equivalences for processes. Theoretical Computer Science* 34, pp. 83–133.

[7] R.J. van Glabbeek (1993): *The linear time – branching time spectrum II; the semantics of sequential systems with silent moves.* In Proc. *CONCUR'93*, LNCS 715, Springer, pp. 66–81.

[8] M. Hennessy (1988): *An Algebraic Theory of Processes*. MIT Press.

[9] M. Hennessy & R. Milner (1985): *Algebraic Laws for Nondeterminism and Concurrency. Journal of the ACM* 32(1), pp. 137–161.

[10] C.A.R. Hoare (1985): *Communicating Sequential Processes*. Prentice-Hall.

[11] C. Jones (1990) *Probabilistic Non-determinism.* Ph.D. Thesis, University of Edinburgh.

[12] B. Jonsson, C. Ho-Stuart & Wang Yi (1994): *Testing and refinement for nondeterministic and probabilistic processes.* In Proc. *FTRTFT'94*, LNCS 863, Springer, pp. 418–430.

[13] B. Jonsson & Wang Yi (1995): *Compositional testing preorders for probabilistic processes.* In Proc. *LICS'95*, IEEE Computer Society Press, pp. 431–441.

[14] B. Jonsson & Wang Yi (2002): *Testing preorders for probabilistic processes can be characterized by simulations. Theoretical Computer Science* 282(1), pp. 33–51.

[15] J. Rutten, M. Kwiatkowska, G. Norman & D. Parker (2004): *Mathematical Techniques for Analyzing Concurrent and Probabilistic Systems,* P. Panangaden and F. van Breugel (eds.), *CRM Monograph Series* 23. American Mathematical Society.

[16] S. Lipschutz (1965): *Schaum's outline of theory and problems of general topology.* McGraw-Hill.

[17] N. Lynch, R. Segala & F.W. Vaandrager (2007): *Observing Branching Structure through Probabilistic Contexts. SIAM Journal on Computing* 37(4), pp. 977–1013.

[18] J. Matoušek (2002): *Lectures on Discrete Geometry*. Springer.

[19] A.K. McIver & C.C. Morgan (2005): *Abstraction, Refinement and Proof for Probabilistic Systems*. Springer.

[20] R. Milner (1989): *Communication and Concurrency*. Prentice-Hall.

[21] E.-R. Olderog & C.A.R. Hoare (1986): *Specification-oriented semantics for communicating processes. Acta Informatica* 23, pp. 9–66.

[22] M. Puterman (1994): *Markov Decision Processes*. Wiley.

[23] R. Segala (1995): *Modeling and Verification of Randomized Distributed Real-Time Systems*. PhD thesis, MIT.

[24] R. Segala (1996): *Testing probabilistic automata.* In Proc. *CONCUR'96*, LNCS 1119, Springer, pp. 299–314.

[25] Wang Yi & K.G. Larsen (1992): *Testing probabilistic and nondeterministic processes.* In Proc. *PSTV'92, IFIP Transactions* C-8, North-Holland, pp. 47–61.

A Further properties of weak derivation

In this section we expose some less obvious properties of derivations, relating to their behaviour at infinity. One important property is that if we associate each state with a weight, which is a value in $[-1, 1]$, then the maximum payoff realisable by following all possible weak derivations can in fact be achieved by some static derivative policy. The property depends on our working within *finitary* pLTSs — that is, ones in which the state space is finite and the (unlifted) transition relation is finite-branching.

It turns out that to prove this property we need a notion of *bounded continuity* of real functions, which we introduce below.

A.1 Bounded continuity

Proposition A.1 (Bounded continuity - nonnegative function) Given a function $f : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{R}_{\geq 0}$ satisfying the conditions

- **C1.** f is monotonic in the second parameter, i.e. $j_1 \leq j_2$ implies $f(i, j_1) \leq f(i, j_2)$ for all $i, j_1, j_2 \in \mathbb{N}$;
- **C2.** for any $i \in \mathbb{N}$, the limit $\lim_{j\rightarrow\infty} f(i, j)$ exists;
- **C3.** the partial sums $S_n = \sum_{i=0}^{n} \lim_{j\rightarrow\infty} f(i, j)$ are bounded, i.e. there exists some $c \in \mathbb{R}_{\geq 0}$ such that $S_n \leq c$ for all $n \geq 0$;

then it holds that

$$\sum_{i=0}^{\infty} \lim_{j\rightarrow\infty} f(i, j) \ = \ \lim_{j\rightarrow\infty} \sum_{i=0}^{\infty} f(i, j).$$

Proof: First we show that, for any given $n \in \mathbb{N}$:

$$\sum_{i=0}^{n} \lim_{j\rightarrow\infty} f(i, j) \ = \ \lim_{j\rightarrow\infty} \sum_{i=0}^{n} f(i, j). \tag{26}$$

Let ϵ be any positive real number. Then, for any $i = 0, \ldots, n$, by **C1** and **C2** the sequence $\{f(i, j)\}_{j=0}^{\infty}$ is nondecreasing and converges to $\lim_{j\rightarrow\infty} f(i, j)$, so there must be a k_ϵ^i such that for all $j \geq k_\epsilon^i$

$$0 \leq \lim_{j\rightarrow\infty} f(i, j) - f(i, j) \leq \frac{\epsilon}{n+1} \ .$$

Let $k_\epsilon := \max\{k_\epsilon^i \mid 0 \leq i \leq n\}$. Then, for all $j \geq k_\epsilon$

$$0 \leq \left(\sum_{i=0}^{n} \lim_{j\rightarrow\infty} f(i, j)\right) - \left(\sum_{i=0}^{n} f(i, j)\right) = \sum_{i=0}^{n} \left(\lim_{j\rightarrow\infty} f(i, j) - f(i, j)\right) \leq \sum_{i=0}^{n} \frac{\epsilon}{n+1} = \epsilon \ .$$

Since $\{\sum_{i=0}^{n} f(i, j)\}_{j=0}^{\infty}$ is a nondecreasing sequence, this yields (26).

By **C3** the sequence $\{S_n\}_{n=0}^{\infty}$ is bounded. Since it is also nondecreasing, it converges to $\ell := \sum_{i=0}^{\infty} \lim_{j\rightarrow\infty} f(i, j)$. Hence the left-hand side of the desired equality exists. For any $i, j \in \mathbb{N}$ we have $f(i, j) \leq \lim_{j\rightarrow\infty} f(i, j)$, so $\sum_{i=0}^{n} f(i, j) \leq S_n \leq \ell$. Since also the sequence $\{\sum_{i=0}^{n} f(i, j)\}_{n=0}^{\infty}$ is nondecreasing, $\sum_{i=0}^{\infty} f(i, j) = \lim_{n\rightarrow\infty} \sum_{i=0}^{n} f(i, j)$ exists and is bounded by ℓ. By **C1** we have that $j_1 \leq j_2$ implies $\sum_{i=0}^{\infty} f(i, j_1) \leq \sum_{i=0}^{\infty} f(i, j_2)$. So also $r := \lim_{j\rightarrow\infty} \sum_{i=0}^{\infty} f(i, j)$ exists and is bounded by ℓ. It remains to show that $\ell \leq r$. For any $j, n \in \mathbb{N}$ we have $\sum_{i=0}^{n} f(i, j) < \sum_{i=0}^{\infty} f(i, j) < r$. Hence $\lim_{j\rightarrow\infty} \sum_{i=0}^{n} f(i, j)$ exists and is bounded by r. By (26) this gives $S_n \leq r$ for any $n \in \mathbb{N}$. Thus $\ell = \lim_{n\rightarrow\infty} S_n \leq r$. □

Proposition A.2 [Bounded continuity - general function] Given a function $f : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{R}$ which satisfies the following conditions

- **C1.** For all $i, j_1, j_2 \in \mathbb{N}$, we have $j_1 \leq j_2$ implies $|f(i, j_1)| \leq |f(i, j_2)|$.
- **C2.** For any $i \in \mathbb{N}$, the limit $\lim_{j\rightarrow\infty} |f(i, j)|$ exists.
- **C3.** For any $n \in \mathbb{N}$, the partial sum $S_n = \sum_{i=0}^{n} \lim_{j\rightarrow\infty} |f(i, j)|$ is bounded, i.e. there exists some $c \in \mathbb{R}_{\geq 0}$ such that $S_n \leq c$ for all $n \geq 0$.
- **C4.** For all $i, j_1, j_2 \in \mathbb{N}$, we have $j_1 \leq j_2$ implies $f(i, j_1) + |f(i, j_1)| \leq f(i, j_2) + |f(i, j_2)|$.

then it holds that

$$\sum_{i=0}^{\infty} \lim_{j\rightarrow\infty} f(i, j) \ = \ \lim_{j\rightarrow\infty} \sum_{i=0}^{\infty} f(i, j).$$

Proof: For any $i, j \in \mathbb{N}$, we have $f(i, j) + |f(i, j)| \leq 2|f(i, j)| \leq 2 \ \lim_{j\rightarrow\infty} |f(i, j)|$ by **C1** and **C2**. Therefore, for any $i \in \mathbb{N}$, the sequence $\{f(i, j) + |f(i, j)|\}_{j=0}^{\infty}$ has a limit. That is, we have the condition

C5. For any $i \in \mathbb{N}$, the limit $\lim_{j\to\infty}(f(i,j) + |f(i,j)|)$ exists.

