Testing finitary probabilistic processes

Yuxin Deng, Rob van Glabbeek, Matthew Hennessy, Carroll Morgan

February 13, 2009

Abstract

This paper provides 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 capture divergence in terms of partial distributions.

Contents

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 only if they can be told apart by some notion of testing.

In our earlier work [3, 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 [9], 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 [21, 3], or pLTS, a generalisation of labelled transition systems [17]. 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 [3, 2] naturally generalises the two preorders of standard testing theory [5] 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* probabilistic tests.

The most significant result of [2] was an alternative characterisation of these preorders as particular forms of coinductively defined *simulations* 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, having finitely many states and displaying finite branching.

The simulation relations 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 part of a distribution to make a τ-move, whereas the remaining part remains 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 $Q_1 = \mathsf{rec}\, x.\, (\tau.x \,{}_{\frac{1}{2}}\!\oplus a.\, \mathbf{0})$ illustrated in Figure 1. Process Q_1 is indistinguishable, using tests, from the simple process $a.\,\mathbf{0}$; we have $Q_1 \simeq_{\mathsf{pmay}} a.\,\mathbf{0}$ and $Q_1 \simeq_{\mathsf{pmust}} a.\,\mathbf{0}$. This is because the process Q_1 will eventually perform the action a with probability 1. However, the action $[\![a.\,\mathbf{0}]\!] \stackrel{a}{\longrightarrow} [\![\,\mathbf{0}\,]\!]$ can not be simulated by a corresponding move $[\![Q_1]\!] \stackrel{\hat{\tau}}{\Longrightarrow} \stackrel{a}{\longrightarrow}$. No matter which distribution Δ we obtain from executing a finite sequence of internal moves $[\![Q_1]\!] \stackrel{\hat{\tau}}{\Longrightarrow} \Delta$, part of it is unable to subsequently perform the action a.

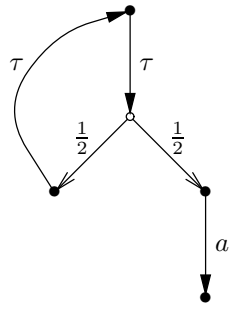


Figure 1: The pLTS of process Q_1

To address this problem we propose a new relation $\Delta \Longrightarrow \Theta$, that indicates 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 $[\![a.\,\mathbf{0}\,]\!]$ will turn out to be a weak derivative of $[\![Q_1]\!]$, $[\![Q_1]\!] \Longrightarrow [\![a.\,\mathbf{0}\,]\!]$, via the infinite sequence of internal moves

$$[\![Q_1]\!] \stackrel{\tau}{\longrightarrow} [\![Q_1 \,{}_{\frac{1}{2}}\!\oplus a.\,\mathbf{0}]\!] \stackrel{\tau}{\longrightarrow} [\![Q_1 \,{}_{\frac{1}{2^2}}\!\oplus a.\,\mathbf{0}\,]\!] \stackrel{\tau}{\longrightarrow} \ldots [\![Q_1 \,{}_{\frac{1}{2^n}}\!\oplus a.\,\mathbf{0}\,]\!] \stackrel{\tau}{\longrightarrow} \ldots\ .$$

One of our contributions here is the significant use of "subdistributions" that sum to *no more than* one [10, 16]. 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 $\varepsilon \stackrel{a}{\longrightarrow} \varepsilon$ for any action a, and a process like $\mathsf{rec}\, x.\, x$ that diverges via an infinite τ path gives rise to the weak transition $\mathsf{rec}\, x.\, x \Longrightarrow \varepsilon$. Our weak transitions relation $\Longrightarrow$ can be regarded as an extension of the one from Lynch, Segala & Vaandrager [15] to partial distributions; the latter, although defined in a very different way, can be obtained from 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 [17]. 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 [17]. 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.

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 [19] we can show that the potentially infinite set $\{\, \Delta \mid [\![P]\!] \Longrightarrow \Delta \,\}$ is nevertheless the convex closure of a *finite* set of subdistributions, from which Theorem 1.1 follows immediately. This in turn implies that the simulation preorder $\sqsubseteq_S$ is compact in the following sense:

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

This key result allows an *inductive* characterisation of the simulation preorder: we can construct a sequence of approximations $\sqsubseteq_S^k$ for $k \geq 0$ with the property that

Theorem 1.3 For finitary distributions Δ and Θ we have $\Delta \sqsubseteq_S \Theta$ if and only if $\Delta \sqsubseteq_S^k \Theta$ for every $k \geq 0$.

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 mimicked by a may-test.

Similar results accrue for must testing: details are given in Section 7.

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

Figure 2: Syntax of pCSP

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 is then given by the two-sorted syntax in Figure 2. It is essentially the finite language of [2, 3], to which has been added 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, as is that of the substitution of terms for free occurrences of variables (with renaming if necessary). By $Q[x \mapsto P]$ we mean the result of substituting the term P for the variable x in Q.

We write pCSP for the set of *closed terms* defined by this grammar, the pCSP *processes*, and sCSP for its subset of pCSP *states*: the sub-sort S above.

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 [18], 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.

The process $s \,\Box\, t$ 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 $P \,\Box\, Q$ 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 $s \,\Box\, (t_1 \,{}_p\!\oplus\, t_2)$ abbreviates the term $(s \,\Box\, t_1) \,{}_p\!\oplus\, (s \,\Box\, t_2)$.

The full language of CSP [1, 9, 18] 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 [18]. 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 $\bigsqcap_{i \in I} P_i$, for some finite index set I.

Our language is interpreted as a *probabilistic labelled transition system* [3, 2]. Essentially the same model has appeared in the literature under different names such as *NP-systems* [11], *probabilistic processes* [12], *simple probabilistic automata* [20], *probabilistic transition systems* [13] etc. Furthermore, there are strong structural similarities with *Markov Decision Processes* [19, 4].

[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.

(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 $\mathsf{Act}, \mathsf{Act}_\tau$ respectively.

Figure 3: Operational semantics of pCSP

We now fix some notation. A (discrete) probability *subdistribution* over a set S is a function $\Delta\colon S \to [0,1]$ 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 the *mass* of Δ, written $|\Delta|$, is $\sum_{s \in \lceil \Delta \rceil} \Delta(s)$. A subdistribution is a (total, or full) *distribution* if $|\Delta| = 1$. We write $\overline{s}$ to denote the point distribution assigning probability 1 to s and 0 to all other elements of S, so that $\lceil \overline{s} \rceil = \{s\}$. With $\mathcal{D}(S)$ we denote the set of subdistributions over S, and with $\mathcal{D}_1(S)$ its subset of full distributions. For $\Delta \in \mathcal{D}(S)$ and f a function with domain S, we write $\mathrm{Exp}_\Delta(f)$, the *expected value* of f over $\Delta \in \mathcal{D}(S)$, for $\sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot f(s)$ whenever the range of f makes the right-hand side well defined. For $\Delta, \Theta \in \mathcal{D}(S)$ we write $\Delta \leq \Theta$ iff $\Delta(s) \leq \Theta(s)$ for all $s \in S$.

When Δ_k for $k \in K$ is a collection, not necessarily finite, of subdistributions then $\sum_{k\in K} \Delta_k$ is the subdistribution given by $(\sum_{k\in K} \Delta_k)(s) := \sum_{k\in K} \Delta_k(s)$ — however because in general the sum could exceed 1 at some s we must view this as a partial operation. If the index set is finite, say $\{1, \cdots, n\}$, we often write the sum as $\Delta_1 + \cdots + \Delta_n$. When p is a real number from $[0,1]$ we use $p \cdot \Delta$ for the subdistribution $(p \cdot \Delta)(s) := p \cdot \Delta(s)$. Finally we use ε to denote the *empty* subdistribution of mass zero.

These operations on subdistributions do not readily adapt themselves to full distributions. But whenever the probabilities p_k sum to 1 and the Δ_k are themselves full distributions, then also $\sum_{k\in K} p_k \cdot \Delta_k$ is a full distribution.

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

(i) S is a set of states,
(ii) L is a set of transition labels,
(iii) relation $\rightarrow$ is a subset of $S \times L \times \mathcal{D}_1(S)$.

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 $(s, \alpha, \Delta) \in \rightarrow$, as well as $s \xrightarrow{\alpha}$ for $\exists \Delta : s \xrightarrow{\alpha} \Delta$ and $s \not\rightarrow$ for $\exists \alpha\colon s \not\xrightarrow{\alpha}$.

The operational semantics of pCSP is defined by a particular pLTS $\langle \mathsf{sCSP}, \mathsf{Act}_\tau, \rightarrow \rangle$, constructed by taking sCSP to be the set of states and $\mathsf{Act}_\tau := \mathsf{Act} \cup \{\tau\}$ to be the set of transition labels; we let a range over Act and α over Act_τ. We interpret pCSP processes P as distributions $[\![P]\!] \in \mathcal{D}_1(\mathsf{sCSP})$ via the function $[\![\_]\!] : \mathsf{pCSP} \to \mathcal{D}_1(\mathsf{sCSP})$ defined below:

$$\begin{array}{rcl} [\![s]\!] & := & \overline{s} \quad \text{for } s \in \mathsf{sCSP} \\ [\![P \,{}_p\!\oplus\, Q]\!] & := & p \cdot [\![P]\!] + (1-p) \cdot [\![Q]\!] \ . \end{array}$$

The transition relation $\rightarrow$ is defined in Figure 3. This is a very slight extension of the rules we used earlier [3, 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 [18], 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 $\Delta(s)$. 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 $(\alpha, \Delta) \in \mathsf{Act}_\tau \times \mathcal{D}_1(\mathsf{sCSP}))$ 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 [4].

Definition 2.2 A subdistribution $\Delta \in \mathcal{D}(S)$ 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 Lifted transitions, and weak moves over distributions

Our intention is to define simulation relations on processes, which are both sound and complete with respect to testing. This has been accomplished in [2] for recursion-free pCSP processes, where it was shown, for instance, that for such processes $P \sqsubseteq_{\text{pmay}} Q$ if and only if Q can (recursively) simulate the ability of P to perform actions. It turns out that to generalise these results requires a careful examination of weak derivations in probabilistic systems of unbounded depth; and that is the purpose of this section.

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 [3, 2], where we didn't 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*.

First some relatively standard terminology. For any subset X of $\mathcal{D}(S)$, with S a set, let $\updownarrow X$, the *convex closure* of X, be the least set satisfying:

(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. However, in general they do not:

Example 3.1 Let $S = \{s_i \mid i \in \mathbb{N}\}$. Then $\updownarrow\{\overline{s_i} \mid i > 2\}$ consists of all total distributions whose support is included in $\{s_i \mid i > 2\}$. However, with a definition of convex closure that requires only binary interpolations of distributions to be included, $\updownarrow\{\overline{s_i} \mid i > 2\}$ 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 $\mathcal{R} \subseteq Y \times \mathcal{D}(S)$ is convex whenever the set $\{\Delta \mid y\ \mathcal{R}\ \Delta\}$ is convex for every y in Y, and $\updownarrow\mathcal{R}$ denotes the smallest convex relation containing $\mathcal{R}$.

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

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

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

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

By construction $\overline{\mathcal{R}}$ is convex. Moreover, because $s(\updownarrow\mathcal{R})\Theta$ implies $\overline{s}\ \overline{\mathcal{R}}\ \Theta$ we have $\overline{\mathcal{R}}=\overline{\updownarrow\mathcal{R}}$, which means that when considering a lifted relation we can w.l.o.g. assume the original relation to have been convex. In fact when $\mathcal{R}$ is indeed convex, we have that $\overline{s}\ \overline{\mathcal{R}}\ \Theta$ and $s\ \mathcal{R}\ \Theta$ 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 simulate any other.

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

Lemma 3.4 $\Delta\ \overline{\mathcal{R}}\ \Theta$ 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 $s_i\ \mathcal{R}\ \Theta_i$,

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

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

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

Proof: This follows immediately from the characterisation in Lemma 3.4. □

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

Remark 3.6 From Lemma 3.4 it also follows that lifting enjoys the following two properties:
(Scaling) If $\Delta \;\overline{\mathcal{R}}\; \Theta$, $p \in \mathbb{R}$ and $|p \cdot \Delta| \leq 1$ then $p \cdot \Delta \;\overline{\mathcal{R}}\; p \cdot \Theta$.
(Additivity) If $\Delta_i \;\overline{\mathcal{R}}\; \Theta_i$ for $i \in I$ and $|\sum_{i \in I} \Delta_i| \leq 1$ then $\sum_{i \in I} \Delta_i \;\overline{\mathcal{R}}\; \sum_{i \in I} \Theta_i$.
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 3.7 Let $\mathcal{R} \subseteq S \times \mathcal{D}(S)$ be a relation from states to subdistributions. Then $f : S \to \mathcal{D}(S)$ is a *choice function for* $\mathcal{R}$, written $f \in_S \mathcal{R}$, if $s \;\mathcal{R}\; f(s)$ for every $s \in S$.

Proposition 3.8 Suppose $\mathcal{R} \subseteq S \times \mathcal{D}(S)$ is a convex relation. Then for any $\Delta \in \mathcal{D}(S)$, $\Delta \;\overline{\mathcal{R}}\; \Theta$ if and only if there is some choice function $f \in_{\lceil \Delta \rceil} \mathcal{R}$ such that $\Theta = \mathrm{Exp}_\Delta(f)$.

Proof: First suppose $\Theta = \mathrm{Exp}_\Delta(f)$ for some choice function $f \in_{\lceil \Delta \rceil} \mathcal{R}$, that is $\Theta = \sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot f(s)$. It now follows from Lemma 3.4 that $\Delta \;\overline{\mathcal{R}}\; \Theta$ since $s \;\mathcal{R}\; f(s)$ for each s.

Conversely suppose $\Delta \;\overline{\mathcal{R}}\; \Theta$; we have to find a choice function $f \in_{\lceil \Delta \rceil} \mathcal{R}$ such that $\Theta = \mathrm{Exp}_\Delta(f)$. Applying Lemma 3.4 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 $s_i \;\mathcal{R}\; \Theta_i$.

Now define the function $f : \lceil \Delta \rceil \to \mathcal{D}(S)$ by letting

$$f(s) = \sum_{\{\, i \in I \mid s_i = s \,\}} \left(\frac{p_i}{\Delta(s)}\right) \cdot \Theta_i$$

Note that $\Delta(s) = \sum_{\{\, i \in I \mid s_i = s \,\}} p_i$ and therefore by convexity $s \;\mathcal{R}\; f(s)$; so f is a choice function for $\mathcal{R}$. 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. □

An important further property is the following:

Proposition 3.9 If $\sum_{i \in I} p_i \cdot \Delta_i \;\overline{\mathcal{R}}\; \Theta$ then $\Theta = \sum_{i \in I} p_i \cdot \Theta_i$ for some subdistributions Θ_i such that $\Delta_i \;\overline{\mathcal{R}}\; \Theta_i$ for $i \in I$.

Proof: Let $\Delta \;\overline{\mathcal{R}}\; \Theta$ where $\Delta = \sum_{i \in I} p_i \cdot \Delta_i$. By Proposition 3.8, using that $\overline{\mathcal{R}} = \overline{\updownarrow \mathcal{R}}$, there is a choice function $f \in_{\lceil \Delta \rceil} \updownarrow \mathcal{R}$ 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 3.8 yields $\Delta_i \;\overline{\mathcal{R}}\; \Theta_i$ for $i \in I$. Finally,

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

The converse to the above is not true in general: from $\Delta \;\overline{\mathcal{R}}\; (\sum_{i \in I} p_i \cdot \Theta_i)$ 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 $\Delta_2 \xrightarrow{a} c$.

In fact a simplified form of Proposition 3.9 holds for un-lifted relations, provided they are convex:

Corollary 3.10 If $(\sum_{i \in I} p_i \cdot \overline{s_i}) \;\overline{\mathcal{R}}\; \Theta$ and $\mathcal{R}$ is convex, then $\Theta = \sum_{i \in I} p_i \cdot \Theta_i$ for subdistributions Θ_i with $s_i \;\mathcal{R}\; \Theta_i$ for $i \in I$.

Proof: Take Δ_i to be $\overline{s_i}$ in Proposition 3.9, whence $\Theta = \sum_{i \in I} p_i \cdot \Theta_i$ for some subdistributions Θ_i such that $\overline{s_i} \;\overline{\mathcal{R}}\; \Theta_i$ for $i \in I$. Because $\mathcal{R}$ is convex, we then have $s_i \;\mathcal{R}\; \Theta_i$ from the remarks following Definition 3.2. □

Lifting satisfies the following monadic property with respect to composition.

Lemma 3.11 Let $\mathcal{R}_1, \mathcal{R}_2 \subseteq S \times \mathcal{D}(S)$. Then the forward relational composition $\overline{\mathcal{R}_1};\overline{\mathcal{R}_2}$ is equal to the lifted composition $\overline{\mathcal{R}_1;\overline{\mathcal{R}_2}}$.

Proof: Suppose $\Delta \ \overline{\mathcal{R}_1};\overline{\mathcal{R}_2} \ \Phi$. Then there is some Θ such that $\Delta \ \overline{\mathcal{R}_1} \ \Theta \ \overline{\mathcal{R}_2} \ \Phi$. By Lemma 3.4 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 $s_i \ \mathcal{R}_1 \ \Theta_i$ for each $i \in I$. By Proposition 3.9 we obtain $\Phi = \sum_{i\in I} p_i \cdot \Phi_i$ with $\Theta_i \ \overline{\mathcal{R}_2} \ \Phi_i$. It follows that $s_i \ \mathcal{R}_1;\overline{\mathcal{R}_2} \ \Phi_i$, and thus $\Delta \ \overline{\mathcal{R}_1;\overline{\mathcal{R}_2}} \ \Phi$. So we have shown that $\overline{\mathcal{R}_1};\overline{\mathcal{R}_2} \ \subseteq \ \overline{\mathcal{R}_1;\overline{\mathcal{R}_2}}$. The other direction can be proved similarly. □

3.2 Weak transitions defined

Definition 3.12 (Weak τ moves to derivatives) Suppose we have subdistributions $\Delta, \ \Delta_k^{\rightarrow}, \Delta_k^{\times}$, for $k \geq 0$, with the following properties:

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

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

We call $\Delta' := \sum_{k=0}^{\infty} \Delta_k^{\times}$ a *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 3.5 it is easily checked that Definition 3.12 is well-defined in that derivatives do not sum to more than one.

Example 3.13 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

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

For $\Delta \xrightarrow{\tau}{}^{\star} \Theta$ 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

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

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 Definition 3.12 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 iff Δ can give rise to a divergent computation.

Example 3.14 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 3.15 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

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

which means that by definition we have

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

which is just $[\![a]\!]$ as claimed. □

Example 3.16 Consider 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 $1/k$ in total. Thus p_k must satisfy $1/k = p_k + (1{-}p_k)/(k{+}1)$, whence by arithmetic we have that $p_k := 1/k^2$ 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 3.17 Let P denote the process $\mathsf{rec}\, x.\, b \,{}_{\frac{1}{2}}\!\oplus\, (x \mid_\emptyset 0)$. Then we have the derivation:

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

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

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

where

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

a distribution with infinite support. □

3.3 Elementary properties of weak derivations

Here we develop some properties of the weak move relation $\Longrightarrow$ which will be important later on in the paper; full proofs and technical lemmas have in some cases been placed in Appendix B. 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 3.9.

Theorem 3.18 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: See Lemma B.3. □

With Theorem 3.18, the relation $\Longrightarrow \subseteq \mathcal{D}(S) \times \mathcal{D}(S)$ can be obtained as the lifting of a relation $\Longrightarrow_S$ from S to $\mathcal{D}(S)$, which is defined by writing $s \Longrightarrow_S \Theta$ just when $\overline{s} \Longrightarrow \Theta$.

Proposition 3.19 $\overline{(\Longrightarrow_S)} = (\Longrightarrow)$.

Proof: That $\Delta \overline{\Longrightarrow_S} \Theta$ implies $\Delta \Longrightarrow \Theta$ is a simple application of Part (i) of Theorem 3.18. 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 lemma 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 $\Delta \overline{\Longrightarrow_S} \Theta$. □

Corollary 3.20 The relations $\Longrightarrow$ is convex.

Proof: This is immediate from its being a lifting. □

We proceed with the important properties of reflexivity and transitivity of weak derivations.

Theorem 3.21 (Reflexivity and transitivity of $\Longrightarrow$) For any $\Delta \in \mathcal{D}(S)$ we have $\Delta \Longrightarrow \Delta$. Moreover, if $\Delta \Longrightarrow \Theta$ and $\Theta \Longrightarrow \Lambda$ then $\Delta \Longrightarrow \Lambda$.

Proof: The first statement is trivial: just take $\Delta_0^{\rightarrow} := \varepsilon$ in Definition 3.12. For the second, see Theorem B.4. □

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 3.22 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 $\Gamma_k, \Gamma_k^{\rightarrow}, \Gamma_k^{\times}, \Psi_k, \Psi_k^{\rightarrow}, \Psi_k^{\times} \in \mathcal{D}(S)$ such that $\Gamma_0 = \Psi_0 = \Phi$ and

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

Proposition 3.23 If $\Phi \Longrightarrow \Gamma$ is an initial segment of $\Phi \Longrightarrow \Psi$, then $\Gamma \Longrightarrow \Psi$.

Proof: To be placed in Appendix B. □

Further properties of weak derivations from finitary subdistributions are given in Section 5.

4 Testing probabilistic processes

We follow our earlier approach [2, 3] to the testing of probabilistic processes. A *test* is simply a process from the language pCSP except that it may use special "success" actions for reporting the outcome. Thus we assume a set Ω of fresh success actions not already in Act_τ. We refer to the augmented language as pCSP^Ω, and the pLTS it generates as plts_t. Now formally a test T is process from this language, and to apply it to process P we form the process $T \mid_{\mathsf{Act}} P$ in which *all* visible actions of P must synchronise with T. Thus this gives rise to a pLTS in which the only possible actions are τ and elements of Ω:

Definition 4.1 A pLTS of the form $\langle S, \Omega_\tau, \longrightarrow\rangle$ is referred to as a *computation structure*.

To determine the outcome of applying a test to a process we therefore have extract some result from a computation structure. However because of the presence of recursion in pCSP these may now be of infinite depth. Consequently we can no longer use our earlier approach [2], because it assumed both processes and tests to be finite.

In the following two subsections we outline two different ways in which outcomes can be calculated from the infinite (but finite-branching) computation structure of $T \mid_{\mathsf{Act}} P$; for finite-state systems they will turn out to be equivalent.

4.1 Extremal testing

Here we assume that tests are allowed to use a *single* success action ω; thus $\Omega = \{\omega\}$. Let S be the set of states in a computation structure. We view the unit interval $[0,1]$ ordered in the standard manner as a complete lattice (with least element 0), and this induces the same structure on the set of functions $S \to [0,1]$; the induced order is given by $f \leq g$ whenever $f(s) \leq g(s)$ for every $s \in S$. Now consider the functional $\mathcal{R}_{\text{min}} : (S{\to}[0,1]) \to (S{\to}[0,1])$ defined by:

$$\mathcal{R}_{\text{min}}(f)(s) = \begin{cases} 1 & \text{if } s \xrightarrow{\omega} \\ 0 & \text{if } s \not\xrightarrow{} \\ \min\{\, f(\Delta) \mid s \xrightarrow{\tau} \Delta \,\} & \text{otherwise} \end{cases}$$

where $f(\Delta) = \text{Exp}_\Delta(f)$. In a similar fashion we can define the functional $\mathcal{R}_{\text{max}} : (S{\to}[0,1]) \to (S{\to}[0,1])$ which uses the $\max$ function in place of $\min$. Both these functions are monotonic, and therefore have least fixed points, which we abbreviate to $\mathbb{V}_{\text{min}}$, $\mathbb{V}_{\text{max}}$ respectively. Furthermore, it can be shown that both $\mathcal{R}_{\text{min}}$ and $\mathcal{R}_{\text{max}}$ are continuous (see Lemma 4.19 below) , so we have the characterisation

$$\mathbb{V}_{\text{min}} = \bigsqcup_{n\in\mathbb{N}} {\mathbb{V}_{\text{min}}}^n \qquad \text{and} \qquad \mathbb{V}_{\text{max}} = \bigsqcup_{n\in\mathbb{N}} {\mathbb{V}_{\text{max}}}^n \tag{2}$$

where both ${\mathbb{V}_{\text{min}}}^0$ and ${\mathbb{V}_{\text{max}}}^0$ denote the *bottom* function $\bot$ defined by $\bot(s) = 0$ for all $s \in S$, and

- ${\mathbb{V}_{\text{min}}}^{(k+1)} = \mathcal{R}_{\text{min}}({\mathbb{V}_{\text{min}}}^k)$
- ${\mathbb{V}_{\text{max}}}^{(k+1)} = \mathcal{R}_{\text{max}}({\mathbb{V}_{\text{max}}}^k)$

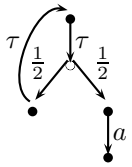
Now for a test T and a process P, we have two ways of defining the outcome of the application of T to P:

$$\mathcal{A}^{\text{e}}_{\text{min}}(T,P) = \mathbb{V}_{\text{min}}(T \mid_{\mathsf{Act}} P)$$
$$\mathcal{A}^{\text{e}}_{\text{max}}(T,P) = \mathbb{V}_{\text{max}}(T \mid_{\mathsf{Act}} P)$$

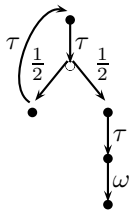
Here $\mathcal{A}^{\text{e}}_{\text{min}}(T,P)$ returns a single probability p, estimating the minimal probability of success; it is a pessimistic estimate. On the other hand $\mathcal{A}^{\text{e}}_{\text{max}}(T,P)$ is optimistic, in that it gives the maximal probability of success.

Definition 4.2

1. $P \sqsubseteq^{e}_{\text{pmay}} Q$ if for every test T, $\mathcal{A}^{\text{e}}_{\text{max}}(T,P) \leq \mathcal{A}^{\text{e}}_{\text{max}}(T,Q)$



(i) The process Q_1



(ii) The computation structure $(a.\omega \mid_{\mathsf{Act}} Q_1)$

Figure 4: Testing the process Q_1

2. $P \sqsubseteq^{e}_{\text{pmust}} Q$ if for every test T, $\mathcal{A}^{\text{e}}_{\text{min}}(T, P) \leq \mathcal{A}^{\text{e}}_{\text{min}}(T, Q)$

Example 4.3 Consider the process $Q_1 = \mathsf{rec}\, x.(\tau.x \oplus_{\frac{1}{2}} a)$, which is described graphically in Figure 4. When we apply the test $T = a.\omega$ to it we get the computation structure also described there. Note that this is deterministic and consequently the pessimistic and optimistic approaches coincide. That is, we have $\mathbb{V}_{\text{max}}(T \mid_{\mathsf{Act}} Q_1) = \mathbb{V}_{\text{min}}(T \mid_{\mathsf{Act}} Q_1) = v$ where v is the least probability –indeed the only probability, in this case– that satisfies

$$v \;=\; \frac{1}{2} \cdot v + \frac{1}{2} \quad \text{, so that } v = 1.$$

In general these solutions are unique just when the process is almost free of divergence, that is *almost (surely) convergent* so that the infinite internal paths have total probability zero. In fact it is possible to show that

$$Q_1 \simeq^{e}_{\text{pmay}} a \qquad\qquad Q_1 \simeq^{e}_{\text{pmust}} a.$$

□

4.2 Resolution-based testing

Here tests are allowed to use a *finite* collection of success actions, $\Omega = \{\omega_1, \ldots \omega_n\}$, although it will be convenient to assume that in any given state *at most* one of the actions ω_i can be executed. Then, when calculating the result of applying the test T to the process P, we use the collection of *resolutions* of $T \mid_{\mathsf{Act}} P$; intuitively a resolution represents a *run* of the combined process $T \mid_{\mathsf{Act}} P$, and as such gives exactly one probability for each success action. So in general the application of T to P will be a *set* of probabilities vectors, *result vectors*, not necessarily finite.

Here we adapt the notion of *resolution* defined in [21, 4] for probabilistic automata, to pLTSs. A computation structure $\langle S, \Omega_\tau, \rightarrow\rangle$ is called *deterministic* if for every $s \in S$ and every $\alpha \in \Omega_\tau$ there is at most one Δ such that $s \xrightarrow{\alpha} \Delta$.

Definition 4.4 A *resolution* of a computation structure $\langle S, \Omega_\tau, \rightarrow\rangle$ is a deterministic computation structure $\langle T, \Omega_\tau, \rightarrow\rangle$ such that there is a resolving function $f \in T \rightarrow S$ which satisfies the following conditions:

1. if $t \xrightarrow{\omega} \Theta$ for some $\omega \in \Omega$ then $f(t) \xrightarrow{\omega} f(\Theta)$
2. if $t \xrightarrow{\tau} \Theta$ then $f(t) \not\xrightarrow{\omega}$ for all $\omega \in \Omega$ and $f(t) \xrightarrow{\tau} f(\Theta)$
3. if $t \not\rightarrow$ then $f(t) \not\rightarrow$

where $f(\Theta)$ is the distribution defined by $f(\Theta)(s) = \sum_{f(t)=s} \Theta(t)$.

The reader is referred to Section 2 of [4] for a detailed discussion of this concept of resolutions, and the manner in which it represents *computation runs 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*. We often use the meta-variable R to

(i) The process Q_2 (ii) The computation structure $(a.\omega \mid_{\mathsf{Act}} Q_2)$

Figure 5: Testing the process Q_2

refer to a resolution, with resolving function f_R, and the computation structure involved will usually be understood from the context; in most cases it is that generated by plts_t.

