Compositional reasoning for Markov decision processes *

Yuxin Deng[1] † Matthew Hennessy[2‡]
[1] Shanghai Jiao Tong University
[2] Trinity College Dublin

Abstract

Markov decision processes (MDPs) have long been used to model quantitative aspects of systems in the presence of uncertainty. However, much of the literature on MDPs takes a monolithic approach, by modelling a system as a particular MDP; properties of the system are then inferred by analysis of that particular MDP. In this paper we develop compositional methods for reasoning about the quantitative behaviour of MDPs. We consider a class of labelled MDPs called weighted MDPs from a process algebraic point of view. For these we define a coinductive simulation-based behavioural preorder which is compositional in the sense that it is preserved by structural operators for constructing MDPs from components.

For finitary convergent processes, which are finite-state and finitely branching systems without divergence, we provide two characterisations of the behavioural preorder. The first uses a novel quantitative probabilistic logic, while the second is in terms of a novel form of testing, in which benefits are accrued during the execution of tests.

1 Introduction

Markov decision processes (MDPs) have long been used to model quantitative aspects of systems in the presence of uncertainty [Put94, RKNP04, BK08]. A comprehensive account of analysis techniques may be found in [Put94], while [RKNP04] provides a good account of *model-checking*.

However much of the literature on MDPs takes a monolithic view of systems; essentially a system is modelled using a particular MDP, and properties of the system are then inferred by analysis of that MDP. Similar phenomenon exists for the related model of weighted automata [DKV09]. In this paper, instead, we would like to develop compositional methods for reasoning about quantitative behaviour of Markov decision processes. This involves devising a method for comparing the behaviour MDPs which is susceptible to compositional analysis; the behaviour of a composite system should be determined by that of its components.

Our starting point is the idea of one system being able to *simulate* another. For example consider the three systems in Figure 1. The first, a two-state machine, continually performs an

*An extended abstract of this paper has been presented at FSEN'11.

†Partially supported by the Natural Science Foundation of China (61173033, 61033002).

‡Supported financially by SFI project no. SFI 06 IN.1 1898.