Moreover, it holds that $\lim_{j\to\infty}(f(i,j) + |f(i,j)|) \;\leq\; 2\; \lim_{j\to\infty}|f(i,j)|$. It follows that

C6. For any $n \in \mathbb{N}$, the partial sum $\sum_{i=0}^{n}\lim_{j\to\infty}(f(i,j) + |f(i,j)|) \;\leq\; 2\; \sum_{i=0}^{n}\lim_{j\to\infty}|f(i,j)| \leq 2c$.

By Proposition A.1 and conditions **C1**, **C2** and **C3**, we infer that

$$\lim_{j\to\infty}\sum_{i=0}^{\infty}|f(i,j)| \;=\; \sum_{i=0}^{\infty}\lim_{j\to\infty}|f(i,j)|. \tag{27}$$

By Proposition A.1 and conditions **C4**, **C5** and **C6**, we infer that

$$\lim_{j\to\infty}\sum_{i=0}^{\infty}(f(i,j) + |f(i,j)|) \;=\; \sum_{i=0}^{\infty}\lim_{j\to\infty}(f(i,j) + |f(i,j)|). \tag{28}$$

Since $\sum_{i=0}^{\infty} f(i,j) \;=\; \sum_{i=0}^{\infty}(f(i,j) + |f(i,j)|) - \sum_{i=0}^{\infty}|f(i,j)|$, we then have

$$\begin{array}{rcl}
\lim_{j\to\infty}\sum_{i=0}^{\infty} f(i,j) & = & \lim_{j\to\infty}(\sum_{i=0}^{\infty}(f(i,j) + |f(i,j)|) - \sum_{i=0}^{\infty}|f(i,j)|) \\
& & \text{[existence of the two limits by (27) and (28)]} \\
& = & \lim_{j\to\infty}\sum_{i=0}^{\infty}(f(i,j) + |f(i,j)|) - \lim_{j\to\infty}\sum_{i=0}^{\infty}|f(i,j)| \\
& & \text{[by (27) and (28)]} \\
& = & \sum_{i=0}^{\infty}\lim_{j\to\infty}(f(i,j) + |f(i,j)|) - \sum_{i=0}^{\infty}\lim_{j\to\infty}|f(i,j)| \\
& = & \sum_{i=0}^{\infty}(\lim_{j\to\infty}(f(i,j) + |f(i,j)|) - \lim_{j\to\infty}|f(i,j)|) \\
& = & \sum_{i=0}^{\infty}\lim_{j\to\infty}(f(i,j) + |f(i,j)| - |f(i,j)|) \\
& = & \sum_{i=0}^{\infty}\lim_{j\to\infty} f(i,j)
\end{array}$$

□

A.2 Realising payoffs

We aim to establish that in a finitary pLTS the maximum payoff realisable by following all possible weak derivations can be attained by using some static derivative policy. In order to do that, we need to formalise some concepts such as discounted weak derivation, discounted payoff etc.

Definition A.3 [Discounted weak derivation] The *discounted weak derivation* $\Delta \Longrightarrow_\delta \Delta'$ for discount factor δ ($0 \leq \delta \leq 1$) is obtained from a weak derivation by discounting each τ transition by δ. That is, there is a collection of $\Delta_k^{\rightarrow}$ and $\Delta_k^{\times}$ satisfying

$$\begin{array}{rcl}
\Delta & = & \Delta_0^{\rightarrow} + \Delta_0^{\times} \\
\Delta_0^{\rightarrow} & \xrightarrow{\tau} & \Delta_1^{\rightarrow} + \Delta_1^{\times} \\
 & \vdots & \\
\Delta_k^{\rightarrow} & \xrightarrow{\tau} & \Delta_{k+1}^{\rightarrow} + \Delta_{k+1}^{\times} \\
 & \vdots &
\end{array}$$

such that $\Delta' = \sum_{k=0}^{\infty}\delta^k\Delta_k^{\times}$.

It is trivial that the relation $\Longrightarrow_1$ coincides with $\Longrightarrow$ given in Definition 3.13.

Below we fix a finite state space $S = \{s_1, ..., s_n\}$ with $n \geq 1$ and deal with vectors. For example, a subdistribution $\Delta \in \mathscr{D}(S)$ can be viewed as the n-dimensional vector $\langle\, \Delta(s_1), ..., \Delta(s_n)\,\rangle$. Similarly, a weight function **w** can be considered as the n-dimensional vector $\langle\, \mathbf{w}(s_1), ..., \mathbf{w}(s_n)\,\rangle$.

Definition A.4 [Discounted payoff] Given a discount δ and weight function **w**, the *discounted payoff function* $\mathbb{P}_{\max}^{\delta,\mathbf{w}} : S \to \mathbb{R}$ is defined by

$$\mathbb{P}_{\max}^{\delta,\mathbf{w}}(s) \;=\; \sup\{\mathbf{w} \centerdot \Delta' \mid \overline{s} \Longrightarrow_\delta \Delta'\}$$

and we will generalise it to be of type $\mathscr{D}(S) \to \mathbb{R}$ by letting $\mathbb{P}_{\max}^{\delta,\mathbf{w}}(\Delta) = \sum_{s\in\lceil\Delta\rceil}\Delta(s)\cdot\mathbb{P}_{\max}^{\delta,\mathbf{w}}(s)$.

Definition A.5 [Max-seeking policy] Given a pLTS, discount δ and weighted function **w**, we say a static derivative policy dpp is *max-seeking* with respect to δ and **w** if for all s the following requirements are met.

1. If $\mathsf{dpp}(s)\uparrow$, then $\mathbf{w}(s) \;\geq\; \delta \cdot \mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta_1)$ for all $s \xrightarrow{\tau} \Delta_1$.
2. If $\mathsf{dpp}(s) = \Delta$ then
   (a) $\delta \cdot \mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta) \;\geq\; \mathbf{w}(s)$ and
   (b) $\mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta) \;\geq\; \mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta_1)$ for all $s \xrightarrow{\tau} \Delta_1$.

Lemma A.6 Given a finitely branching pLTS, discount δ and weighted function $\mathbf{w}$, there always exists a max-seeking policy.

Proof: Given a pLTS, discount δ and weighted function $\mathbf{w}$, the discounted payoff $\mathbb{P}^{\delta,\mathbf{w}}_{\max}(s)$ can be calculated for each state s. Then we can define a derivative policy dpp in the following way. For any state s, if $\mathbf{w}(s) \;\geq\; \delta \cdot \mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta_1)$ for all $s \xrightarrow{\tau} \Delta_1$, then we set dpp undefined at s. Otherwise, we choose a transition $s \xrightarrow{\tau} \Delta$ among the finite number of outgoing transitions from s such that $\mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta) \;\geq\; \mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta_1)$ for all other transitions $s \xrightarrow{\tau} \Delta_1$, and we set $\mathsf{dpp}(s) = \Delta$. □

Given a pLTS, discount δ, weight function $\mathbf{w}$, and derivative policy dpp, we define the function $F^{\delta,\mathsf{dpp},\mathbf{w}} : (S \rightarrow \mathbb{R}) \rightarrow (S \rightarrow \mathbb{R})$ by

$$F^{\delta,\mathsf{dpp},\mathbf{w}} \;:=\; \lambda f.\lambda s. \begin{cases} \mathbf{w}(s) & \text{if } \mathsf{dpp}(s)\uparrow \\ \delta \cdot f(\Delta) & \text{if } \mathsf{dpp}(s) = \Delta \end{cases} \tag{29}$$

where $f(\Delta) = \sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot f(s)$.