Let S denote the set of states in a deterministic computation structure. Then by analogy with the functional $\mathcal{R}_{\text{min}}$ and $\mathcal{R}_{\text{max}}$ of the previous subsection we can define $\mathcal{R} : (S \to [0,1]^\Omega) \to (S \to [0,1]^\Omega)$ by

$$\mathcal{R}(f)(s) = \begin{cases} \vec{\omega_i} & \text{if } s \xrightarrow{\omega_i} \\ \vec{0} & \text{if } s \not\longrightarrow \\ f(\Delta) & \text{if } s \xrightarrow{\tau} \Delta \end{cases} \tag{3}$$

Here we use notation originally introduced in [4] for denoting result vectors in $[0,1]^\Omega$; $\vec{0}$ is the vector which is everywhere 0 while $\vec{\omega_i}$ has 1 in the ω_i^{th} position, but is otherwise 0. Once more this functional has a least fixed point, which we denote by $\mathbb{V}$. When the deterministic computation structure concerned is given by a resolution R, we say that R *realises* the function $\mathbb{V}$.

Now let $\mathcal{A}^\Omega(T,P)$ denote the set

$$\{\, \mathbb{V}(\Delta) \mid f_R(\Delta) = [\![T \mid_{\mathsf{Act}} P]\!], R \text{ a resolution of } \mathsf{plts}_\mathsf{t} \,\}. \tag{4}$$

Definition 4.5

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

These preorders are abbreviated to $P \sqsubseteq_{\text{pmay}} Q$, and $P \sqsubseteq_{\text{pmust}} Q$, when $|\Omega| = 1$.

Example 4.6 Consider the process $Q_2 = \mathsf{rec}\, x.\, \tau.(\tau.x \oplus_{\frac{1}{2}} a) \,\Box\, \tau.(\mathbf{0} \oplus_{\frac{1}{2}} a)$ and the application of the test $T = a.\omega$ to it; this is outlined in Figure 5. In the computation structure of $T \mid_{\mathsf{Act}} Q_2$, for each $k \geq 1$ there is a resolution R_k such that $\mathbb{V}(R_k) = (1 - \frac{1}{2^k})$; intuitively it goes around the loop $(k-1)$ times before at last taking the right hand τ action. Thus $\mathcal{A}^\Omega(T,Q_2)$ 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

$$\mathcal{A}^\Omega(T,Q_2) = \{(1-\frac{1}{2}),\ (1-\frac{1}{2^2}), \ldots, (1-\frac{1}{2^k}), \ldots\ 1\}$$

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

$$\mathcal{A}^\Omega(T,a) \leq_{\text{Ho}} \mathcal{A}^\Omega(T,Q_2) \qquad\qquad \mathcal{A}^\Omega(T,Q_2) \leq_{\text{Sm}} \mathcal{A}^\Omega(T,a)$$

Indeed it is possible to show that

$$a \sqsubseteq_{\text{pmay}} Q_2 \qquad\qquad Q_2 \sqsubseteq_{\text{pmust}} a$$

□

4.3 Comparison

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 4.7 [Monster B] Consider the infinite-state pLTS defined as follows: in addition to the states a and $\mathbf{0}$ it has the infinite set $b_1, b_2, \ldots$, with each of these having two transitions:

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

Now let us compare the state b_1 with the process a. With the test $a.\omega$, using resolutions, we get:

$$\begin{aligned} \mathcal{A}^{\Omega}(a.\omega, b_1) &= \{0, (1-\frac{1}{2}), \ldots, (1-\frac{1}{2^k}), \ldots\} \\ \mathcal{A}^{\Omega}(a.\omega, a) &= \{1\} \end{aligned} \tag{5}$$

which means that $a \not\sqsubseteq^{\Omega}_{\text{pmay}} b_1$.

However when we use extremal testing, the test $a.\omega$ can not distinguish these processes. It is straightforward to see that $\mathbb{V}_{\text{max}}(a.\omega \mid_{\mathsf{Act}} a) = 1$. Moreover for every $n > 0$ once can calculate $\mathbb{V}_{\text{max}}{}^{(n+1)}(a.\omega \mid_{\mathsf{Act}} b_k)$ to be $(1 - \frac{1}{2^{(k+n)}})$, which in turns means that $\mathbb{V}_{\text{max}}(a.\omega \mid_{\mathsf{Act}} b_1)$ also evaluates to 1.

With some more work one can go on to show that no test can distinguish between these processes using optimistic extremal testing, meaning that $a \sqsubseteq^{e}_{\text{pmay}} b_1$.

□

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. It is convenient to break the material up into three subsections. In this first we show, that for resolution-based testing it is sufficient to use a singleton set of success actions, $|\Omega| = 1$. In the second we examine *must* testing, which is easier than the *may* case, which in turn is treated in the final subsection.

4.3.1 Scalar versus Vector testing

In [4] it was shown that for finite-branching probabilistic automata, and resolution-based testing, it is sufficient to use results sets Ω of size 1. We wish to apply this result in our setting, to obtain Theorem 4.11 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.

First of all, we recall the notion of occurrences of actions and the results-gathering function $\mathbb{W}$ given in Definition 5 of [4].

Definition 4.8 Given a fully probabilistic automaton $R = \langle S, \Delta^{\circ}, \rightarrow \rangle$, the probability that R starts with a sequence of actions $\aleph \in \Sigma^*$, is given by $\Pr_{\mathrm{R}}(\aleph, \Delta^{\circ})$, where $\Pr_{\mathrm{R}} : \Sigma^* \times S \rightarrow [0,1]$ is defined inductively:

$$\Pr_{\mathrm{R}}(\varepsilon, s) := 1 \text{ and } \Pr_{\mathrm{R}}(\alpha\aleph, s) := \begin{cases} \Pr_{\mathrm{R}}(\aleph, \Delta) & \text{if } s \xrightarrow{\alpha} \Delta \\ 0 & \text{otherwise} \end{cases}$$

where $\Pr_{\mathrm{R}}(\aleph, \Delta) = \sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot \Pr_{\mathrm{R}}(\aleph, s)$. 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 $\Pr_{\mathrm{R}}(\aleph, s)$ is the probability that R proceeds with sequence $\aleph$ from state s.

Let $\Sigma^{*\alpha}$ be the set of finite sequences in Σ^* that contains α just once, namely at the end. Then the probability that that the fully probabilistic automaton R ever performs an action α is given by $\sum_{\aleph \in \Sigma^{*\alpha}} \Pr_{\mathrm{R}}(\aleph, \Delta^{\circ})$.

Definition 4.9 For a fully probabilistic automaton R, let its success tuple $\mathbb{W}.R \in [0,1]^n$ be such that $(\mathbb{W}.R)_i$ is the probability that R performs the action ω_i.

Then for a (not necessarily fully probabilistic) automaton M we define the set of its success tuples to be those resulting as above from all its resolutions:

$$\mathbb{W}.M \;:=\; \{\mathbb{W}.R \mid R \text{ is a resolution of } M\}.$$

Proposition 4.10 If $R = \langle S, \Delta^\circ, \rightarrow\rangle$ is a fully probabilistic automaton, then $\mathbb{W}.R = \mathbb{V}(\Delta^\circ)$.

Proof: We need to show that $\forall i : (\mathbb{W}.R)_i = (\mathbb{V}(\Delta^\circ))_i$, i.e. $\sum_{\aleph\in\Sigma^*\omega_i} \Pr_{\mathrm{R}}(\aleph, \Delta^\circ) = (\mathbb{V}(\Delta^\circ))_i$, for which it suffices to show that

$$\sum_{\aleph\in\Sigma^*\omega_i} \Pr_{\mathrm{R}}(\aleph, s) = (\mathbb{V}(s))_i \qquad \text{for all } s \in S. \tag{6}$$

Since $\mathbb{V} = \bigsqcup_{n\in\mathbb{N}} \mathbb{V}^n$, we have that $(\mathbb{V}(s))_i = \lim_{n\to\infty}(\mathbb{V}^n)_i$. Letting $\aleph \in \Sigma^*$ be a sequence of actions, we write $|\aleph|$ for its length. The sequence of reals $\{\sum_{\aleph\in\Sigma^*\omega_i,|\aleph|\le n} \Pr_{\mathrm{R}}(\aleph, s)\}_{n=0}^{\infty}$ is increasing and bounded by 1, so it converges and we have $\sum_{\aleph\in\Sigma^*\omega_i} \Pr_{\mathrm{R}}(\aleph, s) = \lim_{n\to\infty} \sum_{\aleph\in\Sigma^*\omega_i,|\aleph|\le n} \Pr_{\mathrm{R}}(\aleph, s)$. We now prove by induction on n that

$$\sum_{\aleph\in\Sigma^*\omega_i,|\aleph|\le n} \Pr_{\mathrm{R}}(\aleph, s) = (\mathbb{V}^n(s))_i \qquad \text{for all } n \in \mathbb{N}. \tag{7}$$

which will yield (6) immediately.

- The base case is $n = 0$. Then $\forall i : \sum_{\aleph\in\Sigma^*\omega_i,|\aleph|\le n} \Pr_{\mathrm{R}}(\aleph, s) = 0$ and $\mathbb{V}^0(s) = \vec{0}$.
- Now supposing (7) holds for some n, we consider the case for $n + 1$. If $s \not\longrightarrow$, then we have

  $$\sum_{\aleph\in\Sigma^*\omega_i,|\aleph|\le n+1} \Pr_{\mathrm{R}}(\aleph, s) = 0 = (\mathbb{V}^{n+1}(s))_i.$$

  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 $\Pr_{\mathrm{R}}(\aleph\omega_i, s) = 0$. In other words, $\sum_{\aleph\in\Sigma^*\omega_i,|\aleph|\le n+1} \Pr_{\mathrm{R}}(\aleph, s) = \Pr_{\mathrm{R}}(\langle\omega_i\rangle, s) = 1$.
  - $\alpha \neq \omega_i$. Then $(\mathbb{V}^{n+1}(s))_i = (\mathbb{V}^n(\Delta))_i$. On the other hand, $\Pr_{\mathrm{R}}(\alpha'\aleph, s) = 0$ if $\alpha \neq \alpha'$. Therefore,

    $$\begin{array}{rcl} \sum_{\aleph\in\Sigma^*\omega_i,|\aleph|\le n+1} \Pr_{\mathrm{R}}(\aleph, s) & = & \sum_{\alpha\aleph\in\Sigma^*\omega_i,|\alpha\aleph|\le n+1} \Pr_{\mathrm{R}}(\alpha\aleph, s) \\ & = & \sum_{\alpha\aleph\in\Sigma^*\omega_i,|\alpha\aleph|\le n+1} \Pr_{\mathrm{R}}(\aleph, \Delta) \\ & = & \sum_{\aleph\in\Sigma^*\omega_i,|\aleph|\le n} \Pr_{\mathrm{R}}(\aleph, \Delta) \\ & = & (\mathbb{V}^n(\Delta))_i \qquad \text{by induction} \\ & = & (\mathbb{V}^{n+1}(s))_i \end{array}$$

□

As a corollary of Proposition 4.10, we have $\mathcal{A}^\Omega(T, P) = \mathbb{W}.(T \mid_{\mathsf{Act}} P)$ for any process P and test T. Therefore, the testing preorders $\sqsubseteq^\Omega_{\text{pmay}}, \sqsubseteq^\Omega_{\text{pmust}}$ defined in Section 4.2 coincides with those in Definition 6 of [4]. Now Theorem 4 of [4] (to be accurate, the variant of that theorem for state-based testing) tells us that when testing finite-state processes it suffices to use a single success action rather than using multiple success actions. That is,

Theorem 4.11 For finite-state processes:

- $P \sqsubseteq^\Omega_{\text{pmay}} Q$ if and only if $P \sqsubseteq_{\text{pmay}} Q$
- $P \sqsubseteq^\Omega_{\text{pmust}} Q$ if and only if $P \sqsubseteq_{\text{pmust}} Q$ □

4.3.2 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. In view of Theorem 4.11, we will restrict our attention to resolution-based testing in which there is only one success action $\omega,\ |\Omega| = 1$.

Let us consider a computation structure $\langle S, \Omega_\tau, \rightarrow\rangle$, 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 $\mathbb{V}_{\text{min}}$, or by using resolutions of the computation structure 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.

Proposition 4.12 $\mathbb{V}_{min}(f_R(\Delta)) \leq \mathbb{V}(\Delta)$, *for any resolution* R.

Proof: Suppose R is a resolution of $\langle S, \Omega_\tau, \rightarrow \rangle$, giving the deterministic computation structure $\langle T, \Omega_\tau, \rightarrow \rangle$. First we show by induction on n that for every state $t \in T$

$$\mathbb{V}_{\min}{}^n(f_R(t)) \leq \mathbb{V}^n(t) \tag{8}$$

For $n = 0$, this is trivial. We consider the inductive step. First if $t \xrightarrow{\omega} \Theta$, then $f_R(t) \xrightarrow{\omega} f(\Theta)$, and thus $\mathbb{V}_{\min}{}^{n+1}(f_R(t)) = 1 = \mathbb{V}^{n+1}(t)$. A similar argument applies if $t \not\xrightarrow{}$, and so let us assume $t \xrightarrow{\tau} \Theta$ for some Θ, and $t \not\xrightarrow{\omega}$.

$$\begin{array}{rcl}
\mathbb{V}_{\min}{}^{(n+1)}(f_R(t)) & = & min\{\mathbb{V}_{\min}{}^n(\Delta) | f_R(t) \xrightarrow{\tau} \Delta\} \\
& \leq & \mathbb{V}_{\min}{}^n(f_R(\Theta)) \\
& = & \sum_{s \in S} f_R(\Theta)(s) \cdot \mathbb{V}_{\min}{}^n(s) \\
& = & \sum_{t' \in T} \Theta(t') \cdot \mathbb{V}_{\min}{}^n(f_R(t')) \\
& \geq & \sum_{t' \in T} \Theta(t') \cdot \mathbb{V}^n(t') \qquad \text{by induction} \\
& = & \mathbb{V}^n(\Theta) \\
& = & \mathbb{V}^{(n+1)}(t)
\end{array}$$

Now by continuity we have from (8) that

$$\mathbb{V}_{\min}(f_R(t)) \leq \mathbb{V}(t) \tag{9}$$

and this result is then easily generalised to distributions:

$$\begin{array}{rcl}
\mathbb{V}_{\min}(f_R(\Delta)) & = & \sum_{s \in S} f_R(\Delta)(s) \cdot \mathbb{V}_{\min}(s) \\
& = & \sum_{t \in T} \Delta(t) \cdot \mathbb{V}_{\min}(f_R(t)) \\
& \leq & \sum_{t \in T} \Delta(t) \cdot \mathbb{V}(t) \qquad \text{by (9)} \\
& = & \mathbb{V}(\Delta)
\end{array}$$

□

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

This new material of Matthew's should be used in the proof of Lemma 5.3.
A resolution R is said to be *static* if its resolving function f_R is injective. Again we refer the reader to [4] for a discussion of power of this restriction. Static restrictions are particularly simple, in that they does not allow states to be resolving into distributions, or computation steps to be interpolated. Moreover they are very easy to describe.

Definition 4.13 A (static) policy for a computation structure $\langle S, \Omega_\tau, \rightarrow \rangle$ is a partial function $\mathsf{pp} : S \rightharpoonup \mathcal{D}_1(S)$ satisfying:

- $s \longrightarrow$ implies $\mathsf{pp}(s)$ is defined
- $s \xrightarrow{\omega}$ implies $s \xrightarrow{\omega} \mathsf{pp}(s)$
- otherwise, if $s \xrightarrow{\tau}$ then $s \xrightarrow{\tau} \mathsf{pp}(s)$

Intuitively a policy pp decides between the choices available at a given state, with a preference for reporting success.

It is easy to see that a policy pp determines a static resolution of the computation structure. This is defined as the deterministic computation structure $\langle S, \Omega_\tau, \rightarrow_{\mathsf{pp}} \rangle$, where $\rightarrow_{\mathsf{pp}}$ is determined by $s \rightarrow_{\mathsf{pp}} f(s)$, and note that the associated resolving function is the identity. Indeed it is possible to show that every static resolution is determined in this manner by some policy.

Proposition 4.14 *In any finite-branching computation structure, there exists a static resolution* R *such that* $\mathbb{V}(f_R^{-1}(\Delta)) = \mathbb{V}_{min}(\Delta)$ *for any distribution* Δ.

Proof: Let Δ be a distribution over the computation structure $\langle S, \Omega_\tau, \rightarrow\rangle$. We exhibit the required resolution by defining a policy over S. We say policy $\mathsf{pp}(-)$ is *min-seeking* if its domain is $\{\, s \in S \mid s \longrightarrow \}$ and it satisfies:

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

Note that by design a min-seeking policy satisfies:

$$\text{if } s \stackrel{\omega}{\not\longrightarrow} \text{ but } s \stackrel{\tau}{\longrightarrow} \text{ then } \mathbb{V}_{\text{min}}(s) = \mathbb{V}_{\text{min}}(\mathsf{pp}(s)) \tag{10}$$

In a finite-branching computation structure it is straightforward to define a min-seeking policy:

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

(ii) Otherwise, if $s \stackrel{\tau}{\longrightarrow}$ let $\{\Delta_1, \ldots \Delta_n\}$ be the finite non-empty set $\{\, \Delta \mid s \stackrel{\tau}{\longrightarrow} \Delta \,\}$. Now let $\mathsf{pp}(s)$ be any Δ_k satisfying the property $\mathbb{V}_{\text{min}}(\Delta_k) \leq \mathbb{V}_{\text{min}}(\Delta_j)$ 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 satisfies the requirements of the proposition. For the sake of clarity let us write $\mathbb{V}_{\mathsf{pp}}(\Delta)$ for the value realised for Δ in the resolution determined by the policy $\mathsf{pp}(-)$.

We already know, from Proposition 4.12, that $\mathbb{V}_{\text{min}}(\Delta) \leq \mathbb{V}_{\mathsf{pp}}(\Delta)$ and so we concentrate on the converse, $\mathbb{V}_{\mathsf{pp}}(\Delta) \leq \mathbb{V}_{\text{min}}(\Delta)$. Recall that the function $\mathbb{V}_{\mathsf{pp}}$ is the least fixpoint of the functional $\mathcal{R}$ defined in (3) above, and interpreted in the resolution determined by $\mathsf{pp}(-)$. So the result follows if we can show that the function $\mathbb{V}_{\text{min}}$ is also a fixed point. Since $|\Omega| = 1$ this amounts to proving

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

However this is a straightforward consequence of (10) above. □

Theorem 4.15 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 $\mathcal{A}^{\Omega}(T, P)$, for any arbitrary test T, there exists some $v' \in \mathcal{A}^{\Omega}(T, Q)$ such that $v' \leq v$. The value v must be of the form $\mathbb{V}(\Delta)$, where there is a resolution R such that $f_R(\Delta) = [\![T \mid_{\mathsf{Act}} P]\!]$. From Proposition 4.12 we know that $\mathbb{V}_{\text{min}}([\![T \mid_{\mathsf{Act}} P]\!]) \leq v$, and now from the hypothesis $P \sqsubseteq^{e}_{\text{pmust}} Q$ we have that $\mathbb{V}_{\text{min}}([\![Q \mid_{\mathsf{Act}} P]\!]) \leq v$. Now employing Proposition 4.14 we can find some other (static) resolution R' and such that $\mathbb{V}(\Theta) = \mathbb{V}_{\text{min}}([\![Q \mid_{\mathsf{Act}} P]\!])$, where Θ is $f_{R'}([\![Q \mid_{\mathsf{Act}} P]\!])$. So we can take the required v' to be $\mathbb{V}(\Theta)$.

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

□

4.3.3 May testing

Here we can try to apply the same proof strategy as in the previous section. The analogue to Proposition 4.12 goes through:

Proposition 4.16 $\mathbb{V}(\Delta) \leq \mathbb{V}_{max}(f_R(\Delta))$ *for any resolution* R*.*

Proof: Similar to the proof of Proposition 4.12 □

However the proof strategy used in Proposition 4.14 can not be used to show that $\mathbb{V}_{\text{max}}$ can be realised by some static resolution, as the following example shows.

Example 4.17 In analogy with the definition used in the proof of Proposition 4.14, we say that a policy $\mathsf{pp}(-)$ is *max-seeking* if its domain is precisely $\{\, s \in S \mid s \longrightarrow \}$, and

$$\text{if } s \stackrel{\omega}{\not\longrightarrow} \text{ but } s \stackrel{\tau}{\longrightarrow} \text{ then } \mathbb{V}_{\max}(\Delta) \leq \mathbb{V}_{\max}(\mathsf{pp}(s)) \;\; \text{whenever } s \stackrel{\tau}{\longrightarrow} \Delta$$

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

To see this, let us turn the (finite-branching) pLTS used in Example 4.7 into a computation structure. Here in addition to the two states ω and $\mathbf{0}$ there is the infinite set $c_1, \ldots c_k, \ldots$ and the transitions

- $c_k \stackrel{\tau}{\longrightarrow} \overline{c_{k+1}}$
- $c_k \stackrel{\tau}{\longrightarrow} [\![\, \mathbf{0} \,{}_{pk}\!\oplus \omega]\!]$, where again pk is the probability $\frac{1}{2^k}$.

One can calculate $\mathbb{V}_{\max}(c_k)$ to be 1 for every k, and a max-seeking policy is given determined by $\mathsf{pp}(c_k) = c_{k+1}$; indeed this is essentially the only such policy. However this resolution does not realise $\mathbb{V}_{\max}$, as $\mathbb{V}_{\mathsf{pp}}(c_k) = 0$. □

Nevertheless we will show that if we restrict attention to finite-branching, and finite-state, computation structures, then there will always exist some static resolution which realises $\mathbb{V}_{\max}$. The proof relies on techniques used in Markow process theory [19], and unlike that of Proposition 4.14 is non-constructive; we simply prove that some such resolution exists, without actually showing how to construct it.

Consider the set of all functions from a finite set S to $[0,1]$, denoted by $\mathcal{F}_S$, and the distance function d over $\mathcal{F}_S$ defined by $d(f,g) = max|f(s) - g(s)|_{s\in S}$. We can check that $(\mathcal{F}_S, d)$ constitutes a complete metric space. Let $\delta \in (0,1)$ be a discount factor. The discounted version of the functional $\mathcal{R}$ in Section 4.2, $\mathcal{R}^\delta : \mathcal{F}_S \to \mathcal{F}_S$ defined by

$$\mathcal{R}^\delta(f)(s) = \begin{cases} 1 & \text{if } s \stackrel{\omega}{\longrightarrow} \\ 0 & \text{if } s \not\longrightarrow \\ \delta \cdot f(\Delta) & \text{otherwise} \end{cases} \tag{11}$$

where $f(\Delta) = \mathrm{Exp}_\Delta(f)$, is a contraction mapping with constant δ. It follows from the Banach fixed point theorem that $\mathcal{R}^\delta$ has a unique fixed point when $\delta < 1$, which we denote by $\mathbb{V}^\delta$. On the other hand, it can be shown that $\mathcal{R}^\delta$ is a continuous function over the complete lattice $\mathcal{F}_S$. So $\mathbb{V}^\delta$, as the least fixed point of $\mathcal{R}^\delta$, 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 $\mathcal{R}^\delta$ over $\bot$. Note that if there is no discount, i.e. $\delta = 1$, we see that $\mathcal{R}^\delta, \mathbb{V}^\delta$ coincides with $\mathcal{R}, \mathbb{V}$ respectively. Similarly, we can define ${\mathbb{V}_{\min}}^\delta$ and ${\mathbb{V}_{\max}}^\delta$.

Let $\mathbb{R}^+$ be the set of non-negative real numbers. The following property is very useful. It says that the directed sup and countable sum can be interchanged. Later on, we only need the first clause since we restrict ourselves to finite-state systems in this paper. The second clause is also proved because it is interesting by itself.

Lemma 4.18 Let S be a set and $(S \to \mathbb{R}^+)$ be the set of all functions from S to $\mathbb{R}^+$. Suppose $\{f_i \mid i \in I\}$ is any directed subset of $(S \to \mathbb{R}^+)$.

1. If S is finite, then
$$\sum_{s\in S} \bigsqcup_{i\in I} f_i(s) \;=\; \bigsqcup_{i\in I} \sum_{s\in S} f_i(s).$$
2. If S is countable and the partial sum $S_n := \sum_{j=1}^n \bigsqcup_{i\in I} f_i(s_j)$ is bounded, i.e. there exists some $c \in \mathbb{R}^+$ such that $S_n \leq c$ for any n, then
$$\sum_{s\in S} \bigsqcup_{i\in I} f_i(s) \;=\; \bigsqcup_{i\in I} \sum_{s\in S} f_i(s).$$

Proof: 1. Since S is finite, we can assume that $|S| = N$ for some $N \in \mathbb{N}$. Let ϵ be any positive real number. For each $s \in S$, there is some index i_s such that $0 \leq \bigsqcup_{i\in I} f_i(s) - f_i(s) \leq \frac{\epsilon}{N}$ for all $i > i_s$. Let $i_S = max\{i_s \mid s \in S\}$. For any $s \in S$, we have $0 \leq \bigsqcup_{i\in I} f_i(s) - f_i(s) \leq \frac{\epsilon}{N}$ for all $i > i_S$. Summing up over all $s \in S$, we get $0 \leq \sum_{s\in S} \bigsqcup_{i\in I} f_i(s) - \sum_{s\in S} f_i(s) \leq \epsilon$ for all $i > i_S$. Therefore, $\bigsqcup_{i\in I} \sum_{s\in S} f_i(s) = \lim_{i\to\infty} \sum_{s\in S} f_i(s) = \sum_{s\in S} \bigsqcup_{i\in I} f_i(s)$.

2. Since the sequence $\{S_n\}_{n\in\mathbb{N}}$ is increasing and bounded, it converges to $\sum_{s\in S}\bigsqcup_{i\in I} f_i(s)$. Let ϵ be any positive real number. We can take a finite subset S' of S which is large enough so that

$$0 \leq \sum_{s\in S}\bigsqcup_{i\in I} f_i(s) - \sum_{s\in S'}\bigsqcup_{i\in I} f_i(s) \leq \frac{\epsilon}{2}. \tag{12}$$

With the same argument as in the proof the first clause, we can choose an index $i_{S'}$ so that

$$0 \leq \sum_{s\in S'}\bigsqcup_{i\in I} f_i(s) - \sum_{s\in S'} f_i(s) \leq \frac{\epsilon}{2} \tag{13}$$

for all $i > i_{S'}$. We observe that $f_i(s) \leq \bigsqcup_{i\in I} f_i(s)$, so the sequence $\{\sum_{j=1}^{n} f_i(s)\}_{n\in\mathbb{N}}$, for any $i \in I$, is increasing and bounded, thus converges to $\sum_{s\in S} f_i(s)$. Therefore, there exists some $N \in \mathbb{N}$ such that

$$0 \leq \sum_{s\in S} f_i(s) - \sum_{j=1}^{N} f_i(s) \leq \frac{\epsilon}{2} \tag{14}$$

for all $i \in I$. Without loss of generality, we assume that $\{s_1, ..., s_N\} \subseteq S'$. It follows from (14) that

$$-\frac{\epsilon}{2} \leq \sum_{s\in S'} f_i(s) - \sum_{s\in S} f_i(s) \leq 0 \tag{15}$$

for all $i \in I$. Take the sum of the three inequalities (12), (13) and (15), we obtain

$$-\frac{\epsilon}{2} \leq \sum_{s\in S}\bigsqcup_{i\in I} f_i(s) - \sum_{s\in S} f_i(s) \leq \epsilon \tag{16}$$

for all $i > i_{S'}$. Therefore, $\bigsqcup_{i\in I}\sum_{s\in S} f_i(s) = \lim_{i\to\infty}\sum_{s\in S} f_i(s) = \sum_{s\in S}\bigsqcup_{i\in I} f_i(s)$.

□

The functionals $\mathcal{R}^\delta$ and $\mathcal{R}^\delta_{\max}$ have the following properties.

Lemma 4.19 1. For any $\delta \in (0, 1]$, the functionals $\mathcal{R}^\delta$ and $\mathcal{R}^\delta_{\max}$ are continuous;

2. If $\delta_1, \delta_2 \in (0, 1]$ and $\delta_1 \leq \delta_2$, then we have $\mathcal{R}^{\delta_1} \leq \mathcal{R}^{\delta_2}$ and $\mathcal{R}^{\delta_1}_{\max} \leq \mathcal{R}^{\delta_2}_{\max}$;

3. Let $\{\delta_n\}_{n\geq 1}$ be a nondecreasing sequence of discount factors converging to 1. It holds that $\bigsqcup_{n\in\mathbb{N}}\mathcal{R}^{\delta_n} = \mathcal{R}$ and $\bigsqcup_{n\in\mathbb{N}}\mathcal{R}^{\delta_n}_{\max} = \mathcal{R}_{\max}$.

Proof: We only consider $\mathcal{R}$, the case for $\mathcal{R}_{\max}$ is similar.