Preprint to appear in Science of Computer Programming

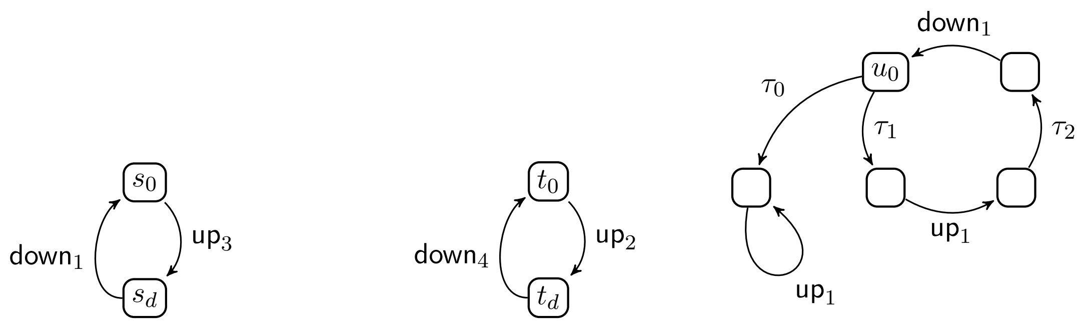

Figure 1: Nondeterministic machines

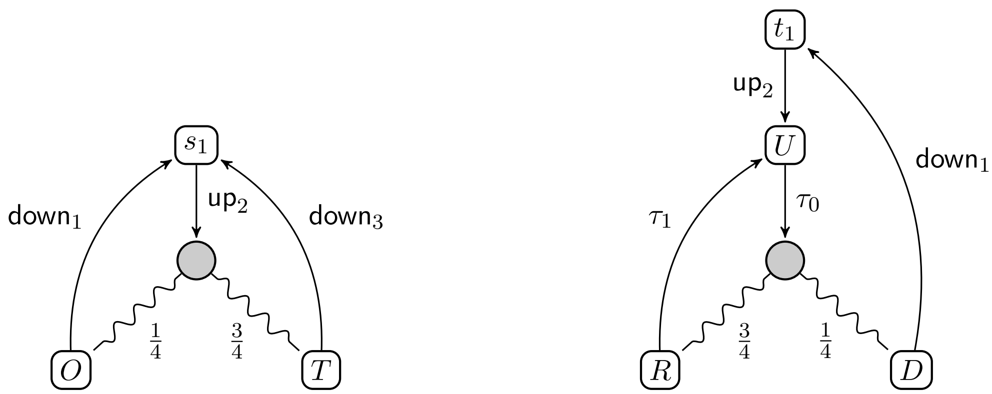

Figure 2: Probabilistic systems

up action, which accrues a benefit of 3 units, followed by a down action, which accrues a benefit of 1. The second machine performs the same actions but with benefits 2 and 4 respectively. In some sense t_0 is an improvement on s_0; intuitively t_0 can simulate the behaviour of s_0 but in so doing accrue more benefits; this is true even if one of its actions up is less beneficial than the corresponding action of s_0. The same is true for the machine u_0; it can also simulate the behaviour of s_0, with more benefit, although in this case some internal weighted actions, denoted by τ, participate in the simulation and add to the accumulation of benefits. In our terminology we will write $s_0 \sqsubseteq_{sim} t_0$, $s_0 \sqsubseteq_{sim} u_0$. However we will have $t_0 \not\sqsubseteq_{sim} u_0$ because although u_0 can simulate the behaviour of t_0 it accumulates less benefit.

Similar informal reasoning can also be applied to probabilistic systems. Consider the systems in Figure 2. Here we have two kinds of nodes; the first as in Figure 1 representing states of the systems, and the second representing probability distributions. For example the first system, from state s_1, can perform the up action with benefit 2 and a quarter of the time it ends up in a state in which down can be performed with benefit only 1. But for the remaining three-quarters it ends up in a state in which down can be performed for the larger benefit 3. The circular darkened node

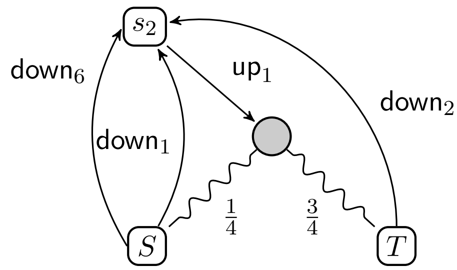

Figure 3: Nondeterministic and probabilistic systems

represents a distribution of states, with its outgoing edges describing the associated probabilities. Again intuitively we can see that s_1 is an improvement on s_0 because it can simulate s_0 and on average accrue slightly more benefits; in our theory we will have $s_0 \sqsubseteq_{sim} s_1$.

The mixture of probabilistic behaviour and internal actions introduces complications. Consider the system t_1 in Figure 2 which after performing an up action probabilistically decides internally whether to perform a down action for benefit 1, or branch back to make another probabilistic choice. However each time it reverts back it accumulates a non-zero benefit via the internal weighted action τ_1, albeit with diminishing probability. Nevertheless it will turn out that for our definition of simulation $s_0 \sqsubseteq_{sim} t_1$ and indeed $s_1 \sqsubseteq_{sim} t_1$.

Systems exhibiting both probabilistic and nondeterministic behaviour require more complicated analysis. Consider the system in Figure 3. After performing the action up it finds itself either in a state in which the action down will accrue the benefit 2, or 25% of the time there will be a nondeterministic choice between it accruing either 1 or 6. In the literature there are numerous mechanisms, such as policies, schedulers, adversaries, etc. [Put94, Seg95, RKNP04] for resolving such choices. Here one can see if this choice systematically leads to the lower benefit 1 then s_2 will not simulate s_0 as it does not accrue sufficient benefits. This is a pessimistic outlook; an optimistic outlook means that the best choices are systematically made. If this is assumed then we will have $s_0 \sqsubseteq_{sim} s_2$; in s_2 one execution of up followed by down will yield on average the benefit $1 + (\frac{3}{4} \cdot 2 + \frac{1}{4} \cdot 6) = 4$.