Lemma A.7 Given a pLTS, discount $\delta < 1$, weight function $\mathbf{w}$, and derivative policy dpp, the function $F^{\delta,\mathsf{dpp},\mathbf{w}}$ has a unique fixed point.

Proof: We first show that the function $F^{\delta,\mathsf{dpp},\mathbf{w}}$ is a contraction mapping. Let f, g be any two functions of type $S \rightarrow \mathbb{R}$.

$$\begin{array}{cl} & |F^{\delta,\mathsf{dpp},\mathbf{w}}(f) - F^{\delta,\mathsf{dpp},\mathbf{w}}(g)| \\ = & \sup\{|F^{\delta,\mathsf{dpp},\mathbf{w}}(f)(s) - F^{\delta,\mathsf{dpp},\mathbf{w}}(g)(s)| \;\mid s \in S\} \\ = & \sup\{|F^{\delta,\mathsf{dpp},\mathbf{w}}(f)(s) - F^{\delta,\mathsf{dpp},\mathbf{w}}(g)(s)\,| \;\mid s \in S \text{ and } \mathsf{dpp}(s)\downarrow\} \\ = & \delta \cdot \sup\{|f(\Delta) - g(\Delta)| \;\mid s \in S \text{ and } \mathsf{dpp}(s) = \Delta \text{ for some } \Delta\} \\ \leq & \delta \cdot \sup\{|f(s') - g(s')| \;\mid s' \in S\} \\ = & \delta \cdot |f - g| \\ < & |f - g| \end{array}$$

By Banach unique fixed point theorem, the function $F^{\delta,\mathsf{dpp},\mathbf{w}}$ has a unique fixed point. □

Lemma A.8 Given a pLTS, discount δ, weight function $\mathbf{w}$, and max-seeking policy dpp, the function $\mathbb{P}^{\delta,\mathbf{w}}_{\max}$ is a fixed point of $F^{\delta,\mathsf{dpp},\mathbf{w}}$.

Proof: We need to show that $F^{\delta,\mathsf{dpp},\mathbf{w}}(\mathbb{P}^{\delta,\mathbf{w}}_{\max})(s) = \mathbb{P}^{\delta,\mathbf{w}}_{\max}(s)$ holds for any state s. We distinguish two cases.

1. If $\mathsf{dpp}(s)\uparrow$, then $F^{\delta,\mathsf{dpp},\mathbf{w}}(\mathbb{P}^{\delta,\mathbf{w}}_{\max})(s) = \mathbf{w}(s) = \mathbb{P}^{\delta,\mathbf{w}}_{\max}(s)$ as expected.
2. If $\mathsf{dpp}(s) = \Delta$, then the arguments are more involved. First note that If $\overline{s} \Longrightarrow_\delta \Delta''$, then by Definition A.3 there exist some $\Delta_0^{\rightarrow}, \Delta_0^{\times}, \Delta_1, \Delta''$ such that $\overline{s} = \Delta_0^{\rightarrow} + \Delta_0^{\times}$, $\Delta_0^{\rightarrow} \xrightarrow{\tau} \Delta_1$, $\Delta_1 \Longrightarrow_\delta \Delta''$ and $\Delta' = \Delta_0^{\times} + \delta \cdot \Delta''$. So we can do the following calculation.

$$\begin{array}{cl} & \mathbb{P}^{\delta,\mathbf{w}}_{\max}(s) \\ = & \sup\{\mathbf{w} \centerdot \Delta' \mid \overline{s} \Longrightarrow_\delta \Delta'\} \\ = & \sup\{\mathbf{w} \centerdot (\Delta_0^{\times} + \delta \cdot \Delta'') \mid \overline{s} = \Delta_0^{\rightarrow} + \Delta_0^{\times}, \Delta_0^{\rightarrow} \xrightarrow{\tau} \Delta_1, \text{ and } \Delta_1 \Longrightarrow_\delta \Delta'' \\ & \qquad \text{for some } \Delta_0^{\rightarrow}, \Delta_0^{\times}, \Delta_1, \Delta''\} \\ = & \sup\{\mathbf{w} \centerdot \Delta_0^{\times} + \delta \cdot \mathbf{w} \centerdot \Delta'' \mid \overline{s} = \Delta_0^{\rightarrow} + \Delta_0^{\times}, \Delta_0^{\rightarrow} \xrightarrow{\tau} \Delta_1, \text{ and } \Delta_1 \Longrightarrow_\delta \Delta'' \\ & \qquad \text{for some } \Delta_0^{\rightarrow}, \Delta_0^{\times}, \Delta_1, \Delta''\} \\ = & \sup\{\mathbf{w} \centerdot \Delta_0^{\times} + \delta \cdot \sup\{\mathbf{w} \centerdot \Delta'' \mid \Delta_1 \Longrightarrow_\delta \Delta''\} \mid \overline{s} = \Delta_0^{\rightarrow} + \Delta_0^{\times} \text{ and } \Delta_0^{\rightarrow} \xrightarrow{\tau} \Delta_1 \\ & \qquad \text{for some } \Delta_0^{\rightarrow}, \Delta_0^{\times}, \Delta_1\} \\ = & \sup\{\mathbf{w} \centerdot \Delta_0^{\times} + \delta \cdot \mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta_1) \mid \overline{s} = \Delta_0^{\rightarrow} + \Delta_0^{\times} \text{ and } \Delta_0^{\rightarrow} \xrightarrow{\tau} \Delta_1 \\ & \qquad \text{for some } \Delta_0^{\rightarrow}, \Delta_0^{\times}, \Delta_1\} \\ = & \sup\{(1-p) \cdot \mathbf{w}(s) + p\delta \cdot \mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta_1) \mid p \in [0,1] \text{ and } \overline{s} \xrightarrow{\tau} \Delta_1 \\ & \qquad \text{for some } \Delta_1\} \qquad [\overline{s} \text{ can be split into } p\overline{s} + (1-p)\overline{s} \text{ only}] \\ = & \sup\{(1-p) \cdot \mathbf{w}(s) + p\delta \cdot \mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta_1) \mid p \in [0,1] \text{ and } s \xrightarrow{\tau} \Delta_1 \\ & \qquad \text{for some } \Delta_1\} \\ = & \sup\{(1-p) \cdot \mathbf{w}(s) + p\delta \cdot \sup\{\mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta_1) \mid s \xrightarrow{\tau} \Delta_1\} \mid p \in [0,1]\} \\ = & \max(\mathbf{w}(s),\;\; \delta \cdot \sup\{\mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta_1) \mid s \xrightarrow{\tau} \Delta_1\}) \\ = & \delta \cdot \mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta) \qquad [\text{as } \mathsf{dpp} \text{ is max-seeking}] \\ = & F^{\delta,\mathsf{dpp},\mathbf{w}}(\mathbb{P}^{\delta,\mathbf{w}}_{\max})(s) \end{array}$$

$\square$

Definition A.9 [Discounted weak SDP-derivation] Let Δ be a subdistribution and dpp a static derivative policy. We define a collection of subdistributions Δ_k as follows.

$$\begin{array}{rcll} \Delta_0 & = & \Delta & \\ \Delta_{k+1} & = & \sum\{\Delta_k(s)\cdot\mathsf{dpp}(s) \mid s\in\lceil\Delta_k\rceil \text{ and } \mathsf{dpp}(s)\downarrow\} & \text{for all } k\geq 0. \end{array}$$

Then $\Delta_k^\times$ is obtained from Δ_k by letting

$$\Delta_k^\times(s) \;=\; \begin{cases} 0 & \text{if } \mathsf{dpp}(s)\downarrow \\ \Delta_k(s) & \text{otherwise} \end{cases}$$

for all $k\geq 0$. Then the *discounted weak SDP-derivation* $\Delta \Longrightarrow_{\delta,\mathsf{dpp}} \Delta'$ determines a unique subdistribution Δ' with $\Delta' = \sum_{k=0}^\infty \delta^k\Delta_k^\times$.