1. Let $f_0 \leq f_1 \leq ...$ be an increasing chain in $S \to [0, 1]$. We need to show that

$$\mathcal{R}^\delta(\bigsqcup_{n\geq 0} f_n) \;=\; \bigsqcup_{n\geq 0}\mathcal{R}^\delta(f_n) \tag{17}$$

For any $s \in S$, we are in one of the following three cases:

(a) $s \xrightarrow{\omega_i}$. We have

$$\begin{array}{lcll} \mathcal{R}^\delta(\bigsqcup_{n\geq 0} f_n)(s) & = & 1 & \text{by (11)} \\ & = & \bigsqcup_{n\geq 0} 1 & \\ & = & \bigsqcup_{n\geq 0}\mathcal{R}^\delta(f_n)(s) & \\ & = & (\bigsqcup_{n\geq 0}\mathcal{R}^\delta(f_n))(s) & \end{array}$$

(b) $s \not\longrightarrow$. Similar to last case. We have

$$\mathcal{R}^{\delta}(\bigsqcup_{n\geq 0} f_n)(s) = 0 = (\bigsqcup_{n\geq 0} \mathcal{R}^{\delta}(f_n))(s).$$

(c) Otherwise, $s \xrightarrow{\alpha} \Delta$ for some action α and distribution $\Delta \in \mathcal{D}_1(S)$. Then we infer that

$$\begin{array}{lcll}
\mathcal{R}^{\delta}(\bigsqcup_{n\geq 0} f_n)(s) & = & \delta \cdot (\bigsqcup_{n\geq 0} f_n)(\Delta) & \text{by (11)} \\
& = & \delta \cdot \sum_{s\in\lceil\Delta\rceil} \Delta(s) \cdot (\bigsqcup_{n\geq 0} f_n)(s) & \\
& = & \delta \cdot \sum_{s\in\lceil\Delta\rceil} \Delta(s) \cdot (\bigsqcup_{n\geq 0} f_n(s)) & \\
& = & \delta \cdot \sum_{s\in\lceil\Delta\rceil} \bigsqcup_{n\geq 0} \Delta(s) \cdot f_n(s) & \\
& = & \delta \cdot \bigsqcup_{n\geq 0} \sum_{s\in\lceil\Delta\rceil} \Delta(s) \cdot f_n(s) & \text{by Lemma 4.18} \\
& = & \delta \cdot \bigsqcup_{n\geq 0} f_n(\Delta) & \\
& = & \bigsqcup_{n\geq 0} \delta \cdot f_n(\Delta) & \\
& = & \bigsqcup_{n\geq 0} \mathcal{R}^{\delta}(f_n)(s) & \\
& = & (\bigsqcup_{n\geq 0} \mathcal{R}^{\delta}(f_n))(s) &
\end{array}$$

2. Straightforward by the definition of $\mathcal{R}$.

3. For any $f \in S \to [0,1]$ and $s \in S$ we show that

$$\mathcal{R}(f)(s) = (\bigsqcup_{n\in\mathbb{N}} \mathcal{R}^{\delta_n})(f)(s). \tag{18}$$

We focus on the non-trivial case that $s \xrightarrow{\alpha} \Delta$ for some action α and distribution $\Delta \in \mathcal{D}_1(S)$.

$$\begin{array}{lcl}
(\bigsqcup_{n\in\mathbb{N}} \mathcal{R}^{\delta_n})(f)(s) & = & \bigsqcup_{n\in\mathbb{N}} \mathcal{R}^{\delta_n}(f)(s) \\
& = & \bigsqcup_{n\in\mathbb{N}} \delta_n \cdot f(\Delta) \\
& = & f(\Delta) \cdot (\bigsqcup_{n\in\mathbb{N}} \delta_n) \\
& = & f(\Delta) \cdot 1 \\
& = & \mathcal{R}(f)(s)
\end{array}$$

□

Lemma 4.20 Let $\{\delta_n\}_{n\geq 1}$ be a nondecreasing sequence of discount factors converging to 1.

- $\mathbb{V} = \bigsqcup_{n\in\mathbb{N}} \mathbb{V}^{\delta_n}$
- $\mathbb{V}_{\max} = \bigsqcup_{n\in\mathbb{N}} \mathbb{V}_{\max}{}^{\delta_n}$

Proof: We only consider $\mathbb{V}$; the case for $\mathbb{V}_{\max}$ 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 $\mathcal{R}$ and $\mathcal{R}^{\delta_n}$ respectively, so we need to prove that

$$\mathit{lfp}(\mathcal{R}) = \bigsqcup_{n\in\mathbb{N}} \mathit{lfp}(\mathcal{R}^{\delta_n}) \tag{19}$$

We now show two inequations.

For any $n \in \mathbb{N}$, we have $\delta_n \leq 1$, so Lemma 4.19 (2) yields $\mathcal{R}^{\delta_n} \leq \mathcal{R}$. It follows that $\mathit{lfp}(\mathcal{R}^{\delta_n}) \leq \mathit{lfp}(\mathcal{R})$, thus $\bigsqcup_{n\in\mathbb{N}} \mathit{lfp}(\mathcal{R}^{\delta_n}) \leq \mathit{lfp}(\mathcal{R})$.

For the other direction, $\mathit{lfp}(\mathcal{R}) \leq \bigsqcup_{n\in\mathbb{N}} \mathit{lfp}(\mathcal{R}^{\delta_n})$, it suffices to show that $\bigsqcup_{n\in\mathbb{N}} \mathit{lfp}(\mathcal{R}^{\delta_n})$ is a pre-fixed point of $\mathcal{R}$, i.e. $\mathcal{R}(\bigsqcup_{n\in\mathbb{N}} \mathit{lfp}(\mathcal{R}^{\delta_n})) \leq \bigsqcup_{n\in\mathbb{N}} \mathit{lfp}(\mathcal{R}^{\delta_n})$, which we derive as follows. Let $\{\delta_n\}_{n\geq 1}$ be a nondecreasing sequence of

discount factors converging to 1.

$$\begin{array}{lll} & \mathcal{R}(\bigsqcup_{n\in\mathbb{N}} lfp(\mathcal{R}^{\delta_n})) & \\ = & (\bigsqcup_{m\in\mathbb{N}} \mathcal{R}^{\delta_m})(\bigsqcup_{n\in\mathbb{N}} lfp(\mathcal{R}^{\delta_n})) & \text{by Lemma 4.19 (3)} \\ = & \bigsqcup_{m\in\mathbb{N}} \mathcal{R}^{\delta_m}(\bigsqcup_{n\in\mathbb{N}} lfp(\mathcal{R}^{\delta_n})) & \\ = & \bigsqcup_{m\in\mathbb{N}} \bigsqcup_{n\in\mathbb{N}} \mathcal{R}^{\delta_m}(lfp(\mathcal{R}^{\delta_n})) & \text{by Lemma 4.19 (1)} \\ = & \bigsqcup_{m\in\mathbb{N}} \bigsqcup_{n\geq m} \mathcal{R}^{\delta_m}(lfp(\mathcal{R}^{\delta_n})) & \\ \leq & \bigsqcup_{m\in\mathbb{N}} \bigsqcup_{n\geq m} \mathcal{R}^{\delta_n}(lfp(\mathcal{R}^{\delta_n})) & \text{by Lemma 4.19 (2)} \\ = & \bigsqcup_{n\in\mathbb{N}} \mathcal{R}^{\delta_n}(lfp(\mathcal{R}^{\delta_n})) & \\ = & \bigsqcup_{n\in\mathbb{N}} lfp(\mathcal{R}^{\delta_n}) & \end{array}$$

This completes the proof of (19). □

In the rest of this section, we consider probabilistic automata with actions τ and ω only.

Lemma 4.21 Suppose $\delta \in (0, 1]$. If $\langle T, \Theta^\circ, \rightarrow\rangle$ is a resolution of $\langle S, \Delta^\circ, \rightarrow\rangle$, then we have $\mathbb{V}^{\delta,n}(\Theta^\circ) \leq \mathbb{V}_{\max}{}^{\delta,n}(\Delta^\circ)$.

Proof: Let $f : T \rightarrow S$ be the resolving function associated with the resolution $\langle T, \Theta^\circ, \rightarrow\rangle$, we show by induction on n that

$$\mathbb{V}_{\max}{}^{\delta,n}(f(t)) \geq \mathbb{V}^{\delta,n}(t) \text{ for any } t \in T \tag{20}$$

The base case $n = 0$ is trivial. We consider the inductive step. If $t \xrightarrow{\omega} \Theta$, then $f(t) \xrightarrow{\omega} f(\Theta)$, thus $\mathbb{V}_{\max}{}^{\delta,n}(f(t)) = 1 = \mathbb{V}^{\delta,n}(t)$. Now suppose $t \xrightarrow{\tau} \Theta$. Then $f(t) \not\xrightarrow{\omega}$ and $f(t) \xrightarrow{\tau} f(\Theta)$. We can infer that

$$\begin{array}{lll} \mathbb{V}_{\max}{}^{\delta,(n+1)}(f(t)) & = & \delta \cdot max\{\mathbb{V}_{\max}{}^{\delta,n}(\Delta) | f(t) \xrightarrow{\tau} \Delta\} \\ & \geq & \delta \cdot \mathbb{V}_{\max}{}^{\delta,n}(f(\Theta)) \\ & = & \delta \cdot \sum_{s\in S} f(\Theta)(s) \cdot \mathbb{V}_{\max}{}^{\delta,n}(s) \\ & = & \delta \cdot \sum_{t'\in T} \Theta(t') \cdot \mathbb{V}_{\max}{}^{\delta,n}(f(t')) \\ & \geq & \delta \cdot \sum_{t'\in T} \Theta(t') \cdot \mathbb{V}^{\delta,n}(t') \quad \text{by induction} \\ & = & \delta \cdot \mathbb{V}^{\delta,n}(\Theta) \\ & = & \mathbb{V}^{\delta,(n+1)}(t) \end{array}$$

So we have proved (20), from which it follows that

$$\mathbb{V}_{\max}{}^{\delta}(f(t)) \geq \mathbb{V}^{\delta}(t) \text{ for any } t \in T \tag{21}$$

Therefore, we have that

$$\begin{array}{lll} \mathbb{V}_{\max}{}^{\delta}(\Delta^\circ) & = & \mathbb{V}_{\max}{}^{\delta}(f(\Theta^\circ)) \\ & = & \sum_{s\in S} f(\Theta^\circ)(s) \cdot \mathbb{V}_{\max}{}^{\delta}(s) \\ & = & \sum_{t\in T} \Theta^\circ(t) \cdot \mathbb{V}_{\max}{}^{\delta}(f(t)) \\ & \geq & \sum_{t\in T} \Theta^\circ(t) \cdot \mathbb{V}^{\delta}(t) \quad \text{by (21)} \\ & = & \mathbb{V}^{\delta}(\Theta^\circ) \end{array}$$

□

We say a resolution of a process is *static* if its associated resolving function is injective.

Lemma 4.22 Suppose $\delta < 1$. Given a probabilistic automaton $\langle S, \Delta^\circ, \rightarrow\rangle$, there exists a static resolution $\langle T, \Theta^\circ, \rightarrow\rangle$ such that $\mathbb{V}_{\max}{}^{\delta}(\Delta^\circ) = \mathbb{V}^{\delta}(\Theta^\circ)$.

Proof: Let $\langle T, \Theta^\circ, \rightarrow\rangle$ be a resolution with an injective resolving function f such that if $t \xrightarrow{\tau} \Theta$ then $\mathbb{V}_{\max}{}^{\delta}(f(\Theta)) = max\{\mathbb{V}_{\max}{}^{\delta}(\Delta) \mid f(t) \xrightarrow{\tau} \Delta\}$. The finite-branching assumption ensures the existence of the such resolving function f.

Let $g : T \to [0,1]$ be the function defined by $g(t) = \mathbb{V}_{\max}{}^{\delta}(f(t))$ for all $t \in T$. Below we show that g is a fixed point of $\mathcal{R}^{\delta}$. If $t \xrightarrow{\omega}$ then $f(t) \xrightarrow{\omega}$. Therefore, $\mathcal{R}^{\delta}(g)(t) = 1 = \mathbb{V}_{\max}{}^{\delta}(f(t)) = g(t)$. Now suppose $t \xrightarrow{\tau} \Theta$. By the definition of f, we have $f(t) \not\xrightarrow{\omega}$, $f(t) \xrightarrow{\tau} f(\Theta)$ with $\mathbb{V}_{\max}{}^{\delta}(f(\Theta)) = max\{\mathbb{V}_{\max}{}^{\delta}(\Delta) \mid f(t) \xrightarrow{\tau} \Delta\}$. Therefore,

$$\begin{array}{rcl}
\mathcal{R}^{\delta}(g)(t) & = & \delta \cdot g(\Theta) \\
& = & \delta \cdot \sum_{t \in T} \Theta(t) \cdot g(t) \\
& = & \delta \cdot \sum_{t \in T} \Theta(t) \cdot \mathbb{V}_{\max}{}^{\delta}(f(t)) \\
& = & \delta \cdot \sum_{s \in S} f(\Theta)(s) \cdot \mathbb{V}_{\max}{}^{\delta}(s) \\
& = & \delta \cdot \mathbb{V}_{\max}{}^{\delta}(f(\Theta)) \\
& = & \delta \cdot max\{\mathbb{V}_{\max}{}^{\delta}(\Delta) | f(t) \xrightarrow{\tau} \Delta\} \\
& = & \mathbb{V}_{\max}{}^{\delta}(f(t)) \\
& = & g(t)
\end{array}$$

Since $\mathcal{R}^{\delta}$ has a unique fixed point $\mathbb{V}^{\delta}$, we derive that g coincides with $\mathbb{V}^{\delta}$, i.e. $\mathbb{V}^{\delta}(t) = g(t) = \mathbb{V}_{\max}{}^{\delta}(f(t))$ for all $t \in T$, from which we can obtain the required result $\mathbb{V}^{\delta}(\Theta^{\circ}) = \mathbb{V}_{\max}{}^{\delta}(\Delta^{\circ})$. □

Theorem 4.23 Given a finite-state probabilistic automaton $\langle S, \Delta^{\circ}, \rightarrow\rangle$, there exists a static resolution $\langle T, \Theta^{\circ}, \rightarrow\rangle$ such that $\mathbb{V}_{\max}(\Delta^{\circ}) = \mathbb{V}(\Theta^{\circ})$.

Proof: By Lemma 4.22, for every discount factor $d \in (0,1)$ there exists a static resolution which achieves the maximum probability of success. Since there are finitely many states in S, there are finitely many 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 $\langle T, \Theta^{\circ}, \longrightarrow\rangle$ with resolving function f_0 such that $\mathbb{V}^{\delta_{n_k}}(t) = \mathbb{V}_{\max}{}^{\delta_{n_k}}(f_0(t))$ for all $t \in T$ and $k = 1, 2, \ldots$. By Lemma 4.20, we have that, for every $t \in T$,

$$\begin{array}{rcl}
\mathbb{V}(t) & = & \bigsqcup_{k \in \mathbb{N}} \mathbb{V}^{\delta_{n_k}}(t) \\
& = & \bigsqcup_{k \in \mathbb{N}} \mathbb{V}_{\max}{}^{\delta_{n_k}}(f_0(t)) \\
& = & \mathbb{V}_{\max}(f_0(t))
\end{array}$$

It follows that $\mathbb{V}(\Theta^{\circ}) = \mathbb{V}_{\max}(\Delta^{\circ})$. □

Along the same line, we can obtain a similar theorem for $\mathbb{V}_{\min}$. However, the use of discount factors is not necessary in this case.

Lemma 4.24 If $\langle T, \Theta^{\circ}, \rightarrow\rangle$ is a resolution of $\langle S, \Delta^{\circ}, \rightarrow\rangle$, then it holds that $\mathbb{V}_{\min}(\Delta^{\circ}) \leq \mathbb{V}(\Theta^{\circ})$.

Proof: Analogous to the proof of Lemma 4.21. □

Theorem 4.25 Given a finite-state probabilistic automaton $\langle S, \Delta^{\circ}, \longrightarrow\rangle$. There exists a static resolution $\langle T, \Theta^{\circ}, \rightarrow\rangle$ such that $\mathbb{V}_{\min}(\Delta^{\circ}) = \mathbb{V}(\Theta^{\circ})$.

Proof: Similar to the proof of Lemma 4.22. Let $\langle T, \Theta^{\circ}, \rightarrow\rangle$ be a resolution with an injective resolving function f such that if $t \xrightarrow{\tau} \Theta$ then $\mathbb{V}_{\min}(f(\Theta)) = min\{\mathbb{V}_{\min}(\Delta) \mid f(t) \xrightarrow{\tau} \Delta\}$.

Let $g : T \to [0,1]$ be the function defined by $g(t) = \mathbb{V}_{\min}(f(t))$. As in the proof of Lemma 4.22, we show that g is a fixed point of $\mathcal{R}$. Since $\mathbb{V}$ is the least fixed point of $\mathcal{R}$, it holds that $\mathbb{V}(t) \leq g(t) = \mathbb{V}_{\min}(f(t))$ for all $t \in T$, from which we can obtain $\mathbb{V}(\Theta^{\circ}) \leq \mathbb{V}_{\min}(\Delta^{\circ})$. Using Lemma 4.24, we derive that $\mathbb{V}_{\min}(\Delta^{\circ}) = \mathbb{V}(\Theta^{\circ})$. □

From Lemmas 4.21 and 4.24 as well as Theorems 4.23 and 4.25, we obtain the following corollary.

Corollary 4.26 Let $M = \langle S, \Delta^{\circ}, \rightarrow\rangle$ be a finitary probabilistic automaton.

$$\begin{array}{rcl}
\mathbb{V}_{\max}(\Delta^{\circ}) & = & max\{\mathbb{V}(\Theta^{\circ}) \mid \Theta^{\circ} \text{ is the initial distribution of a static resolution of } M\} \\
\mathbb{V}_{\min}(\Delta^{\circ}) & = & min\{\mathbb{V}(\Theta^{\circ}) \mid \Theta^{\circ} \text{ is the initial distribution of a static resolution of } M\}
\end{array}$$

The following theorem states that for finitary processes extremal testing yields the same preorders as resolution-based testing.

Theorem 4.27 For finite-state processes:

- $P \sqsubseteq^{e}_{\text{pmay}} Q$ if and only if $P \sqsubseteq_{\text{pmay}} Q$
- $P \sqsubseteq^{e}_{\text{pmust}} Q$ if and only if $P \sqsubseteq_{\text{pmust}} Q$

Proof: An immediate consequence of Corollary 4.26. □

Because of Theorems 4.11 and 4.27, we can choose the variation of the testing preorders which are most convenient to the task at hand. So to prove the soundness of the simulation preorders we will use *extremal* testing, while to prove that modal formulae can be characterised by tests we use the *resolution-based* testing.

4.4 Testing via weak-p derivatives

We now show how resolutions, used in Section 4.2 to determine the outcome of tests, can be realised as the derivatives of a restrictive class of weak-p moves. With this alternative point of view, we can effectively test processes by comparing their possible derivatives directly: in effect the inductive $\mathbb{V}$ is replaced by the induction implicit in the definition of derivative, and the evaluation of ω-move possibilities is done directly on the subdistributions the derivations produce.

Definition 4.28 In a pLTS a state s is called *stable* if $s \stackrel{\tau}{\not\longrightarrow}$, and a subdistribution Θ is called *stable* if every state in its support is stable. We write $\Delta \Longrightarrow\!\!\!\!\!\succ \Delta'$ whenever $\Delta \Longrightarrow \Delta'$ and Δ' is stable, and call Δ' an *extreme* derivative of Δ.

Referring to Definition 3.12, we see this means that in the derivation of Δ' 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: we have $s \in \Delta^{\times}_{k}$ if *and now also* only if $s \stackrel{\tau}{\not\longrightarrow}$.

Lemma 4.29 [Existence 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: The construction of derivatives in Definition 3.12 is simply specialised by choosing at every stage $\Delta^{\times}_{k}$ uniquely to be Δ_k restricted exactly to those states that must stop, i.e. those s for which $s \stackrel{\tau}{\not\longrightarrow}$. Then $\Delta^{\rightarrow}_{k} := \Delta_k - \Delta^{\times}_{k}$ and Δ_{k+1} is chosen arbitrarily so that $\Delta^{\rightarrow}_{k} \stackrel{\tau}{\longrightarrow} \Delta_{k+1}$. That establishes (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. □

It is worth pointing out that the use of subdistributions, rather than distributions, is essential here. For example if Δ diverges, that is if there is an infinite sequence of derivations $\Delta \stackrel{\tau}{\longrightarrow} \Delta_1 \stackrel{\tau}{\longrightarrow} \ldots \Delta_k \stackrel{\tau}{\longrightarrow} \ldots$, then the only extreme derivative of Δ is the empty subdistribution ε.

Lemma 4.30 Let Δ be a subdistribution in a fully probabilistic pLTS. If $\Delta \Longrightarrow\!\!\!\!\!\succ \Delta'$ then $\mathbb{V}^{\Omega}(\Delta) = \mathbb{V}^{\Omega}(\Delta')$.

Proof: Recall that $\mathbb{V}^{\Omega}$ is defined over deterministic pLTSs (only), and that $\mathbb{V} = \bigsqcup_{n\geq 0} \mathbb{V}^{n}$ because its defining functional is continuous, where we are now (in this proof) dropping the superscript $\cdot^{\Omega}$ to reduce clutter: the $\mathbb{V}^{n}$'s are the n^{th} approximants to $\mathbb{V}$. By inspection we have that $s \stackrel{\tau}{\longrightarrow} \Delta$ implies $\mathbb{V}^{n+1}(s) = \mathbb{V}^{n}(\Delta)$, whence by lifting and linearity we get

$$\text{If } \Delta \stackrel{\tau}{\longrightarrow} \Delta' \text{ then } \mathbb{V}^{n+1}(\Delta) = \mathbb{V}^{n}(\Delta') \text{ for all } n \geq 0. \tag{22}$$

Now suppose $\Delta \Longrightarrow\!\!\!\!\succ \Delta'$. Referring to Definition 3.12 and carrying out a straightforward induction based on (22), we have

$$\mathbb{V}^{n+1}(\Delta) \;=\; \mathbb{V}^0(\Delta_{n+1}) + \sum_{k=0}^{n} \mathbb{V}^{n-k+1}(\Delta_k^{\times}) \;=\; \sum_{k=0}^{n} \mathbb{V}^{n-k+1}(\Delta_k^{\times}) \tag{23}$$

for all $n \geq 0$, with the second step depending on $\mathbb{V}^0$'s being identically $\vec{0}$.

Now because Δ' is an extreme derivative we know that all the $\Delta_k^{\times}$'s are stable, whence immediately for each $k \leq n$ we have $\mathbb{V}^{n-k+1}(\Delta_k^{\times}) = \mathbb{V}(\Delta_k^{\times})$, simplifying the last step above to just $\sum_{k=0}^{n} \mathbb{V}(\Delta_k^{\times})$. We conclude by reasoning

$$\begin{array}{lll}
 & \mathbb{V}(\Delta) \;=\; \bigsqcup_{n\geq 0} \mathbb{V}^n(\Delta) \;=\; \bigsqcup_{n\geq 0} \mathbb{V}^{n+1}(\Delta) & \\
= & \bigsqcup_{n\geq 0} \sum_{k=0}^{n} \mathbb{V}(\Delta_k^{\times}) & \text{(23) and immediately above} \\
= & \bigsqcup_{n\geq 0} \mathbb{V}(\sum_{k=0}^{n} \Delta_k^{\times}) & \text{finite linearity of } \mathbb{V} \\
= & \mathbb{V}(\bigsqcup_{n\geq 0} \sum_{k=0}^{n} \Delta_k^{\times}) & \text{least fixed point of continuous functional is itself continuous} \\
= & \mathbb{V}(\sum_{k=0}^{\infty} \Delta_k^{\times}) & \\
= & \mathbb{V}(\Delta') \,. &
\end{array}$$

□

Now according to (4) and Definition 4.5, the outcome of applying a test to a process is determined by the values $\mathbb{V}(\Theta)$ for distributions Θ over particular deterministic pLTSs, namely resolutions. Then determinacy and Lemma 4.30 immediately above ensures that $\mathbb{V}(\Theta)$ coincides with $\mathbb{V}(\Theta_{\succ})$, where $\Theta_{\succ}$ is the unique extreme derivative of Θ. Moreover because of its stability the calculation of $\mathbb{V}(\Theta_{\succ})$ is simply the weighted sum of all the successful states in its support.

The final link between the resolution- and derivation views of testing is that, in an ω-respecting computation structure, resolutions and extreme derivatives are essentially the same thing.

Theorem 4.31 (Resolutions and extreme derivatives) Consider an ω-respecting computation structure $\langle S, \Omega_\tau, \rightarrow\rangle$. Then $\Delta \Longrightarrow\!\!\!\!\succ \Delta'$ if and only if there is a resolution $\langle T, \Omega_\tau, \rightarrow_T\rangle$ with resolving function f and subdistributions Θ, Θ': $\mathcal{D}(T)$ such that

(i) $f(\Theta), f(\Theta') = \Delta, \Delta'$

(ii) $\Theta \Longrightarrow\!\!\!\!\succ_T \Theta'$.

Proof: For only if, suppose $\langle T, \Omega_\tau, \rightarrow_T\rangle$ is a resolution of $\langle S, \Omega_\tau, \rightarrow\rangle$, let Θ be a subdistribution over T, and suppose $\Theta \Longrightarrow\!\!\!\!\succ_T \Theta'$. By linearity we can apply f to that whole derivation structure, preserving its validity and establishing $f(\Theta) \Longrightarrow_T f(\Theta')$.

Now we observe that each state in the support of $f(\Theta')$ is stable: consider any state $t \in \lceil \Theta_k^{\times} \rceil$ for any $k \geq 0$. Since t is stable by assumption, either t is a deadlock state or $t \xrightarrow{\omega} \Gamma$ for some success action $\omega \in \Omega$ and subdistribution Γ. By Definition 4.4, it must be the case that $f(t)$ is a deadlock state in the first case, and that $f(t) \xrightarrow{\omega} f(\Gamma)$ in the second. Additionally, since $\langle S, \mathsf{Act}, \rightarrow\rangle$ is a "non-scooting" computation structure and $f(t) \in S$, we have $f(t) \stackrel{\tau}{\not\rightarrow}$ in the second case. Therefore, $f(t)$ is stable in both cases. Thus, we have in fact $f(\Theta) \Longrightarrow_T f(\Theta')$.

For if, consider an extreme derivation $\Delta \Longrightarrow\!\!\!\!\succ \Delta'$, as given in Definition 3.12 where all $\Delta_k^{\times}$ are assumed to be stable; for convenience we let Δ_k denote $\Delta_k^{\rightarrow} + \Delta_k^{\times}$. To define the corresponding resolution $\langle T, \Omega_\tau, \longrightarrow_T\rangle$ we refer to Definition 4.4. First let the set of states T be $S \times \mathbb{N}$ and the resolving function $f\!:T \rightarrow S$ be given by $f(s,k) = s$. To complete the description we must define the two partial functions $\xrightarrow{\alpha}$, for $\alpha = \omega$ and $\alpha = \tau$. These are always defined so that if $(s,k) \xrightarrow{\alpha} \Gamma$ then the only states in the support of Γ are of the form $(s, k+1)$. In the definition we use $\Theta^{\downarrow k}$, for any subdistribution Θ over S, to be the subdistribution over T given by

$$\Theta^{\downarrow k}(t) = \begin{cases} \Theta(s) & \text{if } t = (s,k) \\ 0 & \text{otherwise} \end{cases}$$

Note that by definition

(a) $f(\Theta^{\downarrow k}) = \Theta$

(b) $\Delta_k^{\downarrow k} = \Delta_k^{\rightarrow\downarrow k} + \Delta_k^{\times\downarrow k}$

The definition of $\xrightarrow{\omega}_T$ is straightforward: its domain consists of states (s,k) where $s \in \lceil \Delta_k^{\times} \rceil$ and is defined by letting $(s,k) \xrightarrow{\omega} \Delta_s^{\downarrow k+1}$ for some arbitrarily chosen $s \xrightarrow{\omega} \Delta_s$.

The definition of $\xrightarrow{\tau}_T$ is more complicated, and is determined by the moves $\Delta_k^{\rightarrow} \xrightarrow{\tau} \Delta_{k+1}$. For a given k this move means that

$$\Delta_k^{\rightarrow} = \sum_{i\in I} p_i \cdot \overline{s_i}, \qquad \Delta_{k+1} = \sum_{i\in I} p_i \cdot \Gamma_i, \qquad s_i \xrightarrow{\tau} \Gamma_i$$

So for each k we let

$$(s,k) \xrightarrow{\tau}_T \sum_{s_i = s} p_i \cdot \Gamma_i^{\downarrow k+1}$$

This definition ensures

1. $(\Delta_k^{\rightarrow})^{\downarrow k} \xrightarrow{\tau}_T (\Delta_{k+1})^{\downarrow k+1}$
2. $(\Delta^{\times})^{\downarrow k}$ is stable.

This completes our definition of the resolution; it remains to find distributions Θ, Θ' over T such that $f(\Theta) = \Delta$, $f(\Theta') = \Delta'$ and $\Theta \Longrightarrow \Theta'$.

Because of (a) (c) and (d) we have the following extreme derivation, which by Part (ii) of Lemma 4.29 is the unique one from $\Delta_0^{\downarrow 0}$:

$$\begin{array}{lcll}
\Delta^{\downarrow 0} & = & (\Delta_0^{\rightarrow})^{\downarrow 0} + (\Delta_0^{\times})^{\downarrow 0} \\
(\Delta_0^{\rightarrow})^{\downarrow 0} & \xrightarrow{\tau}_T & (\Delta_1^{\rightarrow})^{\downarrow 1} + (\Delta_1^{\times})^{\downarrow 1} \\
\vdots & & \vdots \\
(\Delta_k^{\rightarrow})^{\downarrow k} & \xrightarrow{\tau}_T & (\Delta_{k+1}^{\rightarrow})^{\downarrow k+1} + (\Delta_{k+1}^{\times})^{\downarrow k+1} \\
 & & \vdots \\
 & & \hline \\
 & & \Theta' = \sum_{k=0}^{\infty} (\Delta_k^{\times})^{\downarrow k}
\end{array}$$

Letting Θ be $\Delta^{\downarrow 0}$, we see that Note (a) above ensures $f(\Theta) = \Delta$; the same note and the linearity of f applied to distributions also gives $f(\Theta') = \Delta'$.

□

Corollary 4.32 In an ω-respecting computation structure, the following statements hold.