The main contribution of the paper is a coinductively defined behavioural preorder $\sqsubseteq_{sim}$ between MDPs based on simulations which validate the examples discussed informally above. We confine our attention to the optimistic approach to the resolution of nondeterministic choices, although as future work we hope to investigate the pessimistic approach. We also show that this preorder is compositional in the sense that it is preserved by structural operators for constructing MDPs from components. The main operator is one for composing two MDPs in parallel. In $P \mid Q$ the two MPDs P and Q remain independent, execute in parallel and may communicate by synchronising on complementary actions; these internal synchronisations accrue the combined benefits of the associated complementary actions.

For finitary convergent MDPs, which are finite-state and finitely branching systems without divergence, we also provide two characterisations for the behavioural preorder $\sqsubseteq_{sim}$. The first is in terms of a *quantitative probabilistic logic* $\mathcal{L}$. In addition to the standard logical connectives such as conjunction and maximal fixed point this contains a novel quantitative *possibility* modality $\langle\alpha\rangle_w(\phi_1\,{}_p\oplus\,\phi_2)$, where p is some probability between 0 and 1. Intuitively this is satisfied by an MDP which can accrue at least the benefit w by performing the action α, and subsequently satisfy the probabilistic assertion $\phi_1\,{}_p\oplus\,\phi_2$. It turns out that the simulation preorder is completely determined by the logic $\mathcal{L}$. Further evidence of the compatibility between the logic and the simulation relation is the fact that every system P has a *characteristic formula* $\phi(P)$ in the logic which captures its behaviour; informally system Q can simulate P if and only if it satisfies the characteristic formula $\phi(P)$.

Our second characterisation is in terms of a novel form of testing called *benefits* testing. Intuitively a system P can be tested by running it in parallel with another testing system T, and seeing the possible accrued benefits. In the presence of nondeterminism the execution of the combined system $(T \mid P)$ will result in a non-empty set of benefits, $\mathbf{Benefits}(T \mid P)$. Then systems P and Q can be compared by comparing the associated benefit sets $\mathbf{Benefits}(T \mid P)$ and $\mathbf{Benefits}(T \mid Q)$ where T ranges over some collection of possible tests. We show that the simulation preorder $\sqsubseteq_{sim}$ is also determined in this manner by a suitable collection of tests T.

The rest of this paper is organised as follows. Section 2 is devoted to an exposition of our model, which we call *weighted Markov Decision Processes*, *wMDPs*. These correspond to the diagrams we have been using informally in this introduction. The actions in a wMPD take the form $s \xrightarrow{\alpha}_w \Delta$, where α is the label of the action, w its weight, or benefit, and Δ a probability distribution which determines the next state. Following [Seg95, Seg96, DvGHM09], we make extensive use of the generalisation of this next-step relation to actions from distributions to distributions, $\Delta \xrightarrow{\alpha} \Theta$. Furthermore we are interested in *weak* theories, in which internal activity is not directly observable. So we generalise these actions to weak actions, of the form $s \stackrel{\alpha}{\Longrightarrow} \Delta$ and $\Delta \stackrel{\alpha}{\Longrightarrow} \Theta$ respectively, actions in which occurrences of internal actions, denoted by τ, may occur an arbitrary number of times both before and after α. As have already been pointed out by many authors, [LSV07, DvGHM09], in a probabilistic setting we need to allow a potentially infinite number of internal actions to occur, *in the limit*. We follow the formalisation of this idea suggested in [DvGHM09] based on (weighted) hyper-derivations. We outline properties of these hyper-derivations but their proofs, being quite technical, are relegated to an appendix. One particularly significant property is that the set of weak derivatives from a given state, although in general uncountable, in a finite-state wMDP can be generated as the convex-closure of a finite number of derivatives. This is explained in Section 2.4. The proof is very complex, relying on notions such as static policies and payoffs [DvGHM09]. Consequently, again the details are relegated to an appendix.