In other words, if $\Delta \Longrightarrow_{\delta,\mathsf{dpp}} \Delta'$ then Δ comes from the discounted weak derivation $\Delta \Longrightarrow_\delta \Delta'$ which is constructed by following the derivative policy dpp when choosing τ transitions from each state. In the special case when the discount factor $\delta = 1$, we see that $\Longrightarrow_{1,\mathsf{dpp}}$ becomes $\Longrightarrow_{\mathsf{dpp}}$ as defined in page 33.

Definition A.10 [Policy-following payoff] Given a discount δ, weight function $\mathbf{w}$, and derivative policy dpp, the *policy-following payoff function* $\mathbb{P}^{\delta,\mathsf{dpp},\mathbf{w}} : S \to \mathbb{R}$ is defined by

$$\mathbb{P}^{\delta,\mathsf{dpp},\mathbf{w}}(s) \;=\; \mathbf{w}\centerdot\Delta'$$

where Δ is determined by the discounted weak SDP-derivation $\overline{s} \Longrightarrow_{\delta,\mathsf{dpp}} \Delta'$.

Lemma A.11 For any discount δ, weight function $\mathbf{w}$, and derivative policy dpp, the function $\mathbb{P}^{\delta,\mathsf{dpp},\mathbf{w}}$ is a fixed point of $F^{\delta,\mathsf{dpp},\mathbf{w}}$.

Proof: We need to show that $F^{\delta,\mathsf{dpp},\mathbf{w}}(\mathbb{P}^{\delta,\mathsf{dpp},\mathbf{w}})(s) = \mathbb{P}^{\delta,\mathsf{dpp},\mathbf{w}}(s)$ holds for any state s. There are two cases.

1. If $\mathsf{dpp}(s)\uparrow$, then $\overline{s} \Longrightarrow_{\delta,\mathsf{dpp}} \Delta'$ implies $\Delta' = \overline{s}$. Thus, $\mathbb{P}^{\delta,\mathsf{dpp},\mathbf{w}}(s) = \mathbf{w}(s) = F^{\delta,\mathsf{dpp},\mathbf{w}}(\mathbb{P}^{\delta,\mathsf{dpp},\mathbf{w}})(s)$ as required.
2. Suppose $\mathsf{dpp}(s) = \Delta_1$. If $\overline{s} \Longrightarrow_{\delta,\mathsf{dpp}} \Delta'$ then $s \xrightarrow{\tau} \Delta_1$, $\Delta_1 \Longrightarrow_{\delta,\mathsf{dpp}} \Delta''$ and $\Delta' = \delta\Delta''$ for some subdistribution Δ''. Therefore,

$$\begin{array}{rl} & \mathbb{P}^{\delta,\mathsf{dpp},\mathbf{w}}(s) \\ = & \mathbf{w}\centerdot\Delta' \\ = & \mathbf{w}\centerdot\delta\Delta'' \\ = & \delta\cdot\mathbf{w}\centerdot\Delta'' \\ = & \delta\cdot\mathbb{P}^{\delta,\mathsf{dpp},\mathbf{w}}(\Delta_1) \\ = & F^{\delta,\mathsf{dpp},\mathbf{w}}(\mathbb{P}^{\delta,\mathsf{dpp},\mathbf{w}})(s) \end{array}$$

$\square$

Proposition A.12 Let $\delta\in[0,1)$ be a discount and $\mathbf{w}$ a weight function. If dpp is a max-seeking policy with respect to δ and $\mathbf{w}$, then $\mathbb{P}^{\delta,\mathbf{w}}_{\max} = \mathbb{P}^{\delta,\mathsf{dpp},\mathbf{w}}$.

Proof: By Lemma A.7, the function $F^{\delta,\mathsf{dpp},\mathbf{w}}$ has a unique fixed point. By Lemmas A.8 and A.11, both $\mathbb{P}^{\delta,\mathbf{w}}_{\max}$ and $\mathbb{P}^{\delta,\mathsf{dpp},\mathbf{w}}$ are fixed points of the same function $F^{\delta,\mathsf{dpp},\mathbf{w}}$, which means that $\mathbb{P}^{\delta,\mathbf{w}}_{\max}$ and $\mathbb{P}^{\delta,\mathsf{dpp},\mathbf{w}}$ coincide with each other. $\square$

Lemma A.13 Suppose $\overline{s} \Longrightarrow \Delta'$ with $\Delta' = \sum_{i=0}^\infty \Delta_i^\times$ for some properly related $\Delta_i^\times$. Let $\{\delta_j\}_{j=0}^\infty$ be a nondecreasing sequence of discount factors converging to 1. Then for any weight function $\mathbf{w}$ it holds that

$$\mathbf{w}\centerdot\Delta' \;=\; \lim_{j\to\infty}\sum_{i=0}^\infty \delta_j^i(\mathbf{w}\centerdot\Delta_i^\times).$$

Proof: Let $f:\mathbb{N}\times\mathbb{N}\to\mathbb{R}$ be the function defined by $f(i,j) = \delta_j^i(\mathbf{w}\centerdot\Delta_i^\times)$. We check that f satisfies the four conditions in Proposition A.2.

1. f satisfies condition **C1**. For all $i, j_1, j_2\in\mathbb{N}$, if $j_1\leq j_2$ then $\delta_{j_1}^i \leq \delta_{j_2}^i$. It follows that

$$|f(i,j_1)| \;=\; |\delta_{j_1}^i(\mathbf{w}\centerdot\Delta_i^\times)| \;\leq\; |\delta_{j_2}^i(\mathbf{w}\centerdot\Delta_i^\times)| \;=\; |f(i,j_2)|.$$

2. f satisfies condition **C2**. For any $i \in \mathbb{N}$, we have

$$\lim_{j\to\infty} |f(i,j)| \;=\; \lim_{j\to\infty} |\delta_j^i(\mathbf{w} \centerdot \Delta_i^\times)| \;=\; |\mathbf{w} \centerdot \Delta_i^\times|. \tag{30}$$

3. f satisfies condition **C3**. For any $n \in \mathbb{N}$, the partial sum $S_n = \sum_{i=0}^n \lim_{j\to\infty} |f(i,j)|$ is bounded because

$$\sum_{i=0}^n \lim_{j\to\infty} |f(i,j)| \;=\; \sum_{i=0}^n |\mathbf{w} \centerdot \Delta_i^\times| \;\leq\; \sum_{i=0}^\infty |\mathbf{w} \centerdot \Delta_i^\times| \;\leq\; \sum_{i=0}^\infty |\Delta_i^\times| \;=\; |\Delta'|$$

   where the first equality is justified by (30).

4. f satisfies condition **C4**. For any $i, j_1, j_2 \in \mathbb{N}$, if $j_1 \leq j_2$ then

$$\begin{array}{ll} & f(i,j_1) + |f(i,j_1)| \\ = & \delta_{j_1}^i(\mathbf{w} \centerdot \Delta_i^\times) + |\delta_{j_1}^i(\mathbf{w} \centerdot \Delta_i^\times)| \\ = & \delta_{j_1}^i(\mathbf{w} \centerdot \Delta_i^\times + |\mathbf{w} \centerdot \Delta_i^\times|) \\ \leq & \delta_{j_2}^i(\mathbf{w} \centerdot \Delta_i^\times + |\mathbf{w} \centerdot \Delta_i^\times|) \\ = & f(i,j_2) + |f(i,j_2)|. \end{array}$$

Therefore, we can use Proposition A.2 to do the following inference.

$$\begin{array}{ll} & \lim_{j\to\infty} \sum_{i=0}^\infty \delta_j^i(\mathbf{w} \centerdot \Delta_i^\times) \\ = & \sum_{i=0}^\infty \lim_{j\to\infty} \delta_j^i(\mathbf{w} \centerdot \Delta_i^\times) \\ = & \sum_{i=0}^\infty \mathbf{w} \centerdot \Delta_i^\times \\ = & \mathbf{w} \centerdot \sum_{i=0}^\infty \Delta_i^\times \\ = & \mathbf{w} \centerdot \Delta' \end{array}$$

□

Corollary A.14 Let $\{\delta_j\}_{j=0}^\infty$ be a nondecreasing sequence of discount factors converging to 1. For any derivative policy dpp and weight function **w**, it holds that $\mathbb{P}^{1,\mathsf{dpp},\mathbf{w}} = \lim_{j\to\infty} \mathbb{P}^{\delta_j,\mathsf{dpp},\mathbf{w}}$.