1. If $\Delta \Longrightarrow \Delta'$ then there is a resolution Θ of Δ such that $\mathbb{V}^{\Omega}(\Theta) = \mathbb{V}^{\Omega}(\Delta')$.
2. For any resolution Θ of Δ, there exists an extreme derivative Δ' such that $\Delta \Longrightarrow \Delta'$ and $\mathbb{V}^{\Omega}(\Theta) = \mathbb{V}^{\Omega}(\Delta')$.

Proof: Suppose $\Delta \Longrightarrow \Delta'$. By Theorem 4.31, there is a resolution of Δ with resolving function f and a subdistribution Θ such that $\Theta \Longrightarrow \Theta'$ and $f(\Theta') = \Delta'$ and $f(\Theta) = \Delta$; this last statement means that by definition Θ is a resolution of Δ. By Lemma 4.30, we have $\mathbb{V}^{\Omega}(\Theta) = \mathbb{V}^{\Omega}(\Theta')$. Since Θ' and Δ' are extreme derivatives, all the states in their supports are stable. Therefore, for any $t \in \lceil \Theta' \rceil$ we have that (i) $t \xrightarrow{\omega}$ with $\omega \in \Omega$ iff $f(t) \xrightarrow{\omega}$, (ii) $t \not\rightarrow$ iff $f(t) \not\rightarrow$. It follows that $\mathbb{V}^{\Omega}(\Theta') = \mathbb{V}^{\Omega}(\Delta')$. As a result, we obtain that $\mathbb{V}^{\Omega}(\Theta) = \mathbb{V}^{\Omega}(\Delta')$.

To prove 2, suppose Θ is a resolution of Δ; that is there is a resolution as in Definition 4.4 with a resolving function f such that $f(\Theta) = \Theta'$. We know from Lemma 4.29 that there exists a (unique) subdistribution Θ' such that $\Theta \Longrightarrow \Theta'$. By Theorem 4.31 we have that $\Delta = f(\Theta) \Longrightarrow f(\Theta')$. The same arguments as in the other direction show that $\mathbb{V}^{\Omega}(\Theta) = \mathbb{V}^{\Omega}(f(\Theta'))$. □

Definition 4.33 Let Δ be a subdistribution in a computation structure. We write $\mathcal{V}^{\Omega}(\Delta)$ for the set of testing outcomes $\{\mathbb{V}^{\Omega}(\Gamma) \mid \Gamma \text{ is a resolution of } \Delta\}$ obtained from Δ.

Lemma 4.34 Let Δ be a subdistribution in an ω-respecting computation structure.
Then $\mathcal{V}^{\Omega}(\Delta) = \{\mathbb{V}^{\Omega}(\Delta') \mid \Delta \Longrightarrow\!\!\!\!\succ \Delta'\}$.

Proof: This is immediate from Corollary 4.32. □

Finally we have the basis of a derivation-style definition of testing.

Lemma 4.35 Let P be a process and T a compatible test.
Then $\mathcal{A}^{\Omega}(T, P) = \{\mathbb{V}^{\Omega}(\Delta') \mid [\![T \mid_{\mathsf{Act}} P]\!] \Longrightarrow\!\!\!\!\succ \Delta'\}$.

Proof: This is immediate from (4) and Lemma 4.34. □

5 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 the set of derivations from a single starting point is *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 distribuitions in a straightforward way. The other property is "distillation of divergence," allowing us to 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.

Both properties depend on our working within *finitary* pLTSs — that is, ones in which the state space is finite and the (unlifted) transition relation is finite-branching. (The examples of Appendix A show this is necessary, for our approach at least.) Again, some proofs and technical lemmas are deferred to Appendix B.

5.1 Finite generability and closure

Now let us restrict our attention to finitary pLTSs. Here by definition the sets $s\cdot \xrightarrow{a}$ are finite, for every s and a. This of course is no longer true for the lifted relations $\xrightarrow{a}$; nevertheless the sets $\overline{s}\cdot \xrightarrow{a}$ can be finitely represented.

Definition 5.1 A resolution R is said to be *static* if its resolving function f_R is injective. This means that the transition used at a state is always the same one, no matter how often or from where the state is reached, and that it is *pure*, that is not interpolated.

Static resolutions are particularly easy to describe because they are given effectively by subsets of the transition relation of the original pLTS. For us a convenient formulation of this is in terms of the following definition.

Definition 5.2 A *static derivative policy* for a computation structure $\langle S, \mathsf{Act}, \rightarrow\rangle$, or an *SDP*, is a partial function $\mathsf{pp} : S \rightharpoonup \mathcal{D}_1(S)$ satisfying:

- If pp is defined at s then $s \xrightarrow{\tau} \mathsf{pp}(s)$, and
- If $s \not\xrightarrow{\tau}$ then pp is undefined at s.

Intuitively a policy pp decides for each state, once and for all, which of the available τ-choices to take if any: since it chooses a specific transition, or inaction, it does not interpolate; and since it is a function of the state, it makes the same choice on every visit.

There is a close relationship between *SDP*'s and static resolutions. One difference is that (here) we are concerned only with policies for τ transitions (although clearly the idea generalises). The other difference is that an *SDP* can select a "do not take any τ-transition" option, even if some are available, which it does by setting pp to be undefined at s even when s enables some transition. On the other hand, a static resolution must take a transition if it can. In this respect *SDP*'s are close to derivations, which also have the "stopping-here" option.

The great importance for us of *SDP*'s is that they give a particularly simple characterisation of derivatives, provided the state-space is finite and the pLTS is finite-branching. This is essentially a result of Markov Decision Processes [19], which we now translate into our context.

Lemma 5.3 ($\Longrightarrow$ realised by interpolation of finitely many static policies) Suppose $s \Longrightarrow \Delta'$ for some state s and subdistribution Δ'. Then there is a finite index set I, probabilities p_i summing to 1 and static strategies pp_i such that $\Delta' = \sum_{i \in I} p_i \cdot \Delta'_i$ where uniquely $s \Longrightarrow_{\mathsf{pp}_i} \Delta'_i$ for each i.

Proof: See Lemma B.20 and its preceding material. □

Lemma 5.4 (Closure of $\Longrightarrow$) For any state s the set $(s \Longrightarrow)$ of derivatives of s is closed.

Proof: From Lemma 5.3 that set is a subset of the convex closure of the set of subdistributions reachable via $\Longrightarrow$ using one of the finitely many static policies of the pLTS, and that latter set is closed since it is the convex closure of finitely many points. But each of those subdistributions is trivially a derivative itself; and the set of derivatives is convex by Corollary 3.20. Hence the two sets are equal, and thus the former is closed as well. □

Lemma 5.5 [Closure of $\stackrel{a}{\Longrightarrow}$] For any state s the set $\{\Delta' \mid s \stackrel{a}{\Longrightarrow} \Delta'\}$ is closed.

Proof: The relation $\stackrel{a}{\Longrightarrow}$ is a composition of three stages: there must be Δ'_1, Δ'_2 with $s \Longrightarrow \Delta'_1 \stackrel{a}{\longrightarrow} \Delta'_2 \Longrightarrow \Delta'$. For the first stage, from Lemma 5.3 we know any such Δ'_1 is the interpolation of fixed finite set of (other) subdistributions; call them "principal" for that starting point (s) and type of transition ($\Longrightarrow$).

For each state s_1 in the support of some principal subdistribution $\widehat{\Delta}_1$ from the first stage, there are only a finite number of distributions in $(s_1 \stackrel{a}{\longrightarrow})$, because of finite branching. Because $\lceil \widehat{\Delta}_1 \rceil$ itself is finite, there are thus only finitely many subdistributions necessary to generate all of $(\widehat{\Delta}_1 \stackrel{a}{\longrightarrow})$; and because there are only finitely many $\widehat{\Delta}_1$'s, we still need only a finite number of principal subdistributions $\widehat{\Delta}_2$ to generate the results of the first and second stages.

For the third stage we make effectively the same argument as for the second, except that instead of appealing to finite branching of $(\widehat{\Delta}_2 \Longrightarrow)$ instead we use the finite generability that we appealed to in the first stage.

Then Lemma 5.4 applies (analogously) for closure. □

Lemma 5.6 [Zero-one law, static case] If for some static derivative policy pp over a finite-state pLTS there is for some s a derivation $s \Longrightarrow_{\mathsf{pp}} \Delta'$ with $|\Delta'| < 1$ then in fact for some (possibly different) state s_ε we have $s_\varepsilon \Longrightarrow_{\mathsf{pp}} \varepsilon$.

Proof: Suppose that for no state s do we have $s \Longrightarrow_{\mathsf{pp}} \varepsilon$. This means that for every state s there is some other state s' (possibly depending on s), some number $N_s \geq 0$ (no matter how large) and probability $p_s > 0$ (no matter how small) such that s can follow policy pp to reach s' in N_s transitions $\stackrel{\tau}{\longrightarrow}$ with aggregated probability p_s and then can go no further because pp is undefined at s' — for if this were not true, policy pp would generate an unbounded τ-tree of transitions rooted at s with aggregate probability 1.

Set $N \geq 0$ to be the maximum of N_s over all states, and $p > 0$ to be the minimum of p_s over all states. (That $p > 0$ is one place we use finiteness of the state space; the other is in the existence of both the N_s's in the first place, and of their maximum.)

Now pick an arbitrary s and consider the (unique) derivation $s \Longrightarrow_{\mathsf{pp}} \Delta'$. If it is finite, then $|\Delta'|{=}1$ trivially. If not, group the infinite sequence of pp-generated $\stackrel{\tau}{\longrightarrow}_{\mathsf{pp}}$ moves into blocks of N. In any such block the probability of reaching a pp-undefined state is at least $p > 0$, and so the probability of eventually reaching a pp-undefined state (by taking many N-blocks of moves in succession) is in fact 1. Thus $s \Longrightarrow_{\mathsf{pp}} \Delta'$ implies $|\Delta'| = 1$.

Our lemma then is immediate from the contrapositive. □

5.2 Distillation of divergence

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 5.7 [Distillation of divergence, static case]

If for some state s and static derivative policy pp over a finite-state pLTS there is a derivation $s \Longrightarrow_{\mathsf{pp}} \Delta'$ then there is a probability p and full distributions $\Delta'_1, \Delta'_\varepsilon$ such that $s \Longrightarrow (\Delta'_1 \,{}_p\!\oplus \Delta'_\varepsilon)$ and $\Delta' = p \cdot \Delta'_1$ and $\Delta'_\varepsilon \Longrightarrow \varepsilon$.

Proof: See Lemma B.21. □

Lemma 5.8 [Distillation of divergence, general case] For any s, Δ' in a finitary pLTS with $s \Longrightarrow \Delta'$ there is a probability p and full distributions $\Delta'_1, \Delta'_\varepsilon$ such that $s \Longrightarrow (\Delta'_1 \,{}_p\!\oplus \Delta'_\varepsilon)$ and $\Delta' = p \cdot \Delta'_1$ and $\Delta'_\varepsilon \Longrightarrow \varepsilon$.

Proof: From Lemma 5.3 we know that the derivation $s \Longrightarrow \Delta'$ is an interpolation of a finite number of static derivations. Lemma 5.7 then applies to each separately, and the result follows by interpolating the derivations from each static case. □

Corollary 5.9 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 $s' \Longrightarrow \varepsilon$. 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 Lemma 5.8 since any $s' \in \lceil \Delta'_\varepsilon \rceil$ will do. The general result is obtained by dividing the given derivation by $|\Delta|$. □

6 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 coinductive 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 w.r.t. the must testing preorder, are in terms in 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 7.

6.1 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 [6, 2].

Definition 6.1 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 $\Delta \Longrightarrow \Delta^{\text{pre}} \stackrel{a}{\longrightarrow} \Delta^{\text{post}} \Longrightarrow \Delta'$. 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 $\stackrel{\tau}{\longrightarrow}$).
- 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 3.15 we have $[\![Q_1]\!] \stackrel{a}{\Longrightarrow} [\![\mathbf{0}]\!]$, while in Example 3.16 we have $[\![s_2]\!] \stackrel{a}{\Longrightarrow} \frac{1}{2}[\![\mathbf{0}]\!]$ as well as $[\![s_2]\!] \Longrightarrow \stackrel{B}{\not\longrightarrow}$ for any set B not containing a, because $s_2 \Longrightarrow \frac{1}{2}[\![a]\!]$.

Proposition 6.2 The relation $\stackrel{a}{\Longrightarrow}$, for $a \in \mathsf{Act}$, can be obtained as a lifting.

Proof: Relation $\stackrel{a}{\Longrightarrow}$ is in fact $\Longrightarrow \overline{\stackrel{a}{\longrightarrow}} \Longrightarrow$. By Proposition 3.19 this equals $\overline{\Longrightarrow_S} \; \overline{\stackrel{a}{\longrightarrow}} \; \overline{\Longrightarrow_S}$ which, by three applications of Lemma 3.11, equals $\overline{\Longrightarrow_S} \; \overline{\stackrel{a}{\longrightarrow}} \; \overline{\Longrightarrow_S}$ hence $\Longrightarrow_S \overline{\stackrel{a}{\longrightarrow}} \overline{\Longrightarrow_S}$ and finally $\Longrightarrow_S \overline{\stackrel{a}{\longrightarrow}} \Longrightarrow$. □

Corollary 6.3 The relation $\stackrel{a}{\Longrightarrow}$ is convex.

Proof: This is immediate from its being a lifting. □

Definition 6.4 (Failure Simulation Preorder) Define $\sqsupseteq_{FS}$ to be the largest relation in $\mathcal{D}(S) \times \mathcal{D}(S)$ 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 $\Theta'_i \in \mathcal{D}(S)$ with $\Theta \stackrel{\alpha}{\Longrightarrow} (\sum_i p_i \Theta'_i)$ and $\Delta'_i \sqsupseteq_{FS} \Theta'_i$ for each i, and
2. whenever $\Delta \Longrightarrow \not\stackrel{A}{\longrightarrow}$ then also $\Theta \Longrightarrow \not\stackrel{A}{\longrightarrow}$.

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 [5]. This orientation also matches the one used in CSP [9] 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 8. 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 6.5 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 6.4 to conclude that also $\Theta \Longrightarrow \varepsilon$. □

Although the regularity of Definition 6.4 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 6.6 (Failure Similarity) Define $\lhd_{FS}$ to be the largest relation in $S \times \mathcal{D}(S)$ 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 $\Theta' \in \mathcal{D}(S)$ with $\Theta \stackrel{\alpha}{\Longrightarrow} \Theta'$ and $\Delta' \; \overline{\lhd_{FS}} \; \Theta'$, and
2. whenever $\overline{s} \Longrightarrow \not\stackrel{A}{\longrightarrow}$ then $\Theta \Longrightarrow \not\stackrel{A}{\longrightarrow}$.

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

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

Lemma 6.7 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 3.4 that $\quad \Delta = \sum_{i \in I} p_i \cdot \overline{s_i}, \qquad s_i \lhd_{FS} \Theta_i, \qquad \Theta = \sum_{i \in I} p_i \cdot \Theta_i.$

By Propositions 6.2 and 3.9 we know from $\Delta \stackrel{\alpha}{\Longrightarrow} \Delta'$ that $\overline{s_i} \stackrel{\alpha}{\Longrightarrow} \Delta'_i$ for $\Delta'_i \in \mathcal{D}(S)$ 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 $\Theta'_i \in \mathcal{D}(S)$ 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 3.18(i) yield $\Delta' \; \overline{\lhd_{FS}} \; \Theta'$ and $\Theta \stackrel{\alpha}{\Longrightarrow} \Theta'$. □

Lemma 6.8 Suppose $\Delta \; \overline{\lhd_{FS}} \; \Theta$ and $\Delta \Longrightarrow \not\stackrel{A}{\longrightarrow}$. Then $\Theta \Longrightarrow \not\stackrel{A}{\longrightarrow}$.

Proof: Suppose $\Delta \; \overline{\lhd_{FS}} \; \Theta$ and $\Delta \Longrightarrow \Delta' \not\stackrel{A}{\longrightarrow}$. By Lemma 6.7 there exists some Θ' such that $\Theta \Longrightarrow \Theta'$ and $\Delta' \; \overline{\lhd_{FS}} \; \Theta'$. From Lemma 3.4 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 $\Delta' \not\stackrel{A}{\longrightarrow}$, we have that $s_i \not\stackrel{A}{\longrightarrow}$ for all $i \in I$. It follows from $s_i \lhd_{FS} \Theta_i$ that $\Theta_i \Longrightarrow \Theta'_i \not\stackrel{A}{\longrightarrow}$. By Theorem 3.18(i) we obtain that $\sum_{i \in I} p_i \cdot \Theta_i \Longrightarrow \sum_{i \in I} p_i \cdot \Theta'_i \not\stackrel{A}{\longrightarrow}$. By the transitivity of $\Longrightarrow$ we have that $\Theta \Longrightarrow \not\stackrel{A}{\longrightarrow}$. □

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

Proposition 6.9 For $\Delta, \Theta \in \mathcal{D}(S)$ 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 $\lhd'_{FS} \subseteq S \times \mathcal{D}(S)$ 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 $\overline{s} \sqsupseteq_{FS} \Theta_i$ 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 6.4, yielding $\overline{\lhd_{FS}}; \Longrightarrow^{-1} \subseteq \sqsupseteq_{FS}$. So suppose, for given $\Delta, \Theta \in \mathcal{D}(S)$, there is a Θ^{match} with $\Theta \Longrightarrow \Theta^{\text{match}}$ and $\Delta \; \overline{\lhd_{FS}} \; \Theta^{\text{match}}$.

Suppose $\Delta \stackrel{\alpha}{\Longrightarrow} \sum_{i \in I} p_i \cdot \Delta'_i$ for some $\alpha \in \mathsf{Act}_\tau$. By Lemma 6.7 there is some Θ' such that $\Theta^{\text{match}} \stackrel{\alpha}{\Longrightarrow} \Theta'$ and $(\sum_{i \in I} p_i \cdot \Delta'_i) \; \overline{\lhd_{FS}} \; \Theta'$. From Proposition 3.9 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 $\Theta \stackrel{\alpha}{\Longrightarrow} \sum_i p_i \cdot \Theta'_i$ by the transitivity of $\Longrightarrow$ (Theorem 3.21) and $\Delta'_i (\overline{\lhd_{FS}}; \Longrightarrow^{-1}) \Theta'_i$ for each $i \in I$ by the reflexivity of $\Longrightarrow$.

Suppose $\Delta \Longrightarrow \not\stackrel{A}{\rightarrow}$. By Lemma 6.8 we have $\Theta^{\text{match}} \Longrightarrow \not\stackrel{A}{\rightarrow}$. It follows that $\Theta \Longrightarrow \not\stackrel{A}{\rightarrow}$ by the transitivity of $\Longrightarrow$. □

Note the appearance of the "anterior step" $\Theta \Longrightarrow \Theta^{\text{match}}$ in Proposition 6.9 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 6.10 Compare the two processes $P := a \stackrel{1}{\oplus} b$ 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 3.9 that would require $[\![a]\!] \; \overline{\lhd_{FS}} \; [\![Q]\!]$, which must fail since the former's move $\stackrel{a}{\longrightarrow} [\![\mathbf{0}]\!]$ 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 6.11 For $s \in S$ and $\Theta \in \mathcal{D}(S)$ 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 6.9; 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 6.9 and the reflexivity of $\Longrightarrow$.

Example 6.10 shows that $\sqsupseteq_{FS}$ cannot be obtained as the lifting of any relation: it lacks the decomposition property of Proposition 3.9. Nevertheless, $\sqsupseteq_{FS}$ enjoys the property of linearity, as occurs in Definition 3.2:

Lemma 6.12 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 3.18(i)), using Proposition 6.9. □

Example 6.13 (Divergence) From Example 3.14 we know that $[\![\mathsf{rec}\, x.\, x]\!] \Longrightarrow \varepsilon$. This, together with (1) in Section 3.1, and the fact that $\varepsilon \stackrel{A}{\not\rightarrow}$ 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 $\Delta = \Delta_0 \stackrel{\tau}{\longrightarrow} \Delta_1 \stackrel{\tau}{\longrightarrow} \cdots \stackrel{\tau}{\longrightarrow} \cdots$ because — as explained right before Example 3.14 — 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 6.5. □

Example 6.13 shows again that the anterior move in Proposition 6.9 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 3.5 any Θ with $\varepsilon \; \overline{\lhd_{FS}} \; \Theta$ must have $|\Theta| = 0$.

Example 6.14 Referring to the process Q_1 of Example 3.15, with Proposition 6.9 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 $[\![a]\!] \stackrel{a}{\longrightarrow} [\![\mathbf{0}]\!]$ with $[\![Q_1]\!] \Longrightarrow [\![a]\!] \stackrel{a}{\longrightarrow} [\![\mathbf{0}]\!]$. It also enables us to match refusals: if $[\![a]\!] \stackrel{B}{\not\rightarrow}$ then B can not contain the action a, and therefore also $[\![Q_1]\!] \Longrightarrow \not\stackrel{B}{\rightarrow}$.

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 6.15 Let P be the process $a \,{}_{\frac{1}{2}}\!\oplus\, \mathsf{rec}\, x.\, x$ and consider the state s_2 introduced in Example 3.16. 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 6.16 (Subdistributions formally unnecessary) For any subdistribution Δ, let Δ^e denote the (full) distribution defined by

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

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 $\Delta \Longrightarrow \Theta$ on p.9 would be much more complex if expressed with full distributions, as would syntactic manipulations such as those used in the proof of Theorem 3.21.

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 3 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 6.13.

We conclude with an example involving divergence and subdistributions.

Example 6.17 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:

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

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 6.4 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 6.18 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 Proposition 6.2 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 $\overline{s} \Longrightarrow\!\!\!\!\!\!\diagup^{A}$ for some $A \subseteq \mathsf{Act}$, then $\Theta_i \Longrightarrow \Theta'_i \;\diagup\!\!\!\!\!\!\rightarrow^{A}$ for all $i \in I$. By definition we have $\sum_{i\in I} p_i\cdot\Theta'_i \;\diagup\!\!\!\!\!\!\rightarrow^{A}$. Theorem 3.18(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 6.19 The relation $\overline{\lhd_{FS}} \subseteq \mathcal{D}(S) \times \mathcal{D}(S)$ is reflexive and transitive.

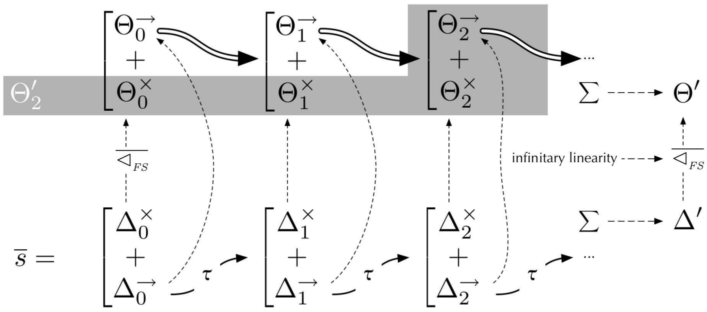


Figure 6: Illustration of Theorem 6.21

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 $s \stackrel{\alpha}{\Longrightarrow} \Delta'$ then there is a Θ' such that $\Theta \stackrel{\alpha}{\Longrightarrow} \Theta'$ and $\Delta' \, \overline{\lhd_{FS}} \, \Theta'$. By Lemma 6.7, 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 3.11 we know that

$$\overline{\lhd_{FS}}; \overline{\lhd_{FS}} = \overline{\lhd_{FS}; \overline{\lhd_{FS}}} \tag{24}$$

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 6.8.

So we established that $\lhd_{FS}; \overline{\lhd_{FS}} \subseteq \lhd_{FS}$. It now follows from Remark 3.3 and (24) that $\overline{\lhd_{FS}}; \overline{\lhd_{FS}} \subseteq \overline{\lhd_{FS}}$. □

6.2 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 6.20 (Simple Failure Similarity) Let $\lhd^s_{FS}$ be the largest relation in $S \times \mathcal{D}(S)$ such that if $s \lhd^s_{FS} \Theta$ 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 $\Delta' \, \overline{\lhd^s_{FS}} \, \Theta'$, and
3. whenever $s \stackrel{A}{\not\longrightarrow}$ then $\Theta \Longrightarrow\!\!\stackrel{A}{\not\longrightarrow}$.

Theorem 6.21 (Equivalence of failure- and simple failure similarity) For finitary distributions $\Delta, \Theta \in \mathcal{D}(S)$ in a pLTS $\langle S, \mathsf{Act}_\tau, \rightarrow\rangle$ we have $\Delta \lhd^s_{FS} \Theta$ iff $\Delta \lhd_{FS} \Theta$.

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 6.20, so that $\lhd_{FS} \subseteq \lhd^s_{FS}$.

For the other direction we need to show that $\lhd^s_{FS}$ satisfies Clause 1 of Definition 6.6 with $\alpha = \tau$, that is

$$\text{if } s \lhd^s_{FS} \Theta \text{ and } \overline{s} \Longrightarrow \Delta' \text{ then there is some } \Theta' \in \mathcal{D}(S) \text{ with } \Theta \Longrightarrow \Theta' \; \Delta' \, \overline{\lhd^s_{FS}} \, \Theta'.$$

Once we have this, the relation $\lhd^s_{FS}$ clearly satisfies both clauses of Definition 6.6, 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 3.12, there must be Δ_k, $\Delta^{\rightarrow}_k$ and $\Delta^{\times}_k$ for $k \geq 0$ such that $\overline{s} = \Delta_0$, $\Delta_k = \Delta^{\rightarrow}_k + \Delta^{\times}_k$, $\Delta^{\rightarrow}_k \stackrel{\tau}{\longrightarrow} \Delta_{k+1}$ 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 3.9 we can define $\Theta =: \Theta^{\times}_0 + \Theta^{\rightarrow}_0$ so that $\Delta^{\{\times,\rightarrow\}}_0 \, \overline{\lhd^s_{FS}} \, \Theta^{\{\times,\rightarrow\}}_0$. Since $\Delta^{\rightarrow}_0 \stackrel{\tau}{\longrightarrow} \Delta_1$ 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 $\Theta_k, \Theta_k^{\rightarrow}, \Theta_k^{\times}$ of subdistributions, for $k \geq 0$, such that $\Theta_0 = \Theta$, $\Delta_k \overline{\lhd^s_{FS}}\ \Theta_k$, $\Theta_k = \Theta_k^{\rightarrow} + \Theta_k^{\times}$, $\Delta_k^{\times} \overline{\lhd^s_{FS}}\ \Theta_k^{\times}$, $\Delta_k^{\rightarrow} \overline{\lhd^s_{FS}}\ \Theta_k^{\rightarrow}$ and $\Theta_k^{\rightarrow} \stackrel{\tau}{\Longrightarrow} \Theta_k$. We define $\Theta' := \sum_i \Theta_i^{\times}$. By Additivity (Remark 3.6) we have $\Delta' \overline{\lhd^s_{FS}}\ \Theta'$. It remains to be shown that $\Theta \Longrightarrow \Theta'$.

For that final step, because $(\Theta \Longrightarrow)$ is closed (Lemma 5.4) 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 $\Theta \Longrightarrow \Theta'_i := (\Theta_i^{\rightarrow} + \sum_{k \leq i} \Theta_k^{\times})$. Since $|\Delta'| = 1$, we must have $\lim_{i\to\infty} |\sum_{k=i}^{\infty} \Delta_i^{\rightarrow}| = 0$ and $\lim_{i\to\infty} |\Delta_i^{\rightarrow}| = 0$, whence by Lemma 3.5, using that $\Delta_i^{\rightarrow} \overline{\lhd^s_{FS}}\ \Theta_i^{\rightarrow}$, also $\lim_{i\to\infty} |\sum_{k=i}^{\infty} \Theta_i^{\rightarrow}| = 0$ and $\lim_{i\to\infty} |\Theta_i^{\rightarrow}| = 0$. 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 Lemma 5.8 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 $\Delta'_{\{1,\varepsilon\}} \overline{\lhd^s_{FS}}\ \Theta'_{\{1,\varepsilon\}}$; 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 3.18(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). □

Add the counterexample from Appendix A.2.3 here.

6.3 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. This is often non-trivial; for example in the development of the testing theory for non-probabilistic processes, this proof obligation caused considerable difficulty and was only achieved by an appeal to König's Lemma (see Lemma 4.4.13 of [7]).

Here, in order to avoid such complications, we introduce yet another version of failure simulation; it modifies Definition 6.20 by checking divergence co-inductively instead of using a predicate.

Definition 6.22 Define $\lhd^c_{FS}$ to be the largest relation in $S \times \mathcal{D}(S)$ such that if $s \lhd^c_{FS} \Theta$ then

1. whenever $s \Longrightarrow \varepsilon$, there are some Δ', Θ' such that $s \Longrightarrow\!\!\stackrel{\tau}{\rightarrow}\!\!\Longrightarrow \Delta' \Longrightarrow \varepsilon$, $\Theta \Longrightarrow\!\!\stackrel{\tau}{\rightarrow}\!\!\Longrightarrow \Theta'$ and $\Delta' \overline{\lhd^c_{FS}}\ \Theta'$; 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 $\Delta' \overline{\lhd^c_{FS}}\ \Theta'$, and
3. whenever $s \stackrel{A}{\not\longrightarrow}$ then $\Theta \Longrightarrow\!\!\stackrel{A}{\not\longrightarrow}$.

Lemma 6.23 The following statements about divergence are equivalent.

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

(2) There is an infinite sequence $\Delta \stackrel{\tau}{\longrightarrow} \Delta_1 \stackrel{\tau}{\longrightarrow} \Delta_2 \stackrel{\tau}{\longrightarrow} \ldots$.

(3) There is an infinite sequence $\Delta \Longrightarrow\!\!\stackrel{\tau}{\rightarrow}\!\!\Longrightarrow \Delta_1 \Longrightarrow\!\!\stackrel{\tau}{\rightarrow}\!\!\Longrightarrow \Delta_2 \Longrightarrow\!\!\stackrel{\tau}{\rightarrow}\!\!\Longrightarrow \ldots$.