Proof: We need to show that $\mathbb{P}^{1,\mathsf{dpp},\mathbf{w}}(s) = \lim_{j\to\infty} \mathbb{P}^{\delta_j,\mathsf{dpp},\mathbf{w}}(s)$, for any state s. Note that for any discount δ_j, each state s enables a unique discounted weak SDP-derivation $\overline{s} \Longrightarrow_{\delta_j,\mathsf{dpp}} \Delta^j$ such that $\Delta^j = \sum_{i=0}^\infty \delta_j^i \Delta_i^\times$ for some properly related $\Delta_i^\times$. Let $\Delta' = \sum_{i=0}^\infty \Delta_i^\times$. We have $\overline{s} \Longrightarrow_{1,\mathsf{dpp}} \Delta'$. Then we can infer that

$$\begin{array}{lll} & \lim_{j\to\infty} \mathbb{P}^{\delta_j,\mathsf{dpp},\mathbf{w}}(s) & \\ = & \lim_{j\to\infty} \mathbf{w} \centerdot \Delta^j & \\ = & \lim_{j\to\infty} \mathbf{w} \centerdot \sum_{i=0}^\infty \delta_j^i \Delta_i^\times & \\ = & \lim_{j\to\infty} \sum_{i=0}^\infty \delta_j^i(\mathbf{w} \centerdot \Delta_i^\times) & \\ = & \mathbf{w} \centerdot \Delta' & \text{by Lemma A.13} \\ = & \mathbb{P}^{1,\mathsf{dpp},\mathbf{w}}(s) & \end{array}$$

□

Theorem A.15 In a finitary pLTS, for any weight function **w** there exists a derivative policy dpp such that $\mathbb{P}^{1,\mathbf{w}}_{\max} = \mathbb{P}^{1,\mathsf{dpp},\mathbf{w}}$.

Proof: Let **w** be a weight function. By Proposition A.12, for every discount factor $d < 1$ there exists a max-seeking derivative policy dpp with respect to δ and **w** such that

$$\mathbb{P}^{\delta,\mathbf{w}}_{\max} = \mathbb{P}^{\delta,\mathsf{dpp},\mathbf{w}}. \tag{31}$$

Since the pLTS is finitary, there are finitely many different static derivative policies. There must exist a derivative policy dpp such that (31) holds for infinitely many discount factors. In other words, for every nondecreasing sequence $\{\delta_n\}_{n=0}^\infty$ converging to 1, there exists a subsequence $\{\delta_{n_j}\}_{j=0}^\infty$ and a derivative policy $\mathsf{dpp}^\star$ such that

$$\mathbb{P}^{\delta_{n_j},\mathbf{w}}_{\max} = \mathbb{P}^{\delta_{n_j},\mathsf{dpp}^\star,\mathbf{w}} \qquad \text{for all } j \geq 0. \tag{32}$$

For any state s, we infer as follows.

$$
\begin{array}{ll}
 & \mathbb{P}^{1,\mathbf{w}}_{\max}(s) \\
= & \sup\{\mathbf{w} \centerdot \Delta' \mid \overline{s} \Longrightarrow \Delta'\} \\
= & \sup\{\lim_{j\to\infty} \sum_{i=0}^{\infty} \delta^i_{n_j} (\mathbf{w} \centerdot \Delta^{\times}_i) \mid \overline{s} \Longrightarrow \Delta' \text{ with } \Delta' = \sum_{i=0}^{\infty} \Delta^{\times}_i\} \qquad \text{by Lemma A.13} \\
= & \lim_{j\to\infty} \sup\{\sum_{i=0}^{\infty} \delta^i_{n_j} (\mathbf{w} \centerdot \Delta^{\times}_i) \mid \overline{s} \Longrightarrow \Delta' \text{ with } \Delta' = \sum_{i=0}^{\infty} \Delta^{\times}_i\} \\
= & \lim_{j\to\infty} \sup\{\mathbf{w} \centerdot \sum_{i=0}^{\infty} \delta^i_{n_j} \Delta^{\times}_i \mid \overline{s} \Longrightarrow \Delta' \text{ with } \Delta' = \sum_{i=0}^{\infty} \Delta^{\times}_i\} \\
= & \lim_{j\to\infty} \sup\{\mathbf{w} \centerdot \Delta'' \mid \overline{s} \Longrightarrow_{\delta_{n_j}} \Delta''\} \\
= & \lim_{j\to\infty} \mathbb{P}^{\delta_{n_j},\mathbf{w}}_{\max}(s) \\
= & \lim_{j\to\infty} \mathbb{P}^{\delta_{n_j},\mathsf{dpp}^\star,\mathbf{w}}(s) \qquad \text{by (32)} \\
= & \mathbb{P}^{1,\mathsf{dpp}^\star,\mathbf{w}}(s) \qquad \text{by Corollary A.14}
\end{array}
$$

□

B Comparison of extremal testing with resolution-based testing

In this section we compare extremal testing with resolution-based testing for finitary pLTSs. We first show that for resolution-based testing it is sufficient to use a singleton set of success actions, i.e. $|\Omega|= 1$. Then we show that the maximum and minimum testing outcomes in a finitary pLTS can be attained by static resolutions; this property plays a key role in establishing that extremal testing coincides with resolution-based testing. Since the must case was already treated in Section 5.2.1, here we concentrate on the may case.

B.1 Scalar versus Vector testing

In [5] it was shown that for finitary pLTSs, and resolution-based testing, it is sufficient to use scalar testing. We wish to apply this result in our setting, to obtain Theorem B.4 below; however to do so we need to demonstrate that the manner in which the value obtained from a resolution used in that paper coincides with our use of least fixed points.

Definition B.1 Let Δ be a subdistribution in a deterministic pLTS $\langle S, \Omega_\tau, \to\rangle$. The probability that Δ starts with a sequence of actions $\aleph \in \Sigma^*$, is given by $\mathrm{Pr_R}(\Delta, \aleph)$, where $\mathrm{Pr_R} : S \times \Sigma^* \to [0, 1]$ is defined inductively:

$$\mathrm{Pr_R}(s, \varepsilon) := 1 \text{ and } \mathrm{Pr_R}(s, \alpha\aleph) := \begin{cases} \mathrm{Pr_R}(\Delta, \aleph) & \text{if } s \xrightarrow{\alpha} \Delta \\ 0 & \text{otherwise} \end{cases}$$

and $\mathrm{Pr_R}(\Delta, \aleph) = \mathrm{Exp}_\Delta(\mathrm{Pr_R}(\_, \aleph))$. The notation ε denotes the empty sequence of actions and $\alpha\aleph$ the sequence starting with $\alpha \in \Sigma$ and continuing with $\aleph \in \Sigma^*$. The value $\mathrm{Pr_R}(s, \aleph)$ is the probability that s starts with a sequence $\aleph$.

Let $\Sigma^{*\alpha}$ be the set of finite sequences in Σ^* that contains α just once, namely at the end. Then the probability that Δ ever reaches an action α is given by $\sum_{\aleph \in \Sigma^{*\alpha}} \mathrm{Pr_R}(\Delta, \aleph)$.

Definition B.2 Let Δ be a subdistribution in a deterministic pLTS $\langle S, \Omega_\tau, \to\rangle$. We define its success tuple $\mathbb{W}(\Delta) \in [0, 1]^n$ be such that $(\mathbb{W}(\Delta))_i$ is the probability that Δ reaches the action ω_i.

Then if Δ is a subdistribution in a (not necessarily deterministic) pLTS $\langle S, \Omega_\tau, \to\rangle$ we define the set of its success tuples to be those resulting as above from all its resolutions:

$$\mathbb{W}(\Delta) \ := \ \{\mathbb{W}(\Theta) \mid \langle R, \Theta, \to_R\rangle \text{ is a resolution of } \Delta\}.$$

Proposition B.3 Let Δ be a subdistribution in a deterministic pLTS $\langle S, \Omega_\tau, \to\rangle$. It holds that $\mathbb{W}(\Delta) = \mathbb{V}(\Delta)$.

Proof: We need to show that $\forall i : (\mathbb{W}(\Delta))_i = (\mathbb{V}(\Delta))_i$, i.e. $\sum_{\aleph \in \Sigma^{*\omega_i}} \mathrm{Pr_R}(\Delta, \aleph) = (\mathbb{V}(\Delta))_i$, for which it suffices to show that

$$\sum_{\aleph \in \Sigma^{*\omega_i}} \mathrm{Pr_R}(s, \aleph) = (\mathbb{V}(s))_i \qquad \text{for all } s \in S. \tag{33}$$

Since $\mathbb{V} = \bigsqcup_{n\in\mathbb{N}} \mathbb{V}^n$, we have that $(\mathbb{V}(s))_i = \bigsqcup_{n\in\mathbb{N}} (\mathbb{V}^n)_i$. Letting $\aleph \in \Sigma^*$ be a sequence of actions, we write $|\aleph|$ for its length. The sequence of reals $\{\sum_{\aleph \in \Sigma^{*\omega_i}, |\aleph| \le n} \mathrm{Pr_R}(s, \aleph)\}_{n=0}^{\infty}$ is nondecreasing and bounded by 1, so it converges and we have $\sum_{\aleph \in \Sigma^{*\omega_i}} \mathrm{Pr_R}(s, \aleph) = \bigsqcup_{n\in\mathbb{N}} \sum_{\aleph \in \Sigma^{*\omega_i}, |\aleph| \le n} \mathrm{Pr_R}(s, \aleph)$. We now prove by induction on n that

$$\sum_{\aleph \in \Sigma^{*\omega_i}, |\aleph| \le n} \mathrm{Pr_R}(s, \aleph) = (\mathbb{V}^n(s))_i \qquad \text{for all } n \in \mathbb{N}. \tag{34}$$

which will yield (33) immediately.

- The base case is $n = 0$. Then $\forall i : \sum_{\aleph\in\Sigma^{*}\omega_i,|\aleph|\leq n} \mathrm{Pr}_{\mathrm{R}}(s,\aleph) = 0$ and $\mathbb{V}^0(s)(\omega_i) = 0$.
- Now supposing (34) holds for some n, we consider the case for $n+1$. If $s \not\longrightarrow$, then we have

$$\sum_{\aleph\in\Sigma^{*}\omega_i,|\aleph|\leq n+1} \mathrm{Pr}_{\mathrm{R}}(s,\aleph) = 0 = (\mathbb{V}^{n+1}(s))_i.$$

If $s \xrightarrow{\alpha} \Delta'$ for some action α and distribution Δ', then there are two possibilities:

  – $\alpha = \omega_i$. We then have $(\mathbb{V}^{n+1}(s))_i = 1$. Note that if $\aleph$ is a finite non-empty sequence without any occurrence of ω_i, then $\mathrm{Pr}_{\mathrm{R}}(s,\aleph\omega_i) = 0$. In other words, $\sum_{\aleph\in\Sigma^{*}\omega_i,|\aleph|\leq n+1} \mathrm{Pr}_{\mathrm{R}}(s,\aleph) = \mathrm{Pr}_{\mathrm{R}}(s,\langle\omega_i\rangle) = 1$.

  – $\alpha \neq \omega_i$. Then $(\mathbb{V}^{n+1}(s))_i = (\mathbb{V}^{n}(\Delta))_i$. On the other hand, $\mathrm{Pr}_{\mathrm{R}}(s,\alpha'\aleph) = 0$ if $\alpha \neq \alpha'$. Therefore,

$$\begin{array}{rcll}\sum_{\aleph\in\Sigma^{*}\omega_i,|\aleph|\leq n+1} \mathrm{Pr}_{\mathrm{R}}(s,\aleph) & = & \sum_{\alpha\aleph\in\Sigma^{*}\omega_i,|\alpha\aleph|\leq n+1} \mathrm{Pr}_{\mathrm{R}}(s,\alpha\aleph) & \\ & = & \sum_{\alpha\aleph\in\Sigma^{*}\omega_i,|\alpha\aleph|\leq n+1} \mathrm{Pr}_{\mathrm{R}}(\Delta',\aleph) & \\ & = & \sum_{\aleph\in\Sigma^{*}\omega_i,|\aleph|\leq n} \mathrm{Pr}_{\mathrm{R}}(\Delta',\aleph) & \\ & = & (\mathbb{V}^{n}(\Delta'))_i & \text{by induction} \\ & = & (\mathbb{V}^{n+1}(s))_i & \end{array}$$

□

As a corollary of Proposition B.3, we have $\mathscr{A}^{\mathrm{r}}(T,P) = \mathbb{W}([\![T \mid_{\mathsf{Act}} P]\!])$ for any process P and test T. Therefore, the testing preorders $\sqsubseteq^{\Omega}_{\mathrm{pmay}}, \sqsubseteq^{\Omega}_{\mathrm{pmust}}$ defined in Section 4.2 coincides with those in Definition 6 of [5]. Now Theorem 4 of [5] (to be accurate, the variant of that theorem for state-based testing) tells us that when testing finitary processes it suffices to use a single success action rather than using multiple success actions. That is,

Theorem B.4 For finitary processes:

- $P \sqsubseteq^{\Omega}_{\mathrm{pmay}} Q$ if and only if $P \sqsubseteq_{\mathrm{pmay}} Q$
- $P \sqsubseteq^{\Omega}_{\mathrm{pmust}} Q$ if and only if $P \sqsubseteq_{\mathrm{pmust}} Q$ □

In view of the above theorem we can assume that only a single success action ω is used in tests, and in this setting we compare extremal testing and resolution-based testing.

B.2 Extremal versus resolution-based testing

Consider the set of all functions from a finite set R to $[0,1]$, denoted by $[0,1]^R$, and the distance function d over $[0,1]^R$ defined by $d(f,g) = \max |f(r) - g(r)|_{r\in R}$. We can check that $([0,1]^R, d)$ constitutes a complete metric space. Let $\delta \in (0,1]$ be a discount factor. Given a deterministic pLTS $\langle R, \{\omega,\tau\}, \rightarrow\rangle$, the discounted version of the functional $\mathscr{R}$ in Section 4.2, $\mathscr{R}^{\delta} : [0,1]^R \rightarrow [0,1]^R$ is defined by

$$\mathscr{R}^{\delta}(f)(r) = \begin{cases} 1 & \text{if } r \xrightarrow{\omega} \\ 0 & \text{if } r \not\xrightarrow{\omega} \text{ and } r \not\xrightarrow{\tau} \\ \delta \cdot f(\Delta) & \text{if } r \not\xrightarrow{\omega} \text{ and } r \xrightarrow{\tau} \Delta \end{cases} \tag{35}$$

where $f(\Delta) = \mathrm{Exp}_{\Delta}(f)$. Below we show that $\mathscr{R}^{\delta}$ is a continuous function over the complete lattice $[0,1]^R$. So the least fixed point of $\mathscr{R}^{\delta}$, denoted by $\mathbb{V}^{\delta}$, has the characterisation $\mathbb{V}^{\delta} = \bigsqcup_{n\in\mathbb{N}} \mathbb{V}^{\delta,n}$, where $\mathbb{V}^{\delta,n}$ is the n-th iteration of $\mathscr{R}^{\delta}$ over $\bot$. Note that if there is no discount, i.e. $\delta = 1$, we see that $\mathscr{R}^{\delta}, \mathbb{V}^{\delta}$ coincides with $\mathscr{R}, \mathbb{V}$ respectively. Similarly, we can define $\mathbb{V}_{\mathrm{sup}}{}^{\delta}$. (There is no need of treating $\mathbb{V}_{\mathrm{inf}}{}^{\delta}$ because the counterpart of Theorem B.8 for $\mathbb{V}_{\mathrm{inf}}$ can be established without using discount.)