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 that has Δ as initial subdistribution and whose states and transitions are subsets of those in L.

(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 6.24 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 $\Delta \Longrightarrow \varepsilon$. By Proposition 3.9(2), 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 $s \Longrightarrow\!\!\xrightarrow{\tau}\!\!\Longrightarrow \Delta'_s \Longrightarrow \varepsilon$, $\Theta_s \Longrightarrow\!\!\xrightarrow{\tau}\!\!\Longrightarrow \Theta'_s$ 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 Proposition 3.9(1), 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 $s \Longrightarrow \varepsilon$. □

Lemma 6.25 $\lhd^c_{FS}$ coincides with $\lhd^s_{FS}$.

Proof: We only need to check that the first clause in Definition 6.20 is equivalent to the first clause in Definition 6.22. For one direction, we consider the relation

$$\mathcal{R} \ := \ \{(s,\Theta) \mid s \Longrightarrow \varepsilon, \Theta \Longrightarrow \varepsilon\}$$

and show $\mathcal{R} \subseteq \lhd^c_{FS}$. Suppose $s \ \mathcal{R} \ \Theta$. By Lemma 6.23 there are two infinite sequences $s \xrightarrow{\tau} \Delta_1 \xrightarrow{\tau} \Delta_2 \xrightarrow{\tau} \ldots$ and $\Theta \xrightarrow{\tau} \Theta_1 \xrightarrow{\tau} \ldots$. Then we have both $\Delta_1 \Longrightarrow \varepsilon$ and $\Theta_1 \Longrightarrow \varepsilon$. Note that $\Delta_1 \Longrightarrow \varepsilon$ if and only if $t \Longrightarrow \varepsilon$ for each $t \in \lceil\Delta_1\rceil$. Therefore, $\Delta_1 \ \overline{\mathcal{R}} \ \Theta_1$ 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 $t \ \mathcal{R} \ \Theta_1$. 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 6.24, we can obtain two infinite sequences $\Delta \Longrightarrow\!\!\xrightarrow{\tau}\!\!\Longrightarrow \Delta_1 \Longrightarrow\!\!\xrightarrow{\tau}\!\!\Longrightarrow \ldots$ and $\Theta \Longrightarrow\!\!\xrightarrow{\tau}\!\!\Longrightarrow \Theta_1 \Longrightarrow\!\!\xrightarrow{\tau}\!\!\Longrightarrow \ldots$ such that $\Delta_i \ \overline{\lhd^c_{FS}} \ \Theta_i$ for all $i \geq 1$. By Lemma 6.23 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 can not rely on clause 2. in Definition 6.22, as the move generated there might actually be empty, as we have seen in Example 3.13. Instead we need a method for generating p-weak moves which contain at least one occurrence of a τ-action.

Definition 6.26 [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 $\Delta \xrightarrow{\tau}_p \Theta$.

First let us recall the following lemma which appeared as Lemma 6.12 in [3]; it still holds in our current setting. The following lemma appeared as Lemma 6.12 in [3]. It still holds in our current setting.

Lemma 6.27 (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 $\mathcal{R}, \mathcal{R}' \subseteq S \times \mathcal{D}(S)$ satisfying $u\mathcal{R}\Psi$ whenever $u = s \mid_A t$ and $\Psi = \Theta \mid_A t$ with $s \ \mathcal{R}' \ \Theta$ and $t \in S$. Then $\Delta \ \overline{\mathcal{R}'} \ \Theta$ and $\Phi \in \mathcal{D}(S)$ implies $(\Delta \mid_A \Phi) \ \overline{\mathcal{R}} \ (\Theta \mid_A \Phi)$. □

Proposition 6.28 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 $\Gamma \ \overline{\mathcal{R}} \ \Psi$, where $\mathcal{R}$ 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 $\Gamma \ \overline{\mathcal{R}} \ \Psi$. 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 6.27(2) and $(s \mid_A \Phi)\ \overline{\mathcal{R}}\ (\Theta \mid_A \Phi)$ by Lemma 6.27(5), whereas in the second case $s \lhd^c_{FS} \Theta$ implies $\Theta \Longrightarrow\xrightarrow{a}\Longrightarrow \Theta'$ for some $\Theta' \in \mathcal{D}(S)$ with $\Delta\ \overline{\lhd^c_{FS}}\ \Theta'$, and we have $\Theta \mid_A t \Longrightarrow\xrightarrow{\tau}\Longrightarrow \Theta' \mid_A \Phi$ by Lemma 6.27(1) and (3), and $(\Delta \mid_A \Phi)\ \overline{\mathcal{R}}\ (\Theta' \mid_A \Phi)$ by Lemma 6.27(5).

The general case now follows using a standard decomposition/recomposition argument. Since $\Delta \mid_A t \xrightarrow{\tau}_p \Gamma$, Lemma 3.4 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 $s_i \in S$, $\Gamma_i \in \mathcal{D}(S)$ and $\sum_{i\in I} p_i \leq 1$. For finitary processes $\lhd^c_{FS}$ is convex by combination of Proposition 6.18, Theorem 6.21 and Lemma 6.25. Hence, since $\Delta\ \overline{\lhd^c_{FS}}\ \Theta$, Corollary 3.10 yields that $\Theta = \sum_{i\in I} p_i{\cdot}\Theta_i$ for some $\Theta_i \in \mathcal{D}(S)$ 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 $\Psi_i \in \mathcal{D}(S)$ such that $\Gamma_i\ \overline{\mathcal{R}}\ \Psi_i$. The required Ψ can be taken to be $\sum_{i\in I} p_i{\cdot}\Psi_i$ as Definition 3.2(2) yields $\Gamma\ \overline{\mathcal{R}}\ \Psi$ and Theorem 3.18(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 6.29 Suppose $\Delta \mid_A t \xrightarrow{\tau} \Gamma$. Then there exists subdistributions $\Delta^{\rightarrow},\ \Delta^{\times},\ \Delta^{\rm next},\ \Gamma^{\times}$ (possibly empty) such that

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

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

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

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

Proof: By Lemma 3.4 $\Delta \mid_A t \xrightarrow{\tau} \Gamma$ 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 $s_i \in S$, $\Gamma_i \in \mathcal{D}(S)$ 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

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

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

Lemma 6.30 If $\Delta \mid_A t \Longrightarrow \varepsilon$ then there is a $\Delta' \in \mathcal{D}(S)$ 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 6.23 there is an infinite sequence

$$\Delta_0 \mid_A t \xrightarrow{\tau} \Psi_1 \xrightarrow{\tau} \Psi_2 \xrightarrow{\tau} \ldots \tag{25}$$

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 6.30.

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 3.9 gives us a $\Gamma_{k+1} \subseteq \Psi_{k+1}$ such that $\Delta_k \mid_A t \xrightarrow{\tau} \Gamma_{k+1}$ and $\Gamma_{k+1} \leq \Psi_{k+1}$. Now apply Lemma 6.30.

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 3.18(ii) that $\Gamma_k^\times \Longrightarrow \varepsilon$. Hence Theorem 3.18(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 6.31 *If* $\Delta \sqsupseteq_{FS} \Theta$ *then* $\Delta \mid_A \Phi \sqsupseteq_{FS} \Theta \mid_A \Phi$.

Proof: We first construct the following relation

$$\mathcal{R} \;:=\; \{(s \mid_A t, \Theta \mid_A t) \mid s \lhd^c_{FS} \Theta\}$$

and check that $\mathcal{R} \subseteq \lhd^c_{FS}$. 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 $s \lhd^c_{FS} \Theta$ and $s \mid_A t \Longrightarrow \varepsilon$. We need to find some Γ, Ψ such that

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

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

By Lemma 6.30 there are $\Delta', \Gamma \in \mathcal{D}(S)$ such that $s \Longrightarrow \Delta'$ and $\Delta' \mid_A t \xrightarrow{\tau}_p \Gamma \Longrightarrow \varepsilon$. Since $\lhd^c_{FS}$ coincides with $\lhd^s_{FS}$ and $\lhd_{FS}$, there must be a $\Theta' \in \mathcal{D}(S)$ such that $\Theta \Longrightarrow \Theta'$ and $\Delta' \;\overline{\lhd^c_{FS}}\; \Theta'$. By Proposition 6.28 we have $\Theta' \mid_A t \Longrightarrow \xrightarrow{\tau} \Longrightarrow \Psi$ for some Ψ such that $\Gamma \;\overline{\mathcal{R}}\; \Psi$. 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 $\Gamma \;\overline{\mathcal{R}}\; \Psi$, which had to be shown.

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

Suppose $\Delta \sqsupseteq_{FS} \Theta$. By definition this means that there is some Θ^{match} such that $\Theta \Longrightarrow \Theta^{\text{match}}$ and $\Delta \;\overline{\lhd^s_{FS}}\; \Theta^{\text{match}}$. By Lemma 6.25 we have $\Delta \;\overline{\lhd^c_{FS}}\; \Theta^{\text{match}}$ and Lemma 6.27(5) yields $(\Delta \mid_A \Phi) \;\overline{\mathcal{R}}\; (\Theta^{\text{match}} \mid_A \Phi)$. Therefore, we have $(\Delta \mid_A \Phi) \;\overline{\lhd^c_{FS}}\; (\Theta^{\text{match}} \mid_A \Phi)$, i.e. $(\Delta \mid_A \Phi) \;\overline{\lhd^s_{FS}}\; (\Theta^{\text{match}} \mid_A \Phi)$ by Lemma 6.25. By Lemma 6.27(1) we also have $(\Theta \mid_A \Phi) \Longrightarrow (\Theta^{\text{match}} \mid_A \Phi)$, which had to be established. □

Proposition 6.32 (Precongruence) If $P \sqsupseteq_{FS} Q$ then $\alpha.P \sqsupseteq_{FS} \alpha.Q$ for $\alpha \in \mathsf{Act}$, and similarly if $P_{1,2} \sqsupseteq_{FS} Q_{1,2}$ respectively then $P_1 \odot P_2 \sqsupseteq_{FS} Q_1 \odot Q_2$ for $\odot$ 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 Lemma 6.31. The other cases are simpler, thus omitted. □

Lemma 6.33 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

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

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 $\mathcal{R} \subseteq \lhd^c_{FS}$.

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 Lemma 6.31, besides that we need to check the requirements $t \xrightarrow{\omega}\!\!\!\!\!/$ and $\Gamma \xrightarrow{\omega}\!\!\!\!\!/$ 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 $(s \mid_{\mathsf{Act}} \Gamma, \Theta \mid_{\mathsf{Act}} \Gamma) \in \overline{\mathcal{R}}$.
   - 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. It follows that $\overline{t} \stackrel{\omega}{\not\longrightarrow}$. 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 $s \lhd^c_{FS} \Theta$, there exists some Θ' such that $\Theta \Longrightarrow \Theta'$ and $\Delta \overline{\lhd^c_{FS}} \Theta'$. 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 $(\Delta \mid_{\mathsf{Act}} t, \Theta' \mid_{\mathsf{Act}} t) \in \overline{\mathcal{R}}$.
     - $s \xrightarrow{a} \Delta$ and $t \xrightarrow{a} \Gamma$ for some action $a \in \mathsf{Act}$. Since $s \lhd^c_{FS} \Theta$, there exists some Θ' such that $\Theta \stackrel{a}{\Longrightarrow} \Theta'$ and $\Delta \overline{\lhd^c_{FS}} \Theta'$. 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 $(\Delta \mid_{\mathsf{Act}} \Gamma, \Theta' \mid_{\mathsf{Act}} \Gamma) \in \overline{\mathcal{R}}$.
   - 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 $s \lhd^c_{FS} \Theta$ 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 $\mathcal{R} \subseteq \lhd^c_{FS}$, from which our expected result can be establishing using similar arguments in the last part of the proof of Lemma 6.31. □

6.4 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 computation structures before extracting values from them, we will be concerned mainly with ω-respecting structures. For those we have the following

Lemma 6.34 Let Δ and Θ be subdistributions in an ω-respecting computation structure. If subdistribution Δ is stable and $\Delta \overline{\lhd^s_{FS}} \Theta$, then $\mathbb{V}(\Delta) \in \mathcal{V}(\Theta)$.

Proof: We first show that if s is stable and $s \lhd^s_{FS} \Theta$ then $\mathbb{V}(s) \in \mathcal{V}(\Theta)$. Since s is stable, we have only two cases:

(i) $s \not\longrightarrow$ Here $\mathbb{V}(s)=0$ and since $s \lhd^s_{FS} \Theta$ we have $\Theta \Longrightarrow \Theta'$ with $\Theta' \not\longrightarrow$, whence in fact $\Theta \Longrightarrow\!\!\!\!\!\rightarrow \Theta'$ and $\mathbb{V}(\Theta') = 0$. Thus from Lemma 4.34 we have $\mathbb{V}(s) = 0 \in \mathcal{V}(\Theta)$.

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

Now for the general case we suppose $\Delta \overline{\lhd^s_{FS}} \Theta$. Use Proposition 3.9 to decompose Θ into $\sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot \Theta_s$ such that $s \lhd^s_{FS} \Theta_s$ for each $s \in \lceil \Delta \rceil$, and recall each such state s is stable. From above we have that $\mathbb{V}(s) \in \mathcal{V}(\Theta_s)$ for those s, and so $\mathbb{V}(\Delta) = \sum_{\in \lceil \Delta \rceil} \Delta(s) \cdot \mathbb{V}(s) \in \sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot \mathcal{V}(\Theta_s) = \mathcal{V}(\Theta)$. □

Lemma 6.35 Let Δ be a subdistribution in an ω-respecting computation structure. If $\Delta \Longrightarrow \Delta'$ then $\mathcal{V}(\Delta') \subseteq \mathcal{V}(\Delta)$.

Proof: Note that if $\Delta' \Longrightarrow\!\!\!\!\!\rightarrow \Delta''$ then $\Delta \Longrightarrow \Delta' \Longrightarrow\!\!\!\!\!\rightarrow \Delta''$, so that every extreme derivative of Δ' is also an extreme derivative of Δ. The result follows from Lemma 4.34. □

Lemma 6.36 Let Δ and Θ be subdistributions in an ω-respecting computation structure. If $\Delta \overline{\lhd^s_{FS}} \Theta$, then we have $\mathcal{V}(\Delta) \subseteq \mathcal{V}(\Theta)$.

Proof: Since $\Delta \overline{\lhd^s_{FS}} \Theta$, for any $\Delta \Longrightarrow \Delta'$ we have the matching transition $\Theta \Longrightarrow \Theta'$ such that $\Delta' \overline{\lhd^s_{FS}} \Theta'$. It follows from Lemmas 6.34 and 6.35 that $\mathbb{V}(\Delta') \in \mathcal{V}(\Theta') \subseteq \mathcal{V}(\Theta)$. By Lemma 4.34 we obtain $\mathcal{V}(\Delta) \subseteq \mathcal{V}(\Theta)$. □

Lemma 6.37 Let Δ and Θ be subdistributions in an ω-respecting computation structure. If $\Theta \sqsubseteq_{FS} \Delta$, then it holds that $\mathcal{V}(\Theta) \supseteq \mathcal{V}(\Delta)$.

Proof: Suppose $\Theta \sqsubseteq_{FS} \Delta$. By Proposition 6.9 and Theorem 6.21, there exists some Θ' such that $\Theta \Longrightarrow \Theta'$ and $\Delta \overline{\lhd^s_{FS}} \Theta'$. By Lemmas 6.36 and 6.35 we obtain $\mathcal{V}(\Delta) \subseteq \mathcal{V}(\Theta') \subseteq \mathcal{V}(\Theta)$. □

Theorem 6.38 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 6.33, for any test T
implies	$\mathcal{V}([P \mid_{\mathsf{Act}} T]) \supseteq \mathcal{V}([Q \mid_{\mathsf{Act}} T])$	$[\cdot]$ is ω-respecting; Lemma 6.37
iff	$\mathcal{A}(T, P) \supseteq \mathcal{A}(T, Q)$	(4)
implies	$\mathcal{A}(T, P) \leq_{\mathrm{Sm}} \mathcal{A}(T, Q)$	Def. Smyth order
iff	$P \sqsubseteq_{\text{pmust}} Q$.	Definition 4.5

□

Corollary 6.39 If $P \sqsubseteq_{FS} Q$ then $P \sqsubseteq^{\Omega}_{\text{pmust}} Q$.

Proof: Section 4.3.1 recalled our earlier result that Ω-testing is reducible to scalar testing. □

7 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 finite approximations. 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 finite-state. From this, completeness follows.

7.1 Inductive characterisation

$\lhd^s_{FS} \lhd_{FS} \lhd^k_{FS} \lhd'_{FS} \lhd^c_{FS}$ The relation $\lhd^s_{FS}$ of Definition 6.20 is given coinductively: it is the largest fixpoint of an equation $\mathcal{R} = \mathcal{F}(\mathcal{R})$. An alternative approach is to use that $\mathcal{F}(-)$ to define $\lhd^s_{FS}$ as a limit of approximants:

Definition 7.1 For every $k \geq 0$ we define the relations $\lhd^k_{FS} \subseteq S \times \mathcal{D}(S)$ as follows:

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

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

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 $\mathrm{rec}\, x.\, a.x$ 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 k a actions in succession and then deadlock. This can be described by the infinitary CCS expression $\sum_{k=0}^{\infty} a^k$. 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 finitely branching state spaces, however, a simple counting argument will show that $\lhd^s_{FS}$ coincides with $\lhd^\infty_{FS}$; see [8, 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 for example the process $\tau.a + \tau.b$ can move via $\Longrightarrow$ to a distribution $[\![a]\!]_p\oplus [\![b]\!]$ for any of the non-denumerably many probabilities $p\in[0,1]$ — and that is hardly finite. Even worse, in probabilistic systems one can have infinite branching over a finite-state system (not possible without probability), just e.g. by picking some infinite subset of $[0,1]$ to be the allowed p's in the example above; that is why we must now impose finite branching explicitly. The effect is what one could call "finitely generated" transitions (in this case by arbitrary interpolation of the two extreme possibilities a and b, two being a finite number), 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 $(\mathcal{D}(S), d_1)$ with $d_1(\Delta,\Theta) = max_{s\in S}|\Delta(s) - \Theta(s)|$ and $(S \to \mathcal{D}(S), d_2)$ with $d_2(f,g) = max_{s\in S} d_1(f(s), g(s))$ are complete. Let X be a subset of either $\mathcal{D}(S)$ or $S \to \mathcal{D}(S)$. Clearly, X is bounded. So if X is closed, it is also compact.

Definition 7.2 A relation $\mathcal{R} \subseteq S \times \mathcal{D}(S)$ is *closed* if for every $s \in S$ the set $s\cdot\ \mathcal{R}$ is closed.

Two examples of closed relations are $\Longrightarrow$ and $\stackrel{a}{\Longrightarrow}$ for any a, as shown by Lemmas 5.4 and 5.5.

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 7.3 Let $\mathcal{R}\subseteq S\times\mathcal{D}(S)$ be closed. Then its set of choice functions $\{\, f : S \to \mathcal{D}(S) \mid f \in_S \mathcal{R} \,\}$ is also closed.

Proof: Straightforward. Need to say what definition of closure we're using wrt a set of functions. □

Corollary 7.4 Let $\mathcal{R}\subseteq S \times \mathcal{D}(S)$ be closed and convex. Then $\overline{\mathcal{R}}$ is also closed.

Proof: For any $\Delta \in \mathcal{D}(S)$, we know from Proposition 3.8 that $\Delta\cdot\ \overline{\mathcal{R}} = \{\mathrm{Exp}_\Delta(f) \mid f \in_{\lceil\Delta\rceil}\mathcal{R}\}$. The function $\mathrm{Exp}_\Delta(-)$ is continuous. By Lemma 7.3 the set of choice functions of $\mathcal{R}$ is closed, and it is also bounded, thus being compact. Its image is also compact, thus being closed. □

Lemma 7.5 Let $\mathcal{R} \subseteq\ S \times \mathcal{D}(S)$ be closed and convex, and $C\ \subseteq\ \mathcal{D}(S)$ be closed. Then the set $\{\, \Delta \mid \Delta\cdot\ \overline{\mathcal{R}} \cap C \neq \emptyset \,\}$ is also closed.

Proof: First define $\mathcal{E} : \mathcal{D}(S) \times (S \to \mathcal{D}(S)) \ \to\ \mathcal{D}(S)$ by $\mathcal{E}(\Theta, f) = \mathrm{Exp}_\Theta(f)$, which is obviously continuous. Then let $F = \{\, f \mid f \in_{\lceil\Delta\rceil}\mathcal{R} \,\}$, which by the previous lemma is closed. Finally let

$$Z = \pi_1(\mathcal{E}^{-1}(C) \cap (\mathcal{D}(S) \times F))$$

where π_1 is the projection on to the first component of a pair. We observe that the continuity of $\mathcal{E}$ ensures that the inverse image of the closed set C is closed. Furthermore, $\mathcal{E}^{-1}(C) \cap (\mathcal{D}(S) \times F)$ 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 $Z = \{\, \Delta \mid \Delta\cdot\ \overline{\mathcal{R}} \cap C \neq \emptyset \,\}$ because

$$\begin{aligned}
\Delta \in Z \ \ &\text{iff } (\Delta, f) \in \mathcal{E}^{-1}(C) \text{ for some } f \in_{\lceil\Delta\rceil}\mathcal{R}\\
&\text{iff } \mathcal{E}(\Delta, f) \in C \text{ for some } f \in_{\lceil\Delta\rceil}\mathcal{R}\\
&\text{iff } \mathrm{Exp}_\Delta(f) \in C \text{ for some } f \in_{\lceil\Delta\rceil}\mathcal{R}\\
&\text{iff } \Delta\ \overline{R}\ \Delta' \text{ for some } \Delta' \in C
\end{aligned}$$

The last line is an application of Proposition 3.8, which requires convexity of $\mathcal{R}$. □

An immediate corollary of this last result is:

Corollary 7.6 In a finitary pLTS the following sets are closed:

(i) $\{\, \Delta \mid \Delta \Longrightarrow \varepsilon \,\}$

(ii) $\{\, \Delta \mid \Delta \stackrel{A}{\not\Longrightarrow} \,\}$

Proof: By Lemma 5.4 and Corollary 3.20 we know 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 7.7 *For every $k \geq 0$ the relation $\lhd_{FS}^k$ is closed and convex.*

Proof: By induction on k. For $k = 0$ it is obvious. So let us assume that $\lhd_{FS}^k$ is closed and convex. We have to show that

$$s\cdot \lhd_{FS}^{(k+1)} \text{ is closed and convex, for every state } s \tag{26}$$

If $s \Longrightarrow \varepsilon$ then this follows from the corollary above, since in this case $s\cdot \lhd_{FS}^{(k+1)}$ coincides with $\{\, \Delta \mid \Delta \Longrightarrow \varepsilon \,\}$. So let us assume that this is not the case.

For every $A \subseteq \mathsf{Act}$ let $R_A = \{\, \Delta \mid \Delta \stackrel{A}{\not\Longrightarrow} \,\}$, which we know by the corollary above to be closed and is obviously convex. Also for every $\Theta,\ \alpha$ let $G_{\Theta,\alpha} = \{\, \Delta \mid (\Delta\cdot \stackrel{\alpha}{\Longrightarrow}) \cap (\Theta\cdot \overline{\lhd_{FS}^k}) \neq \emptyset \,\}$. By Proposition 6.2, the relation $\stackrel{\alpha}{\Longrightarrow}$ is lifted from a closed convex relation. By Corollary 7.4, the assumption that $\lhd_{FS}^k$ is closed and convex implies that $\overline{\lhd_{FS}^k}$ is also closed. So we can appeal to Lemma 7.5 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 $s\cdot \lhd_{FS}^{(k+1)}$ is also closed and convex as it can be written as

$$\cap\{\, R_A \mid s \stackrel{A}{\not\longrightarrow} \,\} \ \cap\ \cap\{\, G_{\Theta,\alpha} \mid s \stackrel{\alpha}{\longrightarrow} \Theta \,\}$$

□

Before the main result of this section we need one more technical result.

Lemma 7.8 Let S be a finite set of states. Suppose $\mathcal{R}^k \subseteq S \times \mathcal{D}(S)$ is a sequence of closed convex relations such that $\mathcal{R}^{(k+1)} \subseteq \mathcal{R}^k$. Then $(\cap_{k=0}^{\infty}\, \overline{\mathcal{R}^k}) \subseteq \overline{(\cap_{k=0}^{\infty}\, \mathcal{R}^k)}$.

Proof: Let $\mathcal{R}^\infty$ denote $(\cap_{k=0}^{\infty}\, \mathcal{R}^k)$, and suppose $\Delta\ \overline{\mathcal{R}^k}\ \Theta$ for every $k \geq 0$. We have to show that $\Delta \overline{\mathcal{R}^\infty} \Theta$.

Let $G = \{\, f : S \to \mathcal{D}(S) \mid \Theta = \mathrm{Exp}_\Delta(f) \,\}$, which is easily seen to be a closed set. For each k let $F^k = \{\, f : S \to \mathcal{D}(S) \mid f \in_{\lceil \Delta \rceil} \mathcal{R}^k \,\}$, which by Lemma 7.3 we also know to be closed. Finally consider the collection of closed sets $H^k = F^k \cap G$; since $\Delta\ \overline{\mathcal{R}^k}\ \Theta$, Proposition 3.8 assures us that all of these are non-empty. Also $H^{(k+1)} \subseteq H^k$ and therefore by the finite-intersection property [14] $\cap_{k=0}^{\infty} H^k$ is also non-empty.

Let f be an arbitrary element of this intersection. For any state s, and for every $k \geq 0,\ s\ \mathcal{R}^k\ f(s)$, that is $s\ \mathcal{R}^\infty\ f(s)$. So f is a choice function for $\mathcal{R}^\infty$, $f \in_{\lceil \Delta \rceil} \mathcal{R}^\infty$. From convexity and Proposition 3.8 it follows that $\Delta\ \overline{\mathcal{R}^\infty}\ \mathrm{Exp}_\Delta(f)$. But from the definition of the G we know that $\Theta = \mathrm{Exp}_\Delta(f)$, and the required result follows. □

Theorem 7.9 In a finitary pLTS, $s \lhd_{FS}^s \Theta$ if and only if $s \lhd_{FS}^\infty \Theta$.

Proof: Since $\lhd_{FS}^s \subseteq \lhd_{FS}^\infty$ it is sufficient to show the opposite inclusion, which by definition holds if $\lhd_{FS}^\infty$ is a failure simulation, viz. if $\lhd_{FS}^\infty \subseteq \mathcal{F}(\lhd_{FS}^\infty)$. Suppose $s \lhd_{FS}^\infty \Theta$, which means that $s \lhd_{FS}^k \Theta$ for every $k \geq 0$. According to Definition 6.20, in order to show $s\ \mathcal{F}(\lhd_{FS}^\infty)\ \Theta$ we have to establish three properties, the first and last of which are trivial (for they are independent on the argument of $\mathcal{F}$).

So suppose $s \stackrel{\alpha}{\longrightarrow} \Delta'$. We have to show that $\Theta \stackrel{\alpha}{\Longrightarrow} \Theta'$ for some Θ' such that $\Delta'\ \overline{\lhd_{FS}^\infty}\ \Theta'$.

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 5.4 and Proposition 7.7 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 7.7, the relations $\lhd^k_{FS}$ are all closed and convex. Therefore, Lemma 7.8 may be applied to them, which enables us to conclude $\Delta' \overline{\lhd^{\infty}_{FS}} \Theta'$. □

Analogously to what we did for $\lhd^s_{FS}$, we also give an inductive characterisation of $\sqsupseteq_{FS}$: For every $k \geq 0$ let $\Delta \sqsupseteq^k_{FS} \Theta$ if there exists a $\Theta \Longrightarrow \Theta^{\text{match}}$ such that $\Delta \overline{\lhd^k_{FS}} \Theta^{\text{match}}$, and let $\sqsupseteq^{\infty}_{FS}$ denote $\bigcap_{k=0}^{\infty} \sqsupseteq^k_{FS}$.

Corollary 7.10 In a finitary pLTS, $\Delta \sqsupseteq_{FS} \Theta$ if and only if $\Delta \sqsupseteq^{\infty}_{FS} \Theta$.

Proof: Since $\lhd^s_{FS} \subseteq \lhd^k_{FS}$ for every $k \geq 0$, it is straightforward to prove one direction: $\Delta \sqsupseteq_{FS} \Theta$ implies $\Delta \sqsupseteq^{\infty}_{FS} \Theta$. For the converse, $\Delta \sqsupseteq^{\infty}_{FS} \Theta$ means that for every k we have some Θ^k satisfying $\Theta \Longrightarrow \Theta^k$ and $\Delta \overline{\lhd^k_{FS}} \Theta^k$. By Proposition 6.9 we have to find some Θ^{∞} such that $\Theta \Longrightarrow \Theta^{\infty}$ and $\Delta \overline{\lhd^k_{FS}} \Theta^{\infty}$. This can be done exactly as in the proof of Theorem 7.9. □

7.2 The modal logic

Let $\mathcal{F}$ be the set of modal formulae defined inductively as follows:

- $\mathbf{div}, \top \in \mathcal{F}$
- $\mathbf{ref}(A) \in \mathcal{F}$ when $A \subseteq \mathsf{Act}$,
- $\langle a \rangle \varphi \in \mathcal{F}$ when $\varphi \in \mathcal{F}$ and $a \in \mathsf{Act}$,
- $\varphi_1 \wedge \varphi_2 \in \mathcal{F}$ when $\varphi_1,\ \varphi_2 \in \mathcal{F}$,
- $\varphi_1 \,{}_p\!\oplus\, \varphi_2 \in \mathcal{F}$ when $\varphi_1, \varphi_2 \in \mathcal{F}$ 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* $\models \subseteq \mathcal{D}(S) \times \mathcal{F}$ is given by:

- $\Delta \models \top$ for any $\Delta \in \mathcal{D}(S)$,
- $\Delta \models \mathbf{div}$ 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 $\Delta_1, \Delta_2 \in \mathcal{D}(S)$ 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 $\Delta \sqsupseteq^{\mathcal{F}} \Theta$ when $\Delta \models \varphi$ implies $\Theta \models \varphi$ for all $\varphi \in \mathcal{F}$ — note the opposing directions. This is because the modal formulae express "bad" properties of our processes, ultimately divergence and refusal: thus $\Theta \sqsubseteq^{\mathcal{F}} \Delta$ means that any bad thing implementation Δ does must have been allowed by the specification Θ.