Lemma B.5
1. For any $\delta \in (0,1]$, the functionals $\mathscr{R}^{\delta}$ and $\mathscr{R}^{\delta}_{\mathrm{max}}$ are continuous;
2. If $\delta_1, \delta_2 \in (0,1]$ and $\delta_1 \leq \delta_2$, then we have $\mathscr{R}^{\delta_1} \leq \mathscr{R}^{\delta_2}$ and $\mathscr{R}^{\delta_1}_{\mathrm{max}} \leq \mathscr{R}^{\delta_2}_{\mathrm{max}}$;
3. Let $\{\delta_n\}_{n\geq 1}$ be a nondecreasing sequence of discount factors converging to 1. Then $\bigsqcup_{n\in\mathbb{N}} \mathscr{R}^{\delta_n} = \mathscr{R}$ and $\bigsqcup_{n\in\mathbb{N}} \mathscr{R}^{\delta_n}_{\mathrm{max}} = \mathscr{R}_{\mathrm{max}}$.

Proof: We only consider $\mathscr{R}$, the case for $\mathscr{R}_{\mathrm{max}}$ is similar.

1. Let $f_0 \leq f_1 \leq ...$ be a nondecreasing chain in $[0,1]^R$. We need to show that

$$\mathscr{R}^{\delta}(\bigsqcup_{n\geq 0} f_n) \;=\; \bigsqcup_{n\geq 0} \mathscr{R}^{\delta}(f_n) \tag{36}$$

   For any $r \in R$, we are in one of the following three cases:

(a) $r \xrightarrow{\omega}$. We have

$$\begin{array}{rcll} \mathscr{R}^\delta(\bigsqcup_{n\geq 0} f_n)(r) & = & 1 & \text{by (35)} \\ & = & \bigsqcup_{n\geq 0} 1 & \\ & = & \bigsqcup_{n\geq 0} \mathscr{R}^\delta(f_n)(r) & \\ & = & (\bigsqcup_{n\geq 0} \mathscr{R}^\delta(f_n))(r) & \end{array}$$

(b) $r \not\xrightarrow{\omega}$ and $r \not\xrightarrow{\tau}$. Similar to last case. We have

$$\mathscr{R}^\delta(\bigsqcup_{n\geq 0} f_n)(r) = 0 = (\bigsqcup_{n\geq 0} \mathscr{R}^\delta(f_n))(r).$$

(c) Otherwise, $r \xrightarrow{\tau} \Delta$ for some distribution $\Delta \in \mathscr{D}_1(R)$. Then we infer that

$$\begin{array}{rcll} \mathscr{R}^\delta(\bigsqcup_{n\geq 0} f_n)(r) & = & \delta \cdot (\bigsqcup_{n\geq 0} f_n)(\Delta) & \text{by (35)} \\ & = & \delta \cdot \sum_{r\in\lceil\Delta\rceil} \Delta(r) \cdot (\bigsqcup_{n\geq 0} f_n)(r) & \\ & = & \delta \cdot \sum_{r\in\lceil\Delta\rceil} \Delta(r) \cdot (\bigsqcup_{n\geq 0} f_n(r)) & \\ & = & \delta \cdot \sum_{r\in\lceil\Delta\rceil} \bigsqcup_{n\geq 0} \Delta(r) \cdot f_n(r) & \\ & = & \delta \cdot \sum_{r\in\lceil\Delta\rceil} \lim_{n\to\infty} \Delta(r) \cdot f_n(r) & \\ & = & \delta \cdot \lim_{n\to\infty} \sum_{r\in\lceil\Delta\rceil} \Delta(r) \cdot f_n(r) & \text{by Proposition A.1} \\ & = & \delta \cdot \bigsqcup_{n\geq 0} \sum_{r\in\lceil\Delta\rceil} \Delta(r) \cdot f_n(r) & \\ & = & \delta \cdot \bigsqcup_{n\geq 0} f_n(\Delta) & \\ & = & \bigsqcup_{n\geq 0} \delta \cdot f_n(\Delta) & \\ & = & \bigsqcup_{n\geq 0} \mathscr{R}^\delta(f_n)(r) & \\ & = & (\bigsqcup_{n\geq 0} \mathscr{R}^\delta(f_n))(r) & \end{array}$$

In the above reasoning, Proposition A.1 can be applied because we can define the function $f : R \times \mathbb{N} \to \mathbb{R}_{\geq 0}$ by letting $f(r, n) = \Delta(r) \cdot f_n(r)$ and check that f satisfies the three conditions in Proposition A.1. If R is finite, we can extend it to a countable set $R' \supseteq R$ and require $f(r', n) = 0$ for all $r' \in R' \backslash R$ and $n \in \mathbb{N}$.

(a) f satisfies condition **C1**. For any $r \in R'$ and $n_1, n_2 \in \mathbb{N}$, if $n_1 \leq n_2$ then $f_{n_1} \leq f_{n_2}$. It follows that

$$f(r, n_1) \ = \ \Delta(r) \cdot f_{n_1}(r) \ \leq \ \Delta(r) \cdot f_{n_2}(r) \ = \ f(r, n_2).$$

(b) f satisfies condition **C2**. For any $r \in R'$, the sequence $\{\Delta(r) \cdot f_n(r)\}_{n=0}^\infty$ is nondecreasing and bounded by $\Delta(r)$. It follows that the limit $\lim_{n\to\infty} f(r, n)$ exists.

(c) f satisfies condition **C3**. For any $R'' \subseteq R'$, the partial sum $\sum_{r\in R''} \lim_{n\to\infty} f(r, n)$ is bounded because

$$\sum_{r\in R''} \lim_{n\to\infty} f(r, n) \ = \ \sum_{r\in R''} \lim_{n\to\infty} \Delta(r) \cdot f_n(r) \ \leq \ \sum_{r\in R''} \Delta(r) \ \leq \ \sum_{r\in R} \Delta(R) \ = \ 1.$$

2. Straightforward by the definition of $\mathscr{R}$.

3. For any $f \in [0, 1]^R$ and $r \in R$ we show that

$$\mathscr{R}(f)(r) = (\bigsqcup_{n\in\mathbb{N}} \mathscr{R}^{\delta_n})(f)(r). \tag{37}$$

We focus on the non-trivial case that $r \xrightarrow{\tau} \Delta$ for some distribution $\Delta \in \mathscr{D}_1(R)$.

$$\begin{array}{rcl} (\bigsqcup_{n\in\mathbb{N}} \mathscr{R}^{\delta_n})(f)(r) & = & \bigsqcup_{n\in\mathbb{N}} \mathscr{R}^{\delta_n}(f)(r) \\ & = & \bigsqcup_{n\in\mathbb{N}} \delta_n \cdot f(\Delta) \\ & = & f(\Delta) \cdot (\bigsqcup_{n\in\mathbb{N}} \delta_n) \\ & = & f(\Delta) \cdot 1 \\ & = & \mathscr{R}(f)(r) \end{array}$$

□

Lemma B.6 Let $\{\delta_n\}_{n\geq 1}$ be a nondecreasing sequence of discount factors converging to 1.

- $\mathbb{V} = \bigsqcup_{n\in\mathbb{N}} \mathbb{V}^{\delta_n}$

- $\mathbb{V}_{\text{sup}} = \bigsqcup_{n\in\mathbb{N}} {\mathbb{V}_{\text{sup}}}^{\delta_n}$

Proof: We only consider $\mathbb{V}$; the case for $\mathbb{V}_{\text{sup}}$ is similar. We use the notation $\mathit{lfp}(f)$ for the least fixed point of the function f over a complete lattice. Recall that $\mathbb{V}$ and $\mathbb{V}^{\delta_n}$ are the least fixed points of $\mathscr{R}$ and $\mathscr{R}^{\delta_n}$ respectively, so we need to prove that

$$\mathit{lfp}(\mathscr{R}) = \bigsqcup_{n\in\mathbb{N}} \mathit{lfp}(\mathscr{R}^{\delta_n}) \tag{38}$$

We now show two inequations.

For any $n \in \mathbb{N}$, we have $\delta_n \leq 1$, so Lemma B.5 (2) yields $\mathscr{R}^{\delta_n} \leq \mathscr{R}$. It follows that $\mathit{lfp}(\mathscr{R}^{\delta_n}) \leq \mathit{lfp}(\mathscr{R})$, thus $\bigsqcup_{n\in\mathbb{N}} \mathit{lfp}(\mathscr{R}^{\delta_n}) \leq \mathit{lfp}(\mathscr{R})$.

For the other direction, $\mathit{lfp}(\mathscr{R}) \leq \bigsqcup_{n\in\mathbb{N}} \mathit{lfp}(\mathscr{R}^{\delta_n})$, it suffices to show that $\bigsqcup_{n\in\mathbb{N}} \mathit{lfp}(\mathscr{R}^{\delta_n})$ is a pre-fixed point of $\mathscr{R}$, i.e. $\mathscr{R}(\bigsqcup_{n\in\mathbb{N}} \mathit{lfp}(\mathscr{R}^{\delta_n})) \leq \bigsqcup_{n\in\mathbb{N}} \mathit{lfp}(\mathscr{R}^{\delta_n})$, which we derive as follows. Let $\{\delta_n\}_{n\geq 1}$ be a nondecreasing sequence of discount factors converging to 1.

$$\begin{array}{cll}
 & \mathscr{R}(\bigsqcup_{n\in\mathbb{N}} \mathit{lfp}(\mathscr{R}^{\delta_n})) & \\
= & (\bigsqcup_{m\in\mathbb{N}} \mathscr{R}^{\delta_m})(\bigsqcup_{n\in\mathbb{N}} \mathit{lfp}(\mathscr{R}^{\delta_n})) & \text{by Lemma B.5 (3)} \\
= & \bigsqcup_{m\in\mathbb{N}} \mathscr{R}^{\delta_m}(\bigsqcup_{n\in\mathbb{N}} \mathit{lfp}(\mathscr{R}^{\delta_n})) & \\
= & \bigsqcup_{m\in\mathbb{N}} \bigsqcup_{n\in\mathbb{N}} \mathscr{R}^{\delta_m}(\mathit{lfp}(\mathscr{R}^{\delta_n})) & \text{by Lemma B.5 (1)} \\
= & \bigsqcup_{m\in\mathbb{N}} \bigsqcup_{n\geq m} \mathscr{R}^{\delta_m}(\mathit{lfp}(\mathscr{R}^{\delta_n})) & \\
\leq & \bigsqcup_{m\in\mathbb{N}} \bigsqcup_{n\geq m} \mathscr{R}^{\delta_n}(\mathit{lfp}(\mathscr{R}^{\delta_n})) & \text{by Lemma B.5 (2)} \\
= & \bigsqcup_{n\in\mathbb{N}} \mathscr{R}^{\delta_n}(\mathit{lfp}(\mathscr{R}^{\delta_n})) & \\
= & \bigsqcup_{n\in\mathbb{N}} \mathit{lfp}(\mathscr{R}^{\delta_n}) &
\end{array}$$

This completes the proof of (38). □

We say a resolution of a pLTS is *static* if its associated resolving function is injective.

Lemma B.7 Suppose $\delta < 1$ and Δ is a subdistribution in a finitely branching pLTS $\langle S, \{\tau, \omega\}, \rightarrow\rangle$. There exists a static resolution $\langle R, \Theta, \rightarrow\rangle$ of Δ with resolving function f such that $\mathbb{V}^{\delta}(r) = {\mathbb{V}_{\text{sup}}}^{\delta}(f(r))$ for all $r \in R$.

Proof: Let $\langle R, \Theta, \rightarrow\rangle$ be a resolution with an injective resolving function f such that

$$\text{if } r \xrightarrow{\tau} \Theta' \text{ then } {\mathbb{V}_{\text{sup}}}^{\delta}(f(\Theta')) = \max\{{\mathbb{V}_{\text{sup}}}^{\delta}(\Delta') \mid f(r) \xrightarrow{\tau} \Delta'\}.$$

The pLTS under consideration is finitely branching, which ensures the existence of the such resolving function f.

Let $g : R \rightarrow [0, 1]$ be the function defined by $g(r) = {\mathbb{V}_{\text{sup}}}^{\delta}(f(r))$ for all $r \in T$. Below we show that g is a fixed point of $\mathscr{R}^{\delta}$. If $r \xrightarrow{\omega}$ then $f(r) \xrightarrow{\omega}$. Therefore, $\mathscr{R}^{\delta}(g)(r) = 1 = {\mathbb{V}_{\text{sup}}}^{\delta}(f(r)) = g(r)$. Now suppose $r \not\xrightarrow{\omega}$ and $r \xrightarrow{\tau} \Theta'$. By the definition of f, we have $f(r) \not\xrightarrow{\omega}$, $f(r) \xrightarrow{\tau} f(\Theta')$ with ${\mathbb{V}_{\text{sup}}}^{\delta}(f(\Theta')) = max\{{\mathbb{V}_{\text{sup}}}^{\delta}(\Delta') \mid f(r) \xrightarrow{\tau} \Delta'\}$. Therefore,

$$\begin{array}{rcl}
\mathscr{R}^{\delta}(g)(r) & = & \delta \cdot g(\Theta') \\
 & = & \delta \cdot \sum_{r'\in R} \Theta(r') \cdot g(r') \\
 & = & \delta \cdot \sum_{r'\in R} \Theta(r') \cdot {\mathbb{V}_{\text{sup}}}^{\delta}(f(r')) \\
 & = & \delta \cdot \sum_{s\in S} f(\Theta')(s) \cdot {\mathbb{V}_{\text{sup}}}^{\delta}(s) \\
 & = & \delta \cdot {\mathbb{V}_{\text{sup}}}^{\delta}(f(\Theta')) \\
 & = & \delta \cdot \max\{{\mathbb{V}_{\text{sup}}}^{\delta}(\Delta') | f(r) \xrightarrow{\tau} \Delta'\} \\
 & = & {\mathbb{V}_{\text{sup}}}^{\delta}(f(r)) \\
 & = & g(r)
\end{array}$$

Since $\delta < 1$, the functional $\mathscr{R}^{\delta}$ is a contraction mapping. It follows from the Banach fixed point theorem that $\mathscr{R}^{\delta}$ has a unique fixed point. So we derive that g coincides with $\mathbb{V}^{\delta}$, i.e. $\mathbb{V}^{\delta}(r) = g(r) = {\mathbb{V}_{\text{sup}}}^{\delta}(f(r))$ for all $r \in R$. □

Theorem B.8 Let Δ be a subdistribution in a finitary pLTS $\langle S, \{\tau, \omega\}, \rightarrow\rangle$. There exists a static resolution $\langle R, \Theta, \rightarrow\rangle$ of Δ such that $\text{Exp}_{\Theta}(\mathbb{V}) = \text{Exp}_{\Delta}(\mathbb{V}_{\text{sup}})$.

Proof: By Lemma B.7, for every discount factor $d \in (0, 1)$ there exists a static resolution which achieves the maximum probability of success. Since the pLTS is finitary, there are finitely many different static resolutions. There must exist a static resolution that achieves the maximum probability of success for infinitely many discount factors. In other words, for every nondecreasing sequence

$\{\delta_n\}_{n\geq 1}$ converging to 1, there exists a subsequence $\{\delta_{n_k}\}_{k\geq 1}$ and a static resolution $\langle R, \Theta, \longrightarrow\rangle$ with resolving function f such that $\mathbb{V}^{\delta_{n_k}}(r) = {\mathbb{V}_{\text{sup}}}^{\delta_{n_k}}(f(r))$ for all $r \in R$ and $k = 1, 2, \ldots$. By Lemma B.6, we have that, for every $r \in R$,

$$\begin{array}{rcl} \mathbb{V}(r) & = & \bigsqcup_{k\in\mathbb{N}} \mathbb{V}^{\delta_{n_k}}(r) \\ & = & \bigsqcup_{k\in\mathbb{N}} {\mathbb{V}_{\text{sup}}}^{\delta_{n_k}}(f(r)) \\ & = & \mathbb{V}_{\text{sup}}(f(r)) \end{array}$$

It follows that $\text{Exp}_{\Theta}(\mathbb{V}) = \text{Exp}_{\Delta}(\mathbb{V}_{\text{sup}})$. □