For pCSP processes we use $P \sqsubseteq^{\mathcal{F}} Q$ to abbreviate $[\![P]\!] \sqsubseteq^{\mathcal{F}} [\![Q]\!]$ 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. Recall that 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 4.3 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 7.11 If $\Delta \sqsupseteq_{FS} \Theta$ then $\Delta \sqsupseteq^{\mathcal{F}} \Theta$.

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 6.5.
- 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 6.4 Case 1 and the induction hypothesis.
- The case when φ is $\mathbf{ref}(A)$ follows by Definition 6.4 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 6.4 Case 1, using $\alpha := \tau$ to infer the existence of suitable $\Theta'_{\{1,2\}}$. □

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 $\sqsubseteq^{\mathcal{F}}$.

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 7.12 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 $\Delta \in \mathcal{D}(S)$ 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 $\varphi^{k+1}_\Delta = (\mathbf{div}) \;{}_{1-\lceil\Delta\rceil}\!\oplus\; \left(\bigoplus_{s \in \lceil\Delta\rceil} \frac{\Delta(s)}{\lceil\Delta\rceil} \cdot \varphi^{k+1}_s\right)$.

Lemma 7.13 For every $k \geq 0$, $s \in S$ and $\Delta \in \mathcal{D}(S)$ 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 7.14 For $k \geq 0$,

(i) $\Theta \models \varphi^k_s$ implies $s \lhd^k_{FS} \Theta$,

(ii) $\Theta \models \varphi^k_\Delta$ implies $\Theta \Longrightarrow \Theta^{\text{match}}$ such that $\Delta \;\overline{\lhd^k_{FS}}\; \Theta^{\text{match}}$,

(iii) $\Theta \models \varphi^k_\Delta$ implies $\Theta \sqsupseteq^k_{FS} \Delta$.

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 $\varphi^{k+1}_s = \mathbf{div}$ and therefore $\Theta \models \mathbf{div}$, which gives the required $\Theta \Longrightarrow \varepsilon$.

Now suppose $s \stackrel{\tau}{\longrightarrow} \Delta$. 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 $\Delta \;\overline{\lhd^k_{FS}}\; \varepsilon$. So let us assume that $\overline{s} \not\Longrightarrow \varepsilon$. Then by the definition of φ^{k+1}_s we must have that $\Theta \models \varphi^k_\Delta$, 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 $\Delta \;\overline{\lhd^k_{FS}}\; \Theta'$.

The matching move for $s \xrightarrow{a} \Theta$ is obtained in a similar manner.

Finally suppose $s \xrightarrow{A}\!\!\!\!\!/$. Since this implies $s \xrightarrow{\tau}\!\!\!\!\!/$, by the definition of φ_s^{k+1} we must have that $\Theta \models \mathbf{ref}(A)$, which actually means that $\Theta \Longrightarrow \xrightarrow{A}\!\!\!\!\!/$.

(ii) By definition $\varphi_\Delta^{k+1} = (\mathbf{div})\ {}_{1-\lceil\Delta\rceil}\!\oplus\ (\bigoplus_{s\in\lceil\Delta\rceil} \frac{\Delta(s)}{\lceil\Delta\rceil}\cdot\varphi_s^{k+1})$ and thus $\Theta \Longrightarrow (1-\lceil\Delta\rceil)\cdot\Theta_{\mathbf{div}} + \sum_{s\in\lceil\Delta\rceil}\Delta(s)\cdot\Theta_s$ 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 3.18(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 $\Delta \lhd_{FS}^{k+1} \sum_{s\in\lceil\Delta\rceil}\Delta(s)\cdot\Theta_s$. □

Theorem 7.15 In a finitary pLTS, $\Delta \sqsupseteq^{\mathcal{F}} \Theta$ if and only if $\Delta \sqsupseteq_{FS} \Theta$.

Proof: One direction follows immediately from Theorem 7.11. For the opposite direction suppose $\Delta \sqsupseteq^{\mathcal{F}} \Theta$. By Lemma 7.13 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 7.10. □

7.3 Characteristic tests for formulae

The import of Theorem 7.15 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 $\mathcal{A}(-,-)$, which are summarised in the following two lemmas 7.16 and 7.19, corresponding to Lemmas 6.7 and 6.8 in [2] respectively.

Lemma 7.16 Let Δ be a pCSP process, and T, T_i be tests.

1. $o \in \mathcal{A}(\omega, \Delta)$ iff $o = |\Delta|\cdot\vec{\omega}$.
2. $\vec{0} \in \mathcal{A}(\tau.\omega, \Delta)$ iff $\Delta \Longrightarrow \varepsilon$.
3. $\vec{0} \in \mathcal{A}(\sum_{a\in A} a.\omega, \Delta)$ iff $\Delta \Longrightarrow \xrightarrow{A}\!\!\!\!\!/$.
4. Suppose the action ω does not occur in the test T. Then $o \in \mathcal{A}(\tau.\omega \,\Box\, a.T, \Delta)$ with $o(\omega) = 0$ iff there is a $\Delta' \in \mathcal{D}(\mathsf{sCSP})$ with $\Delta \stackrel{a}{\Longrightarrow} \Delta'$ and $o \in \mathcal{A}(T, \Delta')$.
5. $o \in \mathcal{A}(T_1\ {}_p\!\oplus T_2, \Delta)$ iff $o = p\cdot o_1 + (1-p)\cdot o_2$ for certain $o_i \in \mathcal{A}(T_i, \Delta)$.
6. $o \in \mathcal{A}(T_1 \sqcap T_2, \Delta)$ if there are a $q \in [0,1]$ and $\Delta_1, \Delta_2 \in \mathcal{D}(\mathsf{sCSP})$ 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 $o_i \in \mathcal{A}(T_i, \Delta_i)$.

Proof:

1. Since $\omega \mid_{\mathsf{Act}} P \xrightarrow{\omega}$, 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 $\mathcal{A}(\omega, \Delta) = \{|\Delta|\cdot\vec{\omega}\}$.
2. ($\Leftarrow$) Assume $\Delta \Longrightarrow \varepsilon$. By Lemma 6.27(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 3.12) have the form $\tau.\omega \mid_{\mathsf{Act}} s$, and thus satisfy $u \xrightarrow{\omega}\!\!\!\!\!/$ 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 $\vec{0} \in \mathcal{A}(\tau.\omega, \Delta)$.

   ($\Rightarrow$) Suppose $\vec{0} \in \mathcal{A}(\tau.\omega, \Delta)$, 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. If follows that all transitions contributing to the derivation $[\tau.\omega \mid_{\mathsf{Act}} \Delta] \Longrightarrow\!\!\!\!\!\rightarrow \Gamma$ 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 6.27(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 $[T \mid_{\mathsf{Act}} \Delta] \Longrightarrow\!\!\!\!\!\!\rightarrow [T \mid_{\mathsf{Act}} \Delta]$ and $\$(T \mid_{\mathsf{Act}} \Delta) = \vec{0}$. Thus we have $\vec{0} \in \mathcal{A}(T, \Delta)$.

   ($\Rightarrow$) Suppose $\vec{0} \in \mathcal{A}(T, \Delta)$. 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 $\Delta' \in \mathcal{D}(\mathsf{sCSP})$ with $\Delta \stackrel{a}{\Longrightarrow} \Delta'$ and $o \in \mathcal{A}(T, \Delta')$. W.l.o.g. we may assume Why? that $\Delta \Longrightarrow \Delta^{\mathrm{pre}} \stackrel{a}{\longrightarrow} \Delta'$. Using Lemma 6.27(1) and (3), and the same reasoning as in the previous cases, $[U \mid_{\mathsf{Act}} \Delta] \Longrightarrow [U \mid_{\mathsf{Act}} \Delta^{\mathrm{pre}}] \stackrel{\tau}{\longrightarrow} [T \mid_{\mathsf{Act}} \Delta'] \Longrightarrow \Gamma$ for a stable subdistribution Γ with $\$\Gamma = o$. It follows that $o \in \mathcal{A}(U, \Delta)$.

   ($\Rightarrow$) Suppose $o \in \mathcal{A}(U, \Delta)$ 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 no state in the support of Γ of the form $\omega \mid_{\mathsf{Act}} t$. Hence there must be a $\Delta' \in \mathcal{D}(\mathsf{sCSP})$ such that $\Delta \Longrightarrow \stackrel{a}{\longrightarrow} \Delta'$ and $[T \mid_{\mathsf{Act}} \Delta'] \Longrightarrow \Gamma$. It follows that $o \in \mathcal{A}(T, \Delta')$.

5. ($\Leftarrow$) Assume $o_i \in \mathcal{A}(T_i, \Delta)$ 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 3.18(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 $o \in \mathcal{A}(T_1 \,{}_p\!\oplus\, T_2, \Delta)$.

   ($\Rightarrow$) Suppose $o \in \mathcal{A}(T_1 \,{}_p\!\oplus\, T_2, \Delta)$. 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 3.18(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 $\$\Gamma_i \in \mathcal{A}(T_i, \Delta)$ for $i = 1, 2$. Moreover, $o = \$\Gamma = p\cdot\$\Gamma_1 + (1{-}p)\cdot\$\Gamma_2$.

6. Suppose $q \in [0, 1]$ and $\Delta_1, \Delta_2 \in \mathcal{D}(\mathsf{pCSP})$ with $\Delta \Longrightarrow q\cdot\Delta_1 + (1{-}q)\cdot\Delta_2$ and $o_i \in \mathcal{A}(T_i, \Delta_i)$. Then there are stable Γ_i with $[T_i \mid_{\mathsf{Act}} \Delta_i] \Longrightarrow \Gamma_i$ and $\$\Gamma_i = o_i$. Now $[(T_1 \sqcap T_2) \mid_{\mathsf{Act}} \Delta] \Longrightarrow q\cdot[(T_1 \sqcap T_2) \mid_{\mathsf{Act}} \Delta_1] + (1{-}q)\cdot[(T_1 \sqcap T_2) \mid_{\mathsf{Act}} \Delta_2] \stackrel{\tau}{\longrightarrow} q\cdot[T_1 \mid_{\mathsf{Act}} \Delta_1] + (1{-}q)\cdot[T_2 \mid_{\mathsf{Act}} \Delta_2] \Longrightarrow q\cdot\Gamma_1 + (1{-}q)\cdot\Gamma_2$. 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 $q\cdot o_1 + (1{-}q)\cdot o_2 \in \mathcal{A}(T_1 \sqcap T_2, \Delta)$. □

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 6.29 and 6.30 respectively.

Lemma 7.17 Suppose $\Delta \mid_A (T_1 \sqcap T_2) \stackrel{\tau}{\longrightarrow} \Gamma$. 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) $\Delta^{\rightarrow} \stackrel{\tau}{\longrightarrow} \Delta^{\mathrm{next}}$

(iii) $\Gamma = \Delta^{\mathrm{next}} \mid_A (T_1 \sqcap T_2) + \Delta_1^{\times} \mid_A T_1 + \Delta_2^{\times} \mid_A T_2$

Proof: By Lemma 3.4 $\Delta \mid_A (T_1 \sqcap T_2) \stackrel{\tau}{\longrightarrow} \Gamma$ implies that

$$\Delta = \sum_{i\in I} p_i\cdot\overline{s_i}, \qquad s_i \mid_A (T_1 \sqcap T_2) \stackrel{\tau}{\longrightarrow} \Gamma_i, \qquad \Gamma = \sum_{i\in I} p_i\cdot\Gamma_i,$$

for certain $s_i \in S$, $\Gamma_i \in \mathcal{D}(\mathsf{sCSP})$ 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 $s_i \stackrel{\tau}{\longrightarrow} \Gamma_i'$. Now let

$$\Delta^{\rightarrow} = \sum_{i\in(I-J_1-J_2)} p_i\cdot\overline{s_i}, \qquad \Delta_k^{\times} = \sum_{i\in J_k} p_i\cdot\overline{s_i}, \qquad \Delta^{\mathrm{next}} = \sum_{i\in(I-J_1-J_2)} p_i\cdot\Gamma_i' \,.$$

where $k = 1, 2$. By construction (i) and (iii) are satisfied, and (ii) follows by property (2) of Definition 3.2. □

Lemma 7.18 If $\Delta \mid_A (T_1 \sqcap T_2) \Longrightarrow\!\!\!\!\!\rightarrow \Psi$ then there are Φ_1 and Φ_2 such that

(i) $\Delta \Longrightarrow \Phi_1 + \Phi_2$

(ii) $\Phi_1 \mid_A T_1 + \Phi_2 \mid_A T_2 \Longrightarrow\!\!\!\!\!\rightarrow \Psi$

Proof: Suppose $\Delta_0 \mid_A (T_1 \sqcap T_2) \Longrightarrow\!\!\!\!\!\rightarrow \Psi$. We know from Definition 3.12 that there is a collection of subdistributions $\Psi_k, \Psi_k^{\rightarrow}, \Psi_k^{\times}$, for $k \geq 0$, satisfying the properties

$$\begin{array}{lcll} \Delta_0 \mid_A (T_1 \sqcap T_2) & = & \Psi_0 = & \Psi_0^{\rightarrow} + \Psi_0^{\times} \\ \Psi_0^{\rightarrow} & \xrightarrow{\tau} & \Psi_1 = & \Psi_1^{\rightarrow} + \Psi_1^{\times} \\ \vdots & & & \vdots \\ \Psi_k^{\rightarrow} & \xrightarrow{\tau} & \Psi_{k+1} = & \Psi_{k+1}^{\rightarrow} + \Psi_{k+1}^{\times} \\ & & & \vdots \\ & & & \hline \\ & & & \Psi = \sum_{k=0}^{\infty} \Psi_k^{\times} \end{array}$$

and Ψ is stable.

Take $\Gamma_0 := \Psi_0$. By induction on $k \geq 0$, we find distributions $\Gamma_{k+1},\ \Delta_k^{\rightarrow},\ \Delta_{k1}^{\times},\ \Delta_{k2}^{\times},\ \Delta_{k+1}$ 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) $\Delta_k = \Delta_k^{\rightarrow} + \Delta_{k1}^{\times} + \Delta_{k2}^{\times}$

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

(v) $\Gamma_{k+1} = \Delta_{k+1} \mid_A (T_1 \sqcap T_2) + \Delta_{k1}^{\times} \mid_A T_1 + \Delta_{k2}^{\times} \mid_A T_2$

Induction Step: Assume we already have Γ_k and Δ_k. Note that $\Delta_k \mid_A (T_1 \sqcap T_2) \leq \Gamma_k \leq \Psi_k = \Psi_k^{\rightarrow} + \Psi_k^{\times}$ and $T_1 \sqcap T_2$ can make a τ move. Since Ψ is stable, we know that either $\Psi_k^{\times} = \varepsilon$ or $\Psi_k^{\times} \stackrel{\tau}{\not\rightarrow}$. In both cases it holds that $\Delta_k \mid_A (T_1 \sqcap T_2) \leq \Psi_k^{\rightarrow}$. Proposition 3.9 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 7.17.

Let $\Phi_1 = \sum_{k=0}^{\infty} \Delta_{k1}^{\times}$ and $\Phi_2 = \sum_{k=0}^{\infty} \Delta_{k2}^{\times}$. By (iii) and (iv) above we obtain a weak τ move $\Delta \Longrightarrow \Phi_1 + \Phi_2$. For $k \geq 0$, let $\Gamma_k^{\rightarrow} := \Delta_k \mid_A (T_1 \sqcap T_2)$, let $\Gamma_0^{\times} := \varepsilon$ and let $\Gamma_{k+1}^{\times} := \Delta_{k1}^{\times} \mid_A T_1 + \Delta_{k2}^{\times} \mid_A T_2$. Moreover, $\Gamma := \Phi_1 \mid_A T_1 + \Phi_2 \mid_A T_2$. Now all conditions of Definition 3.22 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 3.23 we have $\Phi_1 \mid_A T_1 + \Phi_2 \mid_A T_2 \Longrightarrow\!\!\!\!\!\rightarrow \Psi$. □

Lemma 7.19 If $o \in \mathcal{A}(T_1 \sqcap T_2, \Delta)$ then there are a $q \in [0,1]$ and $\Delta_1, \Delta_2 \in \mathcal{D}(\mathsf{sCSP})$ 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 $o_i \in \mathcal{A}(T_i, \Delta_i)$.

Proof: If $o \in \mathcal{A}(T_1 \sqcap T_2, \Delta)$ then there is an extreme derivative Ψ of $[(T_1 \sqcap T_2) \mid_{\mathsf{Act}} \Delta]$ such that $\$\Psi = o$. By Lemma 7.18 there are $\Phi_{1,2}$ such that

(i) $\Delta \Longrightarrow \Phi_1 + \Phi_2$

(ii) and $[T_1 \mid_{\mathsf{Act}} \Phi_1] + [T_2 \mid_{\mathsf{Act}} \Phi_2] \Longrightarrow\!\!\!\!\!\rightarrow \Psi$.

By Theorem 3.18(ii) there are some subdistributions Ψ_1 and Ψ_2 such that $\Psi = \Psi_1 + \Psi_2$ and $T_i \mid_{\mathsf{Act}} \Phi_i \Longrightarrow\!\!\!\!\!\rightarrow \Psi_i$ for $i = 1, 2$. Let $o_i' = \$\Psi_i$. As Ψ_i is stable we obtain $o_i' \in \mathcal{A}(T_i, \Psi_i)$. 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 $o_i \in \mathcal{A}(T_i, \Delta_i)$ for $i = 1, 2$. □

Proposition 7.20 For every formula $\varphi \in \mathcal{F}$ there exists a pair (T_φ, v_φ) with T_φ an Ω-test and $v_\varphi \in [0,1]^\Omega$ such that

$$\Delta \models \varphi \text{ if and only if } \exists o \in \mathcal{A}(T_\varphi, \Delta) : o \leq v_\varphi. \tag{27}$$

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 Ω is countable 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 (27), 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 $\mathcal{F}$ 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:[2]

- 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 a 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 a 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 (27) above holds; the proof relies on Lemmas 7.16 and 7.19.

- Let $\varphi = \top$. For all $\Delta \in \mathcal{D}(\mathsf{sCSP})$ we have $\Delta \models \varphi$ as well as $\exists o \in \mathcal{A}(T_\varphi, \Delta) : o \leq v_\varphi$, using Lemma 7.16(1).
- Let $\varphi = \mathbf{div}$. Suppose $\Delta \models \varphi$. Then we have that $\Delta \Longrightarrow \varepsilon$. By Lemma 7.16(2), $\vec{0} \in \mathcal{A}(T_\varphi, \Delta)$.
  Now suppose $\exists o \in \mathcal{A}(T_\varphi, \Delta) : o \leq v_\varphi$. This implies $o = \vec{0}$, so by Lemma 7.16(2), $\Delta \Longrightarrow \varepsilon$. Hence $\Delta \models \varphi$.
- Let $\varphi = \mathbf{ref}(A)$ with $A \subseteq \mathsf{Act}$. Suppose $\Delta \models \varphi$. Then $\Delta \Longrightarrow \stackrel{A}{\not\longrightarrow}$. By Lemma 7.16(3), $\vec{0} \in \mathcal{A}(T_\varphi, \Delta)$.
  Now suppose $\exists o \in \mathcal{A}(T_\varphi, \Delta) : o \leq v_\varphi$. This implies $o = \vec{0}$, so $\Delta \Longrightarrow \stackrel{A}{\not\longrightarrow}$ by Lemma 7.16(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, $\exists o \in \mathcal{A}(T_\psi, \Delta') : o \leq v_\psi$. By Lemma 7.16(4), $o \in \mathcal{A}(T_\varphi, \Delta)$.
  Now suppose $\exists o \in \mathcal{A}(T_\varphi, \Delta) : o \leq v_\varphi$. This implies $o(\omega) = 0$, so by Lemma 7.16(4) there is a Δ' with $\Delta \stackrel{a}{\Longrightarrow} \Delta'$ and $o \in \mathcal{A}(T_\psi, \Delta')$. 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, $\exists o_i \in \mathcal{A}(T_i, \Delta) : o_i \leq v_i$. Thus $o := p \cdot o_1 + (1-p) \cdot o_2 \in \mathcal{A}(T_\varphi, \Delta)$ by Lemma 7.16(5), and $o \leq v_\varphi$.
  Now suppose $\exists o \in \mathcal{A}(T_\varphi, \Delta) : o \leq v_\varphi$. Then, using Lemma 7.16(5), $o = p \cdot o_1 + (1-p) \cdot o_2$ for certain $o_i \in \mathcal{A}(T_i, \Delta)$. Recall that T_1, T_2 are Ω-disjoint tests. One has $o_i \leq v_i$ 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 $\Delta_1, \Delta_2 \in \mathcal{D}(\mathsf{sCSP})$ 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 $o_i \in \mathcal{A}(T_i, \Delta_i)$ with $o_i \leq v_i$. Hence, there are $o_i' \in \mathcal{A}(T_i', \Delta_i)$ with $o_i' \leq v_i'$. Thus $o := p \cdot o_1' + (1-p) \cdot o_2' \in \mathcal{A}(T_\varphi, \Delta)$ by Lemma 7.16(6), and $o \leq v_\varphi$.

[2] 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 $\exists o \in \mathcal{A}(T_\varphi, \Delta) : o \leq v_\varphi$. Then, by Lemma 7.19, there are $q \in [0,1]$ and $\Delta_1, \Delta_2 \in \mathcal{D}(\mathsf{sCSP})$ 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 $o'_i \in \mathcal{A}(T'_i, \Delta_i)$. 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, $\frac{1}{2}q = q \cdot o'_1(\omega_1) = o(\omega_1) \leq v_\varphi(\omega_1) = p \cdot v'_1(\omega_1) = \frac{1}{2}p$ and likewise $\frac{1}{2}(1-q) = (1-q) \cdot o'_2(\omega_2) = o(\omega_2) \leq v_\varphi(\omega_2) = (1-p) \cdot v'_2(\omega_2) = \frac{1}{2}(1-p)$. Together, these inequalities say that $q = p$. Exactly as in the previous case one obtains $o'_i \leq v'_i$ for both $i = 1, 2$. Given that $T'_i = T_i \, {}_{\frac{1}{2}}\!\oplus \omega_i$, using Lemma 7.16(5), it must be that $o'_i = \frac{1}{2}o_i + \frac{1}{2}\vec{\omega_i}$ for some $o_i \in \mathcal{A}(T_i, \Delta_i)$ with $o_i \leq v_i$. By induction, $\Delta_i \models \varphi_i$ for $i = 1, 2$, and hence $\Delta \models \varphi$. □

Theorem 7.21 If $\Delta \sqsupseteq^{\Omega}_{\text{pmust}} \Theta$ then $\Delta \sqsupseteq^{\mathcal{F}} \Theta$.

Proof: Suppose $\Delta \sqsupseteq^{\Omega}_{\text{pmust}} \Theta$ and $\Delta \models \varphi$ for some $\varphi \in \mathcal{F}$. Let T_φ be a characteristic test of φ with target value v_φ. Then Proposition 7.20 yields $\exists o \in \mathcal{A}(T_\varphi, \Delta) : o \leq v_\varphi$, and hence, given that $\Delta \sqsupseteq^{\Omega}_{\text{pmust}} \Theta$, by the Smyth preorder we have $\exists o' \in \mathcal{A}(T_\varphi, \Theta) : o' \leq v_\varphi$. Thus $\Theta \models \varphi$. □

8 Simulations and may testing

In this section we follow the same strategy as for failure simulations and testing (Section 6) 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 8.1 [Simulation Preorder] Define $\sqsubseteq_S$ to be the largest relation in $\mathcal{D}_1(S) \times \mathcal{D}_1(S)$ 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 8.1, 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 8.2 [Simulation] Define $\lhd_s$ to be the largest relation in $S \times \mathcal{D}_1(S)$ 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 8.2 differs from the analogous Definition 6.20 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. 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 8.3 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 6.21 we can assume total distributions, and so do not need the second part of its proof where divergence is treated. □

8.1 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 computation structures before extracting values from them, we will be concerned mainly with ω-respecting structures, and for those we have the following.

Lemma 8.4 Let Δ and Θ be two distributions. If Δ is stable and $\Delta \mathrel{\overline{\lhd_s}} \Theta$, then $\mathcal{V}(\Delta) \leq_{\mathrm{Ho}} \mathcal{V}(\Theta)$.

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

(i) $s \not\longrightarrow$ Here $\mathcal{V}(s)=\{0\}$ and since $\mathcal{V}(\Theta)$ is not empty we have $\mathcal{V}(s) \leq_{\mathrm{Ho}} \mathcal{V}(\Theta)$.

(ii) $s \xrightarrow{\omega} \Delta'$ for some Δ' Here $\mathcal{V}(s)=\{1\}$ and $\Theta \Longrightarrow \Theta' \xrightarrow{\omega}$ with $\mathcal{V}(\Theta')=\{1\}$. By Lemma 6.35 specialised to full distributions, we have $1 \in \mathcal{V}(\Theta)$. Therefore, $\mathcal{V}(s) \leq_{\mathrm{Ho}} \mathcal{V}(\Theta)$.

Now for the general case we suppose $\Delta \mathrel{\overline{\lhd_s}} \Theta$. Use Proposition 3.9 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 $\mathcal{V}(s) \leq_{\mathrm{Ho}} \mathcal{V}(\Theta_s)$ for those s, and so $\mathcal{V}(\Delta) = \sum_{\in \lceil \Delta \rceil} \Delta(s) \cdot \mathcal{V}(s) \leq_{\mathrm{Ho}} \sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot \mathcal{V}(\Theta_s) = \mathcal{V}(\Theta)$. □

Lemma 8.5 Let Δ and Θ be distributions in an ω-respecting computation structure. If $\Delta \mathrel{\overline{\lhd_s}} \Theta$, then we have $\mathcal{V}(\Delta) \leq_{\mathrm{Ho}} \mathcal{V}(\Theta)$.

Proof: Since $\Delta \mathrel{\overline{\lhd_s}} \Theta$, we consider subdistributions Δ'' with $\Delta \Longrightarrow\!\!\!\!\!\rightarrow \Delta''$; by distillation of divergence (Lemma 5.8) we have full distributions Δ' and $\Delta'_{1,2}$ 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 $\Delta' \mathrel{\overline{\lhd_s}} \Theta'$. By Proposition 3.9, we can find distributions Θ'_1, Θ'_2 such that $\Theta' = \Theta'_1 \,{}_p\!\oplus\, \Theta'_2$ and $\Delta'_{1,2} \mathrel{\overline{\lhd_s}} \Theta'_{1,2}$.

Since $\lceil \Delta'_1 \rceil = \lceil \Delta'' \rceil$ we have that Δ'_1 is stable. It follows from Lemma 8.4 that $\mathcal{V}(\Delta'_1) \leq_{\mathrm{Ho}} \mathcal{V}(\Theta'_1)$. Thus we finish off with

	$\mathcal{V}(\Delta'')$	
$=$	$\mathcal{V}(p \cdot \Delta'_1)$	$\Delta'' = p \cdot \Delta'_1$
$=$	$p \cdot \mathcal{V}(\Delta'_1)$	linearity of $\mathcal{V}$
$\leq_{\mathrm{Ho}}$	$p \cdot \mathcal{V}(\Theta'_1)$	above argument based on distillation
$=$	$\mathcal{V}(p \cdot \Theta'_1)$	linearity of $\mathcal{V}$
$\leq_{\mathrm{Ho}}$	$\mathcal{V}(\Theta')$	$\Theta' = \Theta'_1 \,{}_p\!\oplus\, \Theta'_2$
$\leq_{\mathrm{Ho}}$	$\mathcal{V}(\Theta)$.	Lemma 6.35 specialised to full distributions

Since Δ'' was arbitrary, we have our result. □

Lemma 8.6 Let Δ and Θ be distributions in an ω-respecting computation structure. If $\Delta \sqsubseteq_S \Theta$, then it holds that $\mathcal{V}(\Delta) \leq_{\mathrm{Ho}} \mathcal{V}(\Theta)$.

Proof: Suppose $\Delta \sqsubseteq_S \Theta$. By Lemma 8.3, there exists some Θ^{match} such that $\Theta \Longrightarrow \Theta^{\text{match}}$ and $\Delta \mathrel{\overline{\lhd_s}} \Theta^{\text{match}}$. By Lemmas 8.5 and 6.35 we obtain $\mathcal{V}(\Delta) \leq_{\mathrm{Ho}} \mathcal{V}(\Theta') \subseteq \mathcal{V}(\Theta)$. □

Theorem 8.7 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 6.33 for simulation, for any test T
implies	$\mathcal{V}([P \mid_{\mathsf{Act}} T]) \leq_{\mathrm{Ho}} \mathcal{V}([Q \mid_{\mathsf{Act}} T])$	$[\cdot]$ is ω-respecting; Lemma 8.6
iff	$\mathcal{A}(T, P) \leq_{\mathrm{Ho}} \mathcal{A}(T, Q)$	(4)
iff	$P \sqsubseteq_{\text{pmay}} Q$.	Definition 4.5

□

Corollary 8.8 If $P \sqsubseteq_S Q$ then $P \sqsubseteq^{\Omega}_{\text{pmay}} Q$.

Proof: Section 4.3.1 recalled our earlier result that Ω-testing is reducible to scalar testing. □

8.2 Completeness

Let $\mathcal{L}$ be the subclass of $\mathcal{F}$ by skipping the $\mathbf{div}$ and $\mathbf{ref}(A)$ clauses. We write $P \sqsubseteq^{\mathcal{L}} Q$ just when $[\![P]\!] \models \varphi$ implies $[\![Q]\!] \models \varphi$. We have the counterparts of Theorems 7.15 and 7.21, with similar proofs.

Theorem 8.9 In a finitary pLTS $\Delta \sqsubseteq^{\mathcal{L}} \Theta$ if and only if $\Delta \sqsubseteq_S \Theta$.

Theorem 8.10 If $P \sqsubseteq^{\Omega}_{\text{pmay}} Q$ then $P \sqsubseteq^{\mathcal{L}} Q$.

Corollary 8.11 If $P \sqsubseteq_{\text{pmay}} Q$ then $P \sqsubseteq_S Q$.

Proof: From Theorems 8.9 and 8.10 we know that if $P \sqsubseteq^{\Omega}_{\text{pmay}} Q$ then $P \sqsubseteq_S Q$. Section 4.3.1 recalled our earlier result that Ω-testing is reducible to scalar testing. So the required result follows. □

9 Conclusion and Related Work

In this paper we continued our previous work [3, 4, 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. To do this it was 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.

Segala [21] 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 $\Delta \in \mathcal{D}_1(S)$ and $\Delta \Longrightarrow \Delta'$ we have $|\Delta'| = 1$. In [15] it has been shown that $\sqsubseteq_{TD}$ coincides with a notion of simulation akin to $\sqsubseteq_S$. Other probabilistic extensions of simulation occurring in the literature are reviewed in [3, 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 & C. Zhang (2007): *Remarks on testing probabilistic processes. ENTCS* 172, pp. 359–397.

[4] 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.

[5] R. De Nicola & M. Hennessy (1984): *Testing equivalences for processes. Theoretical Computer Science* 34, pp. 83–133.

[6] 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.

[7] M. Hennessy (1988): *An Algebraic Theory of Processes*. MIT Press.

[8] M. Hennessy & R. Milner (1985): *Algebraic Laws for Nondeterminism and Concurrency. Journal of the ACM* 32(1), pp. 137–161.

[9] C.A.R. Hoare (1985): *Communicating Sequential Processes*. Prentice-Hall.

[10] C. Jones (1990) *Probabilistic Non-determinism.* Ph.D. Thesis, University of Edinburgh.

[11] 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.

[12] B. Jonsson & Wang Yi (1995): *Compositional testing preorders for probabilistic processes.* In Proc. *LICS'95*, IEEE Computer Society Press, pp. 431–441.

[13] B. Jonsson & Wang Yi (2002): *Testing preorders for probabilistic processes can be characterized by simulations. Theoretical Computer Science* 282(1), pp. 33–51.

[14] S. Lipschutz (1965): *Schaum's outline of theory and problems of general topology.* McGraw-Hill.

[15] N. Lynch, R. Segala & F.W. Vaandrager (2007): *Observing Branching Structure through Probabilistic Contexts. SIAM Journal on Computing* 37(4), pp. 977–1013.

[16] A.K. McIver & C.C. Morgan (2005): *Abstraction, Refinement and Proof for Probabilistic Systems*. Springer.

[17] R. Milner (1989): *Communication and Concurrency*. Prentice-Hall.

[18] E.-R. Olderog & C.A.R. Hoare (1986): *Specification-oriented semantics for communicating processes. Acta Informatica* 23, pp. 9–66.

[19] M. Puterman (1994): *Markov Decision Processes*. Wiley.

[20] R. Segala (1995): *Modeling and Verification of Randomized Distributed Real-Time Systems*. PhD thesis, MIT.

[21] R. Segala (1996): *Testing probabilistic automata.* In Proc. *CONCUR'96*, LNCS 1119, Springer, pp. 299–314.

[22] 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 Assorted counter-examples

A.1 Finite state space is necessary; otherwise...

A.1.1 Distillation of divergence

Distillation of divergence is the notion that if a system contains any divergence, no matter how small, it can be "distilled out" into an equivalent presentation in which states either wholly converge or wholly diverge; the relevant lemma is Lemma 5.8. It's used to show the equivalence of failure simulation and failure e-simulation (Theorem 6.21), and to justify the full-distribution approach in the may-case (Section 8).

Example 3.16 is an infinite-state system over states $s_{2,3,\cdots}$ 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.

In spite of that, there is an infinite-state version of Lemma 5.6 (the underlying fact on which Lemma 5.8 depends). In it the assumption is that the partial divergences are *bounded away from zero*, i.e. there is some $\varepsilon > 0$ such that every state converges with probability at least ε. But Example 3.16 fails this assumption, because the $1/k$ probabilities of convergence become arbitrarily small.

A.1.2 Transitivity of failure simulation

I think Section A.2.4 below can be adapted to provide a counter-example here: replace the intermediate infinitely branching but finite-state process (28) by the finitely branching but infinite-state process Example 3.16.

A.1.3 Equivalence of finite and infinite interpolation

Let the state space be the positive integers and consider the infinite interpolant $\sum_i \overline{i}/2^i$ of the set $\{\overline{i} \mid i \geq 1\}$. It cannot be realised by any finite interpolation, thus invalidating Lemma B.1 when the state space is infinite.

A.1.4 Soundness of failure simulation

See Section A.2.5 below.

A.1.5 Pre-congruence of simple failure similarity

We use a modification of Example 3.16 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 $s_k \lhd^s_{FS} (\overline{s_a}\ {}_{\frac{1}{k}}\!\oplus\ \mathbf{0})$ because the move $s_k \xrightarrow{\tau} (\overline{s_a}\ {}_{\frac{1}{k^2}}\!\oplus\ \overline{s_{k+1}})$ can be matched by a move to $(\overline{s_a}\ {}_{\frac{1}{k^2}}\!\oplus (\overline{s_a}\ {}_{\frac{1}{k+1}}\!\oplus\ \mathbf{0}))$ which simplifies to just $(\overline{s_a}\ {}_{\frac{1}{k}}\!\oplus\ \mathbf{0})$ again — i.e. a sufficient $\xrightarrow{\tau}$-simulating move would be the identity resolution of $\Longrightarrow$.

Now $s_2 \mid_a s_a$ 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, which is a contra-indication for the truth of Lemma 6.31 unless it uses finiteness of the state space somewhere. The role of the infinitely many states seems to be to escape distillation of divergence in the system. To get this kind of counter-example, we want to "trick" our constructions into allowing the failure-simulation of an unboundedly deep τ-tree by deadlock. But in a finite-state system such a tree must diverge either with probability one (in which case the divergence clause of $\lhd^s_{FS}$ kicks in, and prevents the tricky simulation because deadlock does not diverge), or with probability zero (in which case it has no substantive effect anyway). Example 3.16 uses infinitely many states to avoid the two extremes.

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 $s_2 \Longrightarrow s_a\ {}_{\frac{1}{2}}\!\oplus\ \varepsilon$, but this move cannot be matched by $s_a\ {}_{\frac{1}{2}}\!\oplus\ \mathbf{0}$.

A.2 Finite branching is necessary; otherwise...

In the non-probabilistic world, it's not possible to have infinite branching if the state space is finite; so finite branching, as a restriction, is imposed only when the state space is infinite. In the probabilistic world, even a two-element state space allows infinite branching of an associated pLTS: consider the pLTS containing transitions $s_1 \xrightarrow{\tau} (s_2 \,{}_p\!\oplus\, s_3)$ for all p in some infinite set.

Here are some examples of what happens if we don't impose finite branching as well as finiteness of the state space.

A.2.1 Closure of derivatives

An obvious outcome is that we lose the guaranteed closure of derivatives: just take the example above where the p's come from some open set like $(0, 1)$.

A.2.2 Distillation of divergence

We also lose distillation of divergence without finite branching. The essence of the counter-example here is that we collapse all the s_k-states of Example 3.16 onto a single state, which then becomes infinitely branching: we take a system comprising just two states and a k-indexed set of transitions

$$s \xrightarrow{\tau}_k (\llbracket\, \mathbf{0} \,\rrbracket_{1/k^2} \oplus \overline{s}) \quad \text{for } k \geq 2, \tag{28}$$

as illustrated in Figure 7. As we have seen, by taking transitions $s \xrightarrow{\tau}_K \cdot \xrightarrow{\tau}_{K+1} \cdot \xrightarrow{\tau}_{K+2} \cdots$ we can achieve divergence with probability $1 - 1/K$ for arbitrary $K \geq 2$; but (contradicting distillation) neither of the two states $s, \mathbf{0}$ wholly diverges.

The zero-one law Lemma 5.6 is not affected by infinite branching, because it is 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 Lemma 5.8: it depends on Lemma 5.3, which in turn requires finite branching.

For those with some knowledge of sequential probabilistic/demonic semantics, there is an instructive comparison to be made. Straightforward generalisation of the invariant/variant principle for loop correctness to probabilistic programs allows a direct proof of the zero-one law in the fully nondeterministic case [16, Sec. 2.6], even in an infinite state space, yet there seems to be no explicit mention of finite branching. This is explained by the fact that probabilistic/demonic sequential programs are guaranteed to be continuous (as predicate transformers) only when finite branching is imposed (by analogy with the well known connection between continuity and bounded nondeterminism for standard programs); and failure of that continuity invalidates the soundness proof of the loop-correctness rule referred to above, since it affects the definition of loop itself (ω-limit, or not).

Finally –pursuing this point– note that deterministic (but probabilistic) outcomes with infinite support are unproblematic: for example the sequential program $b, n := \mathsf{true}, 0;$ **while** b **do** $n := n{+}1; b := \mathsf{true} \,{}_{1/2}\!\oplus\, \mathsf{false}$ **end** is perfectly legitimate (i.e. continuous), even though it has infinitely many potential final values of n.

A.2.3 Equivalence of failure simulation and extended failure simulation

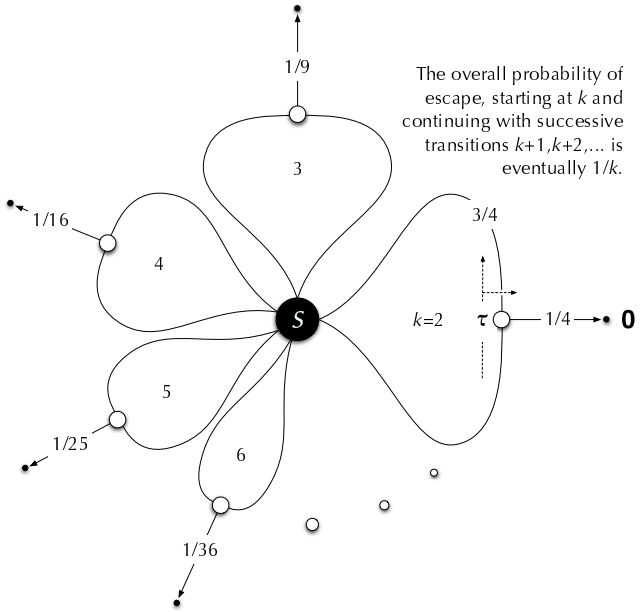
By relating both s and $\mathbf{0}$ (simulated states) to $\mathbf{0}$ (simulating state) we can see that s from (28) above is failure-simulated by just $\mathbf{0}$ itself. Yet we have for example $s \Longrightarrow \llbracket\, \mathbf{0} \,\rrbracket_{1/2} \oplus \varepsilon$, a move which cannot be matched by any extended failure-simulating $\Longrightarrow$-move from $\mathbf{0}$ — and that means that Theorem 6.21 fails in this case. As we'd expect, its proof refers to Lemma 5.8 which in turn refers to Lemma 5.3 where finite branching is assumed.

A.2.4 Transitivity of failure simulation

Define a process $t_0 \xrightarrow{\tau} (\mathbf{0} \,{}_{1/2}\!\oplus\, \overline{t_1})$ and $t_1 \xrightarrow{\tau} \overline{t_1}$, and argue that $t_0 \lhd^s_{FS} \overline{s}$ from (28) above; but also –as observed there– we have $s \lhd^s_{FS} \overline{\mathbf{0}}$. Yet we do not have $t_0 \lhd^s_{FS} \overline{\mathbf{0}}$.

A.2.5 Soundness of failure simulation

Given Section A.2.4, failure simulation and the failure-testing preorder cannot coincide for infinitary processes, since the preorder is transitive trivially. Is it soundness, or completeness that fails? If we compare (28) with $\mathbf{0}$, we have failure simulation of the former by the latter (Appendix A.2.3). But the test $\tau.\omega$ gives $(0, 1]$ as an outcome set for the former and just $\{1\}$ for the latter, breaking the preorder.



There are two states s (large black node) and **0** (small black nodes); the white nodes indicate probabilistic choice; and all transitions are internal. 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.

Example starting from $k = 3$

Loop last used	3	4	5	6	7	8	9	10	$\cdots 15$	$\cdots 20$	$\cdots 30$	$\cdots 40$
Overall escape probability	0.11	0.17	0.20	0.22	0.24	0.25	0.26	0.27	0.29	0.30	0.31	0.32

Figure 7: Infinitely branching flower, from Section A.2.2.

A.2.6 Coincidence of failure simulation and the limit of its approximation

Rob used the following example to show the necessity of the finite branching condition.

Consider a PLTS with four states $s, t, u, v, 0$ and the transitions are

- $s \xrightarrow{a} \overline{0}\,_{\frac{1}{2}}\!\oplus \overline{s}$
- $t \xrightarrow{a} \overline{0}$, $t \xrightarrow{a} \overline{t}$
- $u \xrightarrow{a} \overline{u}$
- $v \xrightarrow{\tau} \overline{u}\,_{p}\!\oplus \overline{t}$ 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 $s \lhd^{k}_{FS} \overline{v}$ for all $k \geq 0$ but we do not have $s \lhd^{s}_{FS} \overline{v}$.

We first observe that $s \lhd^{s}_{FS} \overline{v}$ 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 $s \lhd^{k}_{FS} \overline{v}$ 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

$$s \lhd^{k}_{FS} (\overline{u}\,_{\frac{1}{2^k}}\!\oplus \overline{t}). \tag{29}$$

The base case $k = 0$ is trivial. So suppose we already have (29). We now show that $s \lhd^{(k+1)}_{FS} (\overline{u}\,_{\frac{1}{2^{k+1}}}\!\oplus \overline{t})$. 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 $(\overline{u}\,_{\frac{1}{2^{k+1}}}\!\oplus \overline{t}) \xrightarrow{a} (\overline{0}\,_{\frac{1}{2}}\!\oplus (\overline{u}\,_{\frac{1}{2^k}}\!\oplus \overline{t}))$.

B Technical lemmas and proofs for Section 3

B.1 Infinitary properties of lifting

Lemma B.1 [Infinite interpolation in finite dimensional space] In finite dimensional Euclidean space, infinite interpolation reduces to finite interpolation.

Proof: Suppose we have a set X in N-dimensional Euclidean space and a point x such that x is an infinite interpolation of points in X but is not a finite interpolation of any such points. Since the set of finite interpolants of X is convex (in the usual sense), the point x can therefore be non-strictly separated from it: there is a hyperplane H (thus of dimension $N-1$) containing x such that all the finite interpolants of X lie non-strictly on one side of H. Since x is however an (infinite) interpolant of X, and is in H with all of X on one side of H, in fact x must be an (infinite) interpolant of $X \cap H$. But x is not, of course, a *finite* interpolant of $X \cap H$ (since otherwise it would have been a finite interpolant of X in the first place).

Since $X \cap H$ is a space of (at least) one dimension lower than our original, the argument reduces eventually to a simple line, where it holds trivially. □

Lemma B.2 [Infinitary linearity of lifting] Let $\mathcal{R}$ be a relation between S and $\mathcal{D}(S)$ and, as usual, let $\overline{\mathcal{R}}$ be its lifted version, thus a relation in $\mathcal{D}(S) \times \mathcal{D}(S)$. Let I be an index set, possibly infinite. Then $\Delta_i \;\overline{\mathcal{R}}\; \Theta_i$ for all i implies that $(\sum_i p_i \cdot \Delta_i) \;\overline{\mathcal{R}}\; (\sum_i p_i \cdot \Theta_i)$ where $\sum_i p_i \leq 1$.

Proof: This is immediate from the current definition of lifting.

We note first that $\overline{s} \;\overline{\mathcal{R}}\; \Theta$ just when $\Theta = \sum_j p^j \cdot \Theta^j$ for j ranging over a *finite* index set J such that $s \;\mathcal{R}\; \Theta^j$ for each index j and $1 = \sum_j p^j$. Because we are in a finite state space, however, we can by Lemma B.1 allow J to be infinite.

Now more generally we have that $\Delta \;\overline{\mathcal{R}}\; \Theta$ just when $\Theta = \sum_s \sum_{j_s} p^{j_s} \cdot \Theta^{j_s}$ for j_s ranging over a (possibly infinite) index set J_s, and s itself ranging over the support of Δ, such that $s \;\mathcal{R}\; \Theta^{j_s}$ for each j_s and $\Delta(s) = \sum_{j_s} p^{j_s}$.

The result then follows by straightforward (if intricate) rearrangement of absolutely convergent series:

$$\begin{array}{lll} & \Delta_i \;\overline{\mathcal{R}}\; \Theta_i \quad \text{for all } i & \\ \text{iff} & \quad \Theta_i = \sum_s \sum_{j^i_s} p^{j^i_s} \cdot \Theta^{j^i_s} & \\ \text{and} & s \;\mathcal{R}\; \Theta^{j^i_s} & \text{for each } s \text{ in the support of } \Delta_i \\ \text{and} & \Delta_i(s) = \sum_{j^i_s} p^{j^i_s} & \end{array}$$

implies $\sum_i p_i{\cdot}\Theta_i = \sum_i \sum_s \sum_{j_s^i} p_i{\cdot}p^{j_s^i}{\cdot}\Theta^{j_s^i}$

and $s \mathcal{R} \Theta^{j_s^i}$ for each s in the support of some Δ_i

and $\sum_i p_i{\cdot}\Delta_i(s) = \sum_i \sum_{j_s^i} p_i{\cdot}p^{j_s^i}$

implies some cheating here… rearrangement; suitable definitions of p^{j_s} and Θ^{j_s}

$\sum_i p_i{\cdot}\Theta_i = \sum_s \sum_{j_s} p^{j_s}{\cdot}\Theta^{j_s}$

and $s \mathcal{R} \Theta^{j_s}$ for each s in the support of $\sum_i p_i{\cdot}\Theta_i$

and $(\sum_i p_i{\cdot}\Delta_i)(s) = \sum_{j_s} p^{j_s}$

implies $(\sum_i p_i{\cdot}\Delta_i)\ \overline{\mathcal{R}}\ (\sum_i p_i{\cdot}\Theta_i)$.

□

B.2 Elementary properties of weak derivations

Lemma B.3 Suppose $p_1 + p_2 \leq 1$. Then

(i) $\Delta_{1,2} \Longrightarrow \Delta'_{1,2}$ implies $(p_1{\cdot}\Delta_1 + p_2{\cdot}\Delta_2) \Longrightarrow (p_1{\cdot}\Delta'_1 + p_2{\cdot}\Delta'_2)$.

(ii) If $(p_1{\cdot}\Delta_1 + p_2{\cdot}\Delta_2) \Longrightarrow \Delta'$ then $\Delta' = p_1{\cdot}\Delta'_1 + p_2{\cdot}\Delta'_2$ for subdistributions $\Delta'_{1,2}$ such that $\Delta_{1,2} \Longrightarrow \Delta'_{1,2}$.

Proof: For (i) we note that lifted transitions $\xrightarrow{\tau}$ have that property directly from Clause (2) of Definition 3.2. Thus the structures whose existence is implied by Definition 3.12 for $\Delta_1 \Longrightarrow \Delta'_1$ and $\Delta_2 \Longrightarrow \Delta'_2$ separately can be added with the p_1, p_2 scaling given to form a single composite structure establishing $(p_1{\cdot}\Delta_1 + p_2{\cdot}\Delta_2) \Longrightarrow (p_1{\cdot}\Delta'_1 + p_2{\cdot}\Delta'_2)$.

For (ii) we use an inductive argument, here presented informally. To avoid confusion of subscripts we will effect some renaming and simplification in the demonstrandum, making it read

$$\text{If } \Gamma + \Lambda \Longrightarrow \Pi \text{ then } \Pi = \Pi^{\Gamma} + \Pi^{\Lambda} \text{ with } \Gamma, \Lambda \Longrightarrow \Pi^{\Gamma,\Lambda}. \tag{30}$$

(The $p_{1,2}$ will be re-introduced further below.)

Now from Definition 3.12 we have $\Gamma + \Lambda = \Pi_0 = \Pi_0^{\times} + \Pi_0^{\rightarrow}$ for some $\Pi_0^{\times}, \Pi_0^{\rightarrow}$ with, further, that $\Pi_0^{\rightarrow} \xrightarrow{\tau} \Pi_1$ for some Π_1. Define

$$\begin{array}{rcl} \Gamma^{\rightarrow} & := & \Gamma \sqcap \Pi_0^{\rightarrow} \\ \Gamma^{\times} & := & \Gamma - \Gamma^{\rightarrow} \\ \Lambda^{\times} & := & \Lambda \sqcap \Pi_0^{\times} \\ \Lambda^{\rightarrow} & := & \Lambda - \Lambda^{\times} \ , \end{array} \tag{31}$$

and then check these elementary facts: that $\Gamma^{\times} + \Gamma^{\rightarrow} = \Gamma$ and $\Lambda^{\times} + \Lambda^{\rightarrow} = \Lambda$, and that all the introduced subdistributions are in their proper ranges. What remains is to show that they combine properly, and for that we fix a state s and distinguish two cases: either (a) $\Pi_0^{\rightarrow}.s \geq \Gamma.s$ or (b) $\Pi_0^{\rightarrow}.s \leq \Gamma.s$. In Case (a) the definitions (31) simplify to $\Gamma^{\rightarrow}.s, \Gamma^{\times}.s, \Lambda^{\times}.s, \Lambda^{\rightarrow}.s := \Gamma.s, 0, \Pi_0^{\times}.s, (\Lambda.s - \Pi_0^{\times}.s)$, whence immediately $\Gamma^{\rightarrow}.s + \Lambda^{\rightarrow}.s = \Pi_0^{\rightarrow}.s$ and $\Gamma^{\times}.s + \Lambda^{\times}.s = \Pi_0^{\times}.s$. Case (b) is similar.

Because $\Pi_0^{\rightarrow}$ is $\xrightarrow{\tau}$-enabled, we see that $s \xrightarrow{\tau}\!\!\!\!\!/$ implies $\Pi_0^{\rightarrow}.s = 0$ whence also $\Gamma^{\rightarrow}.s = \Lambda^{\rightarrow}.s = 0$, so that both $\Gamma^{\rightarrow}, \Lambda^{\rightarrow}$ are $\xrightarrow{\tau}$-enabled also. Thus we appeal to Proposition 3.9 to find Γ_1, Λ_1 with $\Gamma^{\rightarrow} \xrightarrow{\tau} \Gamma_1$ and $\Lambda^{\rightarrow} \xrightarrow{\tau} \Lambda_1$ 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 (here the informal induction) to induce derivation structures for Γ, Λ separately that establish (30) when added together as in Part (i) of this lemma.

Finally, we let Γ, Λ, Π be $p_1{\cdot}\Delta_1, p_2{\cdot}\Delta_2, \Delta'$, and scale the induced derivation structures up by $1/p_1, 1/p_2$ respectively. Because later subdistributions in those structures can never be bigger than earlier ones, the proper bounding of $\Delta_{1,2}$ themselves guarantees that the up-scaling does not make any subsequent subdistributions too big. □

Theorem B.4 [Transitivity of $\Longrightarrow$] In a finite-state pLTS, 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

$$\Delta = \Delta_0, \qquad \Delta_k = \Delta_k^{\times} + \Delta_k^{\rightarrow}, \qquad \Delta_k^{\rightarrow} \xrightarrow{\tau} \Delta_{k+1}, \qquad \Theta = \sum_{k=0}^{\infty} \Delta_k^{\times}. \tag{32}$$

Since $\Theta = \Delta_0^{\times} + \sum_{k\geq 1} \Delta_k^{\times}$ and $\Theta \Longrightarrow \Lambda$, by Theorem 3.18 there are $\Lambda_0, \Lambda_1^{\geq}$ such that

$$\Delta_0^{\times} \Longrightarrow \Lambda_0, \qquad \sum_{k\geq 1} \Delta_k^{\times} \Longrightarrow \Lambda_1^{\geq}, \qquad \Lambda = \Lambda_0 + \Lambda_1^{\geq} \ .$$

Using Theorem 3.18 again, we have $\Lambda_1, \Lambda_2^{\geq}$ such that

$$\Delta_1^{\times} \Longrightarrow \Lambda_1, \qquad \sum_{k\geq 2} \Delta_k^{\times} \Longrightarrow \Lambda_2^{\geq}, \qquad \Lambda_1^{\geq} = \Lambda_1 + \Lambda_2^{\geq},$$

thus in combination $\Lambda = \Lambda_0 + \Lambda_1 + \Lambda_2^{\geq}$. Continuing this process we have that

$$\Delta_k^{\times} \Longrightarrow \Lambda_k, \qquad \sum_{j>k} \Delta_j^{\times} \Longrightarrow \Lambda_{k+1}^{\geq}, \qquad \Lambda = \sum_{j=0}^{k} \Lambda_j + \Lambda_{k+1}^{\geq} \tag{33}$$

for all $k \geq 0$. Lemma 3.5 and Proposition 3.19 ensure that $\Delta \Longrightarrow \Theta$ implies $|\Delta| \geq |\Theta|$ for any subdistributions Δ and Θ, and therefore that $|\sum_{j>k} \Delta_j^{\times}| \geq |\Lambda_{k+1}^{\geq}|$ for all $k \geq 0$. But since $\Theta = \sum_{k=0}^{\infty} \Delta_k^{\times}$ from (32), we know that the tail sum $\sum_{j>k} \Delta_j^{\times}$ converges to ε when k approaches ∞, and therefore that $\lim_{k\to\infty} \Lambda_k^{\geq} = \varepsilon$. Thus by taking that limit we conclude that

$$\Lambda \quad = \quad \sum_{k=0}^{\infty} \Lambda_k \ . \tag{34}$$

Now 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

$$\Delta_k^{\times} = \Delta_{k0}, \qquad \Delta_{kl} = \Delta_{kl}^{\times} + \Delta_{kl}^{\rightarrow}, \qquad \Delta_{kl}^{\rightarrow} \xrightarrow{\tau} \Delta_{k,l+1} \qquad \Lambda_k = \sum_{l\geq 0} \Delta_{kl}^{\times}. \tag{35}$$

Therefore we can put all this together with

$$\Lambda \quad = \quad \sum_{k=0}^{\infty} \Lambda_k \quad = \quad \sum_{k,l\geq 0} \Delta_{kl}^{\times} \quad = \quad \sum_{i\geq 0} \left(\sum_{k,l|k+l=i} \Delta_{kl}^{\times} \right) \ , \tag{36}$$

where the last step is a straightforward diagonalisation.

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

$$\Delta'_i = \Delta'^{\times}_i + \Delta'^{\rightarrow}_i, \qquad \Delta'^{\times}_i = \sum_{k,l|k+l=i} \Delta_{kl}^{\times}, \qquad \Delta'^{\rightarrow}_i = (\sum_{k,l|k+l=i} \Delta_{kl}^{\rightarrow}) + \Delta_i^{\rightarrow} \ , \tag{37}$$

so that from (36) we have immediately that

$$\Lambda \quad = \quad \sum_{i\geq 0} \Delta'^{\times}_i \ . \tag{38}$$

We now show that

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

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

from which, with (38), we will have $\Delta \Longrightarrow \Lambda$ as required. For (i) we observe that

	Δ	
$=$	Δ_0	(32)
$=$	$\Delta_0^{\times} + \Delta_0^{\rightarrow}$	(32)
$=$	$\Delta_{00} + \Delta_0^{\rightarrow}$	(35)
$=$	$\Delta_{00}^{\times} + \Delta_{00}^{\rightarrow} + \Delta_0^{\rightarrow}$	(35)
$=$	$(\sum_{k,l|k+l=0} \Delta_{kl}^{\times}) + (\sum_{k,l|k+l=0} \Delta_{kl}^{\rightarrow}) + \Delta_0^{\rightarrow}$	index arithmetic
$=$	$\Delta'^{\times}_0 + \Delta'^{\rightarrow}_0$	(37)
$=$	Δ'_0 .	(37)

For (ii) we observe that

	$\Delta'^{\rightarrow}_i$	
$=$	$(\sum_{k,l|k+l=i} \Delta_{kl}^{\rightarrow}) + \Delta_i^{\rightarrow}$	(37)
$\xrightarrow{\tau}$	$(\sum_{k,l|k+l=i} \Delta_{k,l+1}) + \Delta_{i+1}$	(32), (35), Proposition 3.9
$=$	$(\sum_{k,l|k+l=i} (\Delta_{k,l+1}^{\times} + \Delta_{k,l+1}^{\rightarrow})) + \Delta_{i+1}^{\times} + \Delta_{i+1}^{\rightarrow}$	(32), (35)

$$
\begin{array}{lll}
= & (\sum_{k,l|k+l=i} \Delta^{\times}_{k,l+1}) + \Delta^{\times}_{i+1} + (\sum_{k,l|k+l=i} \Delta^{\rightarrow}_{k,l+1}) + \Delta^{\rightarrow}_{i+1} & \text{rearrange} \\
= & (\sum_{k,l|k+l=i} \Delta^{\times}_{k,l+1}) + \Delta_{i+1,0} + (\sum_{k,l|k+l=i} \Delta^{\rightarrow}_{k,l+1}) + \Delta^{\rightarrow}_{i+1} & (35) \\
= & (\sum_{k,l|k+l=i} \Delta^{\times}_{k,l+1}) + \Delta^{\times}_{i+1,0} + \Delta^{\rightarrow}_{i+1,0} + (\sum_{k,l|k+l=i} \Delta^{\rightarrow}_{k,l+1}) + \Delta_{i+1} & (35) \\
= & (\sum_{k,l|k+l=i+1} \Delta^{\times}_{kl}) + (\sum_{k,l|k+l=i+1} \Delta^{\rightarrow}_{kl}) + \Delta^{\rightarrow}_{i+1} & \text{index arithmetic} \\
= & \Delta'^{\times}_{i+1} + \Delta'^{\rightarrow}_{i+1} & (37) \\
= & \Delta'_{i+1}\ , & (37)
\end{array}
$$

which concludes the proof. □

B.3 Structural properties of weak derivations

Definition B.5 [Reward function] Let a *reward function* be a function $\$: S \to [-1, 1]$ from the state space into the real interval $[-1, 1]$. [3]

Define empty suprema over sets of reward-like values to be $-\infty$. This is only a convenience in the middle of calculations; the infinities always disappear in the end. Write $\$.\Delta$ for the expected value of reward function $\$$ over distribution Δ. Write $(s \Longrightarrow)$ for $\{\Delta' \mid s \Longrightarrow \Delta'\}$ and write $\mathbb{W}_{\sqcap}.\$.s$ for $\sqcap\{\$\Delta' \mid s \Longrightarrow \Delta'\}$. Write $(s \Longrightarrow_{\mathsf{pp}})$ for the single subdistribution that results from using policy pp to construct $(s \Longrightarrow)$; note that $(\Longrightarrow_{\mathsf{pp}})$ is a function (and is total), whereas $(\Longrightarrow)$ is in general a relation (also total). This function is made precise in Definition B.8 below.

Lemma B.6 [Linearity of $\mathbb{W}_{\sqcap}$ and $\mathbb{W}_{\sqcup}$] For any reward function $\$$ the associated $\mathbb{W}_{\sqcap}.\$$ and $\mathbb{W}_{\sqcup}.\$$ are linear.

Proof: We prove the binary case for $\mathbb{W}_{\sqcap}$: extension to any finite linear combination is straightforward, and the argument for $\mathbb{W}_{\sqcup}$ is the same.

Suppose $\Delta = \Delta_1 \,{}_p\!\oplus\, \Delta_2$ for some p and $\Delta_{\{1,2\}}$. Define $m_{\{1,2\}} := \mathbb{W}_{\sqcap}.\$.\Delta_{\{1,2\}}$ respectively. Then for any $\varepsilon > 0$ there are $\Delta'_{\{1,2\}}$ with $\Delta_{\{1,2\}} \Longrightarrow \Delta'_{\{1,2\}}$ and $\$.\Delta'_{\{1,2\}} \leq m_{\{1,2\}} + \varepsilon$. By linearity of $\Longrightarrow$ we therefore have $\Delta' := \Delta'_1 \,{}_p\!\oplus\, \Delta'_2$ with $\Delta \Longrightarrow \Delta'$ and $\$.\Delta' \leq (m_1 \,{}_p\!\oplus\, m_2) + \varepsilon$. Since ε is arbitrary, the result follows. □

Our overall aim is to establish that every element in $(s \Longrightarrow)$ is the interpolation of a finite number of static derivatives Δ'_i, that is in each case $s \Longrightarrow_{\mathsf{pp}_i} \Delta'_i$. In order to do that, however, we must introduce the notion of discounted derivatives.

Definition B.7 [Discounted derivatives] Define $s \Longrightarrow_{\delta} \Delta'$ for *discount* $0 \leq \delta \leq 1$ by analogy with our earlier definition Definition 3.12 of derivative except that each $\xrightarrow{\tau}$-move discounts its outcome by δ. That is, we revise that earlier definition by including multiplications by δ at the appropriate points:

$$
\begin{array}{lll}
\Delta = \Delta^{\rightarrow}_0 + \Delta^{\times}_0 & & \text{— The } \times \text{ component stops “here,”} \\
\delta\cdot\Delta^{\rightarrow}_0 \xrightarrow{\tau} \Delta^{\rightarrow}_1 + \Delta^{\times}_1 & & \text{— but the } \rightarrow \text{ component moves on, discounted by } \delta. \\
\vdots \qquad\qquad\quad \vdots & & \\
\delta\cdot\Delta^{\rightarrow}_k \xrightarrow{\tau} \Delta^{\rightarrow}_{(k+1)} + \Delta^{\times}_{(k+1)} & & \\
\qquad\qquad\qquad\quad \vdots & & \\
\hline
\text{In total:} \quad \Delta' & & \text{— Finally, all the stopped components are summed.}
\end{array}
$$

Then we call $\Delta' := \sum_{k=0}^{\infty} \Delta^{\times}_k$ a *δ-discounted derivative* of Δ, and write $\Delta \Longrightarrow_{\delta} \Delta'$ to mean that Δ can make a *weak δ-discounted τ move* to a δ-discounted derivative Δ'.

Note that $(\Longrightarrow_1)$ and $(\Longrightarrow)$ agree trivially.

Finally, we combine the two notions of static policy and discount:

Definition B.8 [Discounted *SDP*-derivatives] Define $s \Longrightarrow_{\delta,\mathsf{pp}} \Delta'$ as follows:

$$
\begin{array}{lll}
\Delta_0 & = & \Delta \\
\Delta^{\times}_k.s & = & 0 \text{ if } \mathsf{pp} \text{ defined at } s \text{ else } \Delta_k.s \\
\Delta_{k+1} & = & \displaystyle\sum_{\substack{s \,\in\, \lceil \Delta_k \rceil \\ \mathsf{pp} \text{ defined at } s}} \delta\cdot \mathsf{pp}(s)
\end{array}
$$

[3] In later sections we will restrict reward functions to the non-negatives [0,1]; but here we need the more general range. Maybe more explanations for the reason of needing this general range?

Then, as usual, we sum the stopped distributions to define $\Delta' := \sum_{k=0}^{\infty} \Delta_k^{\times}$.

Note that $(s \Longrightarrow_{\mathsf{pp}})$, which we defined informally above, is given by $(s \Longrightarrow_{1,\mathsf{pp}})$.

Definition B.9 [discounted payoff] For reward function $\$$ and discount δ define the discounted maximising payoff function $\mathbb{W}^{\delta}_{\sqcup}.\$.s := \sqcup\{\$\Delta' \mid s \Longrightarrow_{\delta} \Delta'\}$.

The payoff function $\mathbb{W}^{\delta}_{\sqcup}$, applied to s, thus gives the supremal expected reward determined by $\$$ over all distributions reached by starting at s and using $\Longrightarrow_{\delta}$. Pre-calculating this function is the key technique in defining important policies: for us, they will be the maximising ones.

For the moment we will just show that the function $\mathbb{W}^{\delta}_{\sqcup}.\$$ (over S) is a fixed point.

Lemma B.10 [$\mathbb{W}^{\delta}_{\sqcup}.\$$ as a fixed point] In the usual context (suitable $S, \delta, \$$, underlying pLTS and $s \in S$) we have

$$\mathbb{W}^{\delta}_{\sqcup}.\$.s \quad = \quad \$s \;\sqcup\; \delta \cdot \sqcup \{\mathbb{W}^{\delta}_{\sqcup}.\$.\Delta' \mid s \xrightarrow{\tau} \Delta'\}\,,$$

that is $\mathbb{W}^{\delta}_{\sqcup}.\$$ for any $\delta, \$$ (including $\delta = 1$) is a fixed point of the function $(\lambda W.(\lambda s.\$s \sqcup \delta \cdot \sqcup \{W.\Delta' \mid s \xrightarrow{\tau} \Delta'\}))$.

Note that if $s \not\xrightarrow{\tau}$ on the right-hand side (for some argument) then the supremum is empty, giving $-\infty$, as we mentioned above; multiplied by δ it is still $-\infty$; then $\sqcup$'d with $\$s$ the $-\infty$ disappears. (Doing this avoids a case analysis in at least two places.)

Proof: We rely on splitting $\Longrightarrow_{\delta}$ up in a "first one step, and then the rest" -style. This gives us first

$$\begin{array}{ll} & (s \Longrightarrow_{\delta}) \\ = & \{\Delta_0^{\times} + \delta\cdot\Delta' \mid \overline{s} = \Delta_0^{\times} + \Delta_0^{\rightarrow} \wedge \Delta_0^{\rightarrow} \xrightarrow{\tau} \Delta_1 \wedge \Delta_1 \Longrightarrow_{\delta} \Delta' \quad \text{for some } \Delta_0^{\times}, \Delta_0^{\rightarrow}, \Delta_1\}\,, \end{array}$$

which allows us to calculate further as follows:

$$\begin{array}{lll} & \mathbb{W}^{\delta}_{\sqcup}.\$.s & \\ = & \sqcup\{\$\Delta' \mid s \Longrightarrow_{\delta} \Delta'\} & \text{definition} \\ = & \sqcup\{\$.(\Delta_0^{\times} + \delta\cdot\Delta') \mid \overline{s} = \Delta_0^{\times} + \Delta_0^{\rightarrow} \wedge \Delta_0^{\rightarrow} \xrightarrow{\tau} \Delta_1 \wedge \Delta_1 \Longrightarrow_{\delta} \Delta' \quad \text{for some } \Delta_0^{\rightarrow}, \Delta_1\} & \text{above} \\ = & \sqcup\{\$.\Delta_0^{\times} + \delta\cdot\$\Delta' \mid \overline{s} = \Delta_0^{\times} + \Delta_0^{\rightarrow} \wedge \Delta_0^{\rightarrow} \xrightarrow{\tau} \Delta_1 \wedge \Delta_1 \Longrightarrow_{\delta} \Delta' \quad \text{for some } \Delta_0^{\rightarrow}, \Delta_1\} & \\ = & \sqcup\{\$.\Delta_0^{\times} + \delta\cdot\sqcup\{\$\Delta' \mid \overline{s} = \Delta_0^{\times} + \Delta_0^{\rightarrow} \wedge \Delta_0^{\rightarrow} \xrightarrow{\tau} \Delta_1 \wedge \Delta_1 \Longrightarrow_{\delta} \Delta' \quad \text{for some } \Delta_0^{\rightarrow}, \Delta_1\}\} & \\ = & & \text{lift } \mathbb{W}^{\delta}_{\sqcup}.\$\text{; linearity of } \Longrightarrow_{\delta} \\ & \sqcup\{\$.\Delta_0^{\times} + \delta\cdot\sqcup\{\mathbb{W}^{\delta}_{\sqcup}.\$.\Delta_1 \mid \overline{s} = \Delta_0^{\times} + \Delta_0^{\rightarrow} \wedge \Delta_0^{\rightarrow} \xrightarrow{\tau} \Delta_1 \quad \text{for some } \Delta_0^{\rightarrow}\}\} & \\ = & \sqcup\{\$s \;{}_p\!\oplus\; \delta\cdot\sqcup\{\mathbb{W}^{\delta}_{\sqcup}.\$.\Delta_1 \mid s \xrightarrow{\tau} \Delta_1\} \quad \text{for some } 0 \le p \le 1\} & \overline{s} \text{ can be split only into } \overline{s} \;{}_p\!\oplus\; \overline{s} \\ = & \$s \;\sqcup\; \delta\cdot\sqcup\{\mathbb{W}^{\delta}_{\sqcup}.\$.\Delta_1 \mid s \xrightarrow{\tau} \Delta_1\}\,, & \sqcup \text{ over interpolation of scalars must be one or the other} \end{array}$$

as required. □

Given that Lemma B.10 expresses $\mathbb{W}^{\delta}_{\sqcup}.\$$ as a fixed point, and works for $\mathbb{W}^{1}_{\sqcup}.\$$ in particular, one could ask why we bother with the δ. The answer is that in the $\delta{=}1$ case it's not clear which fixed point we have, whereas for $\delta{<}1$ the fixed point is unique, as we will see by considering optimal discounted policies.

Definition B.11 [max-seeking policy] Given a pLTS, reward function $\$$ and discount δ say that a *SDP* given by pp is *max-seeking* just when for all s we have

1. if pp is undefined at s then $\$s \ge \delta\cdot\mathbb{W}^{\delta}_{\sqcup}.\$.\Delta'$ for all Δ' with $s \xrightarrow{\tau} \Delta'$.
2. if pp is defined at s then both
   (a) $\delta\cdot\mathbb{W}^{\delta}_{\sqcup}.\$.(\mathsf{pp}\,.s) \ge \$s$ and
   (b) $\mathbb{W}^{\delta}_{\sqcup}.\$.(\mathsf{pp}\,.s) \ge \mathbb{W}^{\delta}_{\sqcup}.\$.\Delta'$ for all Δ' with $s \xrightarrow{\tau} \Delta'$.

What a max-seeking policy does is to evaluate $\mathbb{W}^{\delta}_{\sqcup}.\$$ "in advance" (given $\$, \delta$), and then label every state s with the expected payoff value $\mathbb{W}^{\delta}_{\sqcup}.\$.s$ that $\mathbb{W}^{\delta}_{\sqcup}.\$$ assigns to it. The policy at any state s is then to compare the reward $\$s$ at s itself with the expected label values $\mathbb{W}^{\delta}_{\sqcup}.\$.s'$ over each distribution of states s' that are successors of s, and then to select the greatest among all those. (Note the use of finite branching.) We will see that in a properly discounted system (i.e. with $\delta < 1$) such a policy maximises the payoff.

In case that seems obvious, we now give an undiscounted case (i.e. δ=1) where a max-seeking policy doesn't work. Take the system

$$\begin{array}{lll} s_0 & \xrightarrow{\tau} & \overline{s_0} \\ s_0 & \xrightarrow{\tau} & \overline{s_0}\,{}_{1/2}\!\oplus\, \overline{s_1} \\ s_1 & \not\longrightarrow & \end{array}$$

with $\$s_0 = 0$ and $\$s_1 = 1$,

and observe that from both states a payoff of 1 is attainable eventually; that is, we have $\mathbb{W}^\delta_\sqcup.\$.s_{0,1} = 1$ and both states will be $\mathbb{W}^\delta_\sqcup.\$$ -labelled with 1. At s_0 therefore the policy compares "stay here," yielding immediate payoff $\$s_0 = 0$, with "move to s_0" with $\mathbb{W}$-label 1 and "move to $\overline{s_0}\,{}_{1/2}\!\oplus\, \overline{s_1}$" also with $\mathbb{W}$-label 1. Clearly one of the latter two is chosen — but which? If it is the first, then in fact the overall payoff will be 0 because of divergence — the maximum is not attained, and the policy has failed.

On the other hand, with a proper discount the payoff at s_0 would be given by $\mathbb{W}^\delta_\sqcup.\$.s_0 = \delta\cdot(\mathbb{W}^\delta_\sqcup.\$.s_0\ {}_{1/2}\!\oplus\ 1)$, that is $\delta/(2-\delta)$. Then the reward (immediate payoff) at s_0 is still 0, the expected payoff for taking $s_0 \xrightarrow{\tau} s_0$ has become $\delta^2/(2-\delta)$ and the expected payoff for taking $s_0 \xrightarrow{\tau} (\overline{s_0}\,{}_{1/2}\!\oplus\, \overline{s_1})$ has become $\delta/(2-\delta)$. Although the two expressions are equal for δ=1, when $0<\delta<1$ the last choice is strictly greatest, and so the policy succeeds by taking that choice every time (therefore being static). Note that although here there is a unique max-seeking policy, in general there could be many.

We will now show that in a properly discounted system (δ<1), choosing one of the max-seeking *SDP*'s and following it consistently will achieve the maximum expected reward, just as immediately above. (One cannot change *SDP*'s "mid-stream," however.) The proof relies on showing two expressions are fixed points of the same function, and on then using the discount to argue via contraction that in fact the function's fixed point is unique.

We begin by noting that following a policy cannot increase the payoff.

Definition B.12 [Policy-following payoff] Define $\mathbb{W}^{\delta,\mathsf{pp}}.\$.s$ to be $\$\Delta'$ where $s \Longrightarrow_{\delta,\mathsf{pp}} \Delta'$. Note that Δ' is unique.

Lemma B.13 [Policy-following does not increase payoff] For any policy pp and discount δ we have $\mathbb{W}^{\delta,\mathsf{pp}} \le \mathbb{W}^\delta_\sqcup$.

Proof: Trivially $s \Longrightarrow_{\delta,\mathsf{pp}} \Delta'$ implies $s \Longrightarrow_\delta \Delta'$, whence we have $\mathbb{W}^{\delta,\mathsf{pp}}.\$.s \le \mathbb{W}^\delta_\sqcup.\$.s$ because $\mathbb{W}_\sqcup$ is defined as a supremum. The desired inequality follows because $\delta, \$, s$ are arbitrary. □

Additionally we can develop a fixed-point characterisation of $\mathbb{W}^{\delta,\mathsf{pp}}$ in the style of Lemma B.10.

Lemma B.14 [$\mathbb{W}^{\delta,\mathsf{pp}}$ as fixed point] For any discount δ and policy pp we have

$$\mathbb{W}^{\delta,\mathsf{pp}}.\$.s \quad = \quad \$s \text{ if } \mathsf{pp} \text{ undefined at } s \text{ else } \delta\cdot\mathbb{W}^{\delta,\mathsf{pp}}.\$.(\mathsf{pp}\,.s)\ . \tag{39}$$

Proof: We proceed as in Lemma B.10, noting first that we have

$$s \Longrightarrow_{\delta,\mathsf{pp}} (\overline{s} \text{ if } \mathsf{pp} \text{ undefined at } s \text{ else } \delta\cdot\Delta') \qquad \text{where } (\mathsf{pp}\,.s) \Longrightarrow_{\delta,\mathsf{pp}} \Delta'\ , \tag{40}$$

where again we exploit that Δ' is unique. Then we calculate

$$\begin{array}{lll} & \mathbb{W}^{\delta,\mathsf{pp}}.\$.s & \\ = & \$.(\overline{s} \text{ if } \mathsf{pp} \text{ undefined at } s \text{ else } \delta\cdot\Delta') \qquad \text{where } (\mathsf{pp}\,.s) \Longrightarrow_{\delta,\mathsf{pp}} \Delta' & \text{Definition B.12, (40)} \\ = & \$s \text{ if } \mathsf{pp} \text{ undefined at } s \text{ else } \delta\cdot\$\Delta' \qquad \text{where } (\mathsf{pp}\,.s) \Longrightarrow_{\delta,\mathsf{pp}} \Delta' & \\ = & \$s \text{ if } \mathsf{pp} \text{ undefined at } s \text{ else } \delta\cdot\mathbb{W}^{\delta,\mathsf{pp}}.\$.(\mathsf{pp}\,.s)\ . & \end{array}$$

□

Now we concentrate on max-seeking policies, and show that $\mathbb{W}^\delta_\sqcup.\$$ satisfies the fixed-point equation (39) provided the policy pp it refers to is max-seeking.

Lemma B.15 [fixed-point for max-seeking policy] Let pp be a max-seeking policy. Then

$$\mathbb{W}^\delta_\sqcup.\$.s \quad = \quad \$s \text{ if } \mathsf{pp} \text{ is undefined at } s \text{ else } \delta\cdot\mathbb{W}^\delta_\sqcup.\$.(\mathsf{pp}\,.s)\ . \tag{41}$$

Proof: We calculate

$$\begin{array}{lll} & \$s \text{ if } \mathsf{pp} \text{ is undefined at } s \text{ else } \delta\cdot\mathbb{W}^\delta_\sqcup.\$.(\mathsf{pp}\,.s) & \\ = & \$s \ \sqcup\ \delta\cdot\sqcup\{\mathbb{W}^\delta_\sqcup.\$.\Delta' \mid s \xrightarrow{\tau} \Delta'\} & \mathsf{pp} \text{ is max-seeking} \\ = & \mathbb{W}^\delta_\sqcup.\$.s\ . & \text{Lemma B.10} \end{array}$$

□

Putting the above together, and exploiting contraction, gives us our first important proposition concerning max-seeking policies.

Proposition B.16 [discounted max-seeking policies are optimal] If pp is a max-seeking policy with respect to δ and $\$$, and $\delta < 1$, then $\mathbb{W}^{\delta,\mathsf{pp}}.\$ = \mathbb{W}^{\delta}_{\sqcup}.\$$.

Proof: We use the results above to recall that each side is a fixed point of the equation

$$\mathcal{W}_{\delta,\$} \quad = \quad (\lambda s.\ \$s \text{ if } \mathsf{pp} \text{ undefined at } s \text{ else } \delta \cdot \mathcal{W}_{\delta,\$}.(\mathsf{pp}\,.s))\ , \tag{42}$$

and then argue that the fixed point is unique because of contraction and completeness. The first point is immediate from (41) and (39).

For the second, we recall that we are working with functions from S into $[-1,1]$ and show that (42) is a contraction mapping. Write $\widehat{\mathcal{W}}^{i}_{\delta,\$}$ for Equation 42's right-hand side applied to some $\mathcal{W}^{i}_{\delta,\$}$. Then we have

$$\begin{array}{lll}
 & \text{dist}(\widehat{\mathcal{W}}^{1}_{\delta,\$}, \widehat{\mathcal{W}}^{2}_{\delta,\$}) & \\
= & \sqcup\{\text{abs}(\widehat{\mathcal{W}}^{1}_{\delta,\$}.s - \widehat{\mathcal{W}}^{2}_{\delta,\$}.s) \mid s \in S\} & \text{definition of metric dist} \\
= & \sqcup\{\text{abs}(\widehat{\mathcal{W}}^{1}_{\delta,\$}.s - \widehat{\mathcal{W}}^{2}_{\delta,\$}.s) \mid s \in S \wedge \mathsf{pp} \text{ defined at } s\} & \text{equal if } \mathsf{pp} \text{ undefined at } s \\
= & \delta \cdot \sqcup\{\text{abs}(\mathcal{W}^{1}_{\delta,\$}.(\mathsf{pp}\,.s) - \mathcal{W}^{2}_{\delta,\$}.(\mathsf{pp}\,.s)) \mid s \in S \wedge \mathsf{pp} \text{ defined at } s\} & (42) \\
\leq & \delta \cdot \sqcup\{\text{abs}(\mathcal{W}^{1}_{\delta,\$}.s' - \mathcal{W}^{2}_{\delta,\$}.s') \mid s' \in S\} & \text{arithmetic; } \mathsf{pp}\,.s \text{ is a subdistribution} \\
= & \delta \cdot \text{dist}(\mathcal{W}^{1}_{\delta,\$}, \mathcal{W}^{2}_{\delta,\$}) & \text{definition of metric dist} \\
< & \text{dist}(\mathcal{W}^{1}_{\delta,\$}, \mathcal{W}^{2}_{\delta,\$})\ , & \delta < 1
\end{array}$$

which gives us the contraction, whence the uniqueness and finally the equality. □

The last step is to show that we can achieve optimality with a *SDP* even without discounting; the argument closely follows the MDP approach [19, p284]. First we need to investigate the behaviour of discounted outcomes as the discount δ tends to 1.

Lemma B.17 Suppose we have $s \Longrightarrow \Delta'$ for given s, Δ' and thus $\Delta' = \sum_i \Delta^{\times}_i$ for some properly related sequence of subdistributions $\Delta^{\times}_i$. Then for any reward function $\$$ we have $\$\Delta' = \lim_{\delta \to 1} \sum_i \delta^i \cdot \$\Delta^{\times}_i$.

Proof: Note that, since $\$$ and the $\Delta^{\times}_i$ are fixed, this series is of the form $\sum_i a_i \delta^i$ with $\sum_i |a_i| \leq 1$ — that is, although the coefficients a_i can individually be negative as well as positive, as a series they are absolutely convergent.

Taking limits in δ on both sides then gives us

$$\begin{array}{lll}
 & \lim_{\delta \to 1} \sum_i \delta^i \cdot \$\Delta^{\times}_i & \\
= & \sum_i \lim_{\delta \to 1} \delta^i \cdot \$\Delta^{\times}_i & \text{absolute convergence} \\
= & \sum_i \$\Delta^{\times}_i & \\
= & \$.(\sum_i \Delta^{\times}_i) & \\
= & \$\Delta'\ . &
\end{array}$$

□

That leads immediately to this corollary.

Corollary B.18 For any $\$$, pp we have $\lim_{\delta \to 1} \mathbb{W}^{\delta,\mathsf{pp}}.\$ \ = \ \mathbb{W}^{1,\mathsf{pp}}.\$$.

Proof: Given δ, let the sequence $\Delta^{\times}_i$ be generated as in Definition B.8 by $s \Longrightarrow_{\delta,\mathsf{pp}}$, and apply Lemma B.17 immediately above. □

Now fix the reward $\$$ and suppose we have an infinite sequence of discounts tending to 1. From Proposition B.16 we know that for each of these discounts there is an optimal *SDP*; but we also know from the fact the state space is finite that there are only a finite number of *SDP*'s to choose from along the sequence. Thus at least one occurs infinitely often in connection with our sequence of discounts: let that optimal policy be pp and let $\delta_1, \delta_2, \cdots$ be the infinite (sub-)sequence of discounts associated with it. Note that $\lim_n \delta_n = 1$.

With that preparation we now have our main theorem concerning existence of optimal strategies.

Theorem B.19 In a finitary pLTS we have for any reward function $\$$ that $\mathbb{W}_{\sqcup}.\$ = \mathbb{W}^{\mathsf{pp}}.\$$ for some static policy pp. Note that pp can depend on $\$$.

Proof: By definition we have for any s that $\mathbb{W}_{\sqcup}.\$.s = \sqcup\{\$\Delta' \mid s \Longrightarrow \Delta'\}$. Pick one such Δ' and suppose that it is $\Longrightarrow$-generated as $\sum_i \Delta_i^{\times}$. We now reason as follows:

$$
\begin{array}{lll}
 & \$\Delta' & \\
= & \lim_{n\to\infty} \sum_i \delta_n^i \cdot \$\Delta_i^{\times} & \text{Lemma B.17} \\
\leq & \lim_{n\to\infty} \mathbb{W}^{\delta_n,\mathsf{pp}}.\$.s & \text{from above: that same } \mathsf{pp} \text{ is optimal for all } n \\
= & \mathbb{W}^{1,\mathsf{pp}}.\$.s & \text{Corollary B.18} \\
= & \mathbb{W}^{\mathsf{pp}}.\$.s\ . &
\end{array}
$$

Since Δ' was chosen arbitrarily we have therefore $\mathbb{W}_{\sqcup}.\$.s \leq \mathbb{W}^{\mathsf{pp}}.\$.s$, and the other direction is trivial. □

Lemma B.20 [$\Longrightarrow$ realised by interpolation of finitely many static policies] Suppose $s \Longrightarrow \Delta'$ for some state s and subdistribution Δ'. Then there is a finite index set I, probabilities p_i summing to 1 and static strategies pp_i such that $\Delta' = \sum_{i\in I} p_i \cdot \Delta_i'$ where uniquely $s \Longrightarrow_{\mathsf{pp}_i} \Delta_i'$ for each i.

Proof: For simplicity let I index *all* static policies in our pLTS: this is possible because from finiteness of the state-space and finite branching we know there are only finitely many such strategies; and it loses no generality.

Suppose for a contradiction that $s \Longrightarrow \Delta'$ for some Δ' that does not lie in the closed and convex set of interpolants X of the finite set $\{\Delta_i' \mid s \Longrightarrow_{\mathsf{pp}_i} \Delta_i'\}$. Thus Δ' can be separated from X by a hyperplane H whose normal can be scaled into $[-1, 1]$ because we are in finitely many dimensions. Use that scaled normal to define a reward function $\$_H$ such that $\$_H.\Delta' > c$ but $\$_H.\Delta'_X < c$ for all $\Delta'_X \in X$ and in particular $\$_H.\Delta_i' < c$ for all Δ_i', where c is the constant term of the hyperplane.

Since $\$_H.\Delta' > c$ we must have $\mathbb{W}_{\sqcup}.\$_H.s > c$ also; yet $\mathbb{W}^{\mathsf{pp}_i}.\$_H.s < c$ for all i, contradicting Theorem B.19. Thus there can be no such Δ'. □

Lemma B.21 [Distillation of divergence, static case] If for some state s and static derivative policy pp over a finite-state pLTS there is a derivation $s \Longrightarrow_{\mathsf{pp}} \Delta'$ then there is a probability p and full distributions $\Delta_1', \Delta_\varepsilon'$ such that $s \Longrightarrow (\Delta_1' \ {}_p\!\oplus\ \Delta_\varepsilon')$ and $\Delta' = p\cdot\Delta_1'$ and $\Delta_\varepsilon' \Longrightarrow \varepsilon$.

Proof: Make a modified policy pp' by setting $\mathsf{pp}'.s = \mathsf{pp}.s$ except when $s \Longrightarrow_{\mathsf{pp}} \varepsilon$, in which case make pp' undefined at s. Recalling the definition Definition B.8 of $\Longrightarrow_{\mathsf{pp}'}$, determine the unique Δ'' so that $s \Longrightarrow_{\mathsf{pp}'} \Delta''$.

Now use Lemma 5.6 to observe that $|\Delta''|{=}1$ because no state wholly diverges. We conclude by splitting Δ'' up into $\Delta_1'' + \Delta_\varepsilon''$ so that the support of Δ_ε'' is all the pp-diverging states and Δ_1'' is supported by all the rest, and finally determine $\Delta_1', \Delta_\varepsilon'$ and p by normalisation of those. □