Then, still in Section 2, we turn our attention to a subclass of wMDPs, called *bounded wMDPs*. In an arbitrary wMDP if $\Delta \stackrel{\tau}{\Longrightarrow}_w \Theta$ then w may in general be infinite because of an indefinite accumulation of weights during an infinite internal computation. In bounded wMDPs we are guaranteed that such w's will always be a finite real number. Such wMDPs are the main focus of the paper, and their properties are studied in Section 2.5.

Section 3 is devoted to our notion of simulation, called *amortised weighted simulation*, based on ideas from [KAK05]. In the first subsection we give the definition and some examples. The formal simulation preorder $\lhd$ is defined coinductively but in Section 3.2 we show that in bounded wMDPs it can also be defined as the intersection of an infinite sequence of inductively defined relations. This result depends on compactness arguments, which we are able to employ in bounded wMDPs because of the finite generability property alluded to above. Then in Section 3.3 we show that the simulation preorder can be captured by a very simple modal logic, again if we restrict attention to bounded wMDPs. This logic is quantitative in the sense that satisfying formulae depends to some extent on the benefits which a process can accrue. The logical characterisation in turn depends on the approximation result from Section 3.2.

In Section 4 we offer another justification for our simulation based on testing [NH84]. Because of the presence of weights or benefits in wMDPs we are able to use a novel form of (may) testing in which benefits are accrued as tests are applied to processes; then processes can be compared in terms of their ability to accumulate benefits. In section 4.1 this idea, *benefits testing*, is explained in detail and we also show that it is preserved by the simulation preorder. More interesting is the result, for bounded wMDPs, that the preorder is completely determined by these tests. This proof requires a digression, in Section 4.2, into a more standard testing framework. Here we extend the ideas on [Seg96, DvGMZ07] by developing a version of multi-success testing suitable for wMDPs. In a non-trivial theorem we show that in bounded wMDPs both testing preorders, benefit-based and multi-success, coincide. The interest in multi-success testing is that we can mimic the results in [DvGHM09] to show that this form of testing can be captured by the modal logic of the previous section. Since we already know that the modal logic determines the simulation preorder we have therefore also established the soundness and completeness of benefits testing for the simulation preorder.

Section 4 ends with a short discussion of another natural form of testing, *expected benefits testing*, in which the average weight of each path of a computation leading to a success is associated with a test. By means of a simple example we show that the simulation preorder is not sound for this form of testing.

2 Weighted Markov decision processes

2.1 Introduction

There is considerable variation in the literature in the formal definition of a (labelled) Markov decision process [RKNP04, Put94]. For the purpose of this paper we use Definition 2.1.

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

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

are.

For $\Delta \in \mathcal{D}_{sub}(S)$ and f a function with domain S, we write $\text{Exp}_\Delta(f)$, the *expected value* of f over $\Delta \in \mathcal{D}_{sub}(S)$, for $\sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot f(s)$. More generally suppose $f : S^k \to T$. This is lifted to a function $f^\dagger : \mathcal{D}_{sub}(S)^k \to \mathcal{D}_{sub}(T)$ by letting $f^\dagger(\Delta_1, \ldots, \Delta_n)(t) = \sum_{t = f(s_1,\ldots,s_k)} \Delta_1(s_1) \cdot \ldots \cdot \Delta_k(s_k)$. We will often abbreviate the lifted function $f^\dagger$ to simply f.

Definition 2.1 [Weighted Markov decision process] A *weighted Markov decision process* or wMDP is a 4-tuple $\langle S, \mathsf{A}, W, \longrightarrow \rangle$ where S is a set of states, A a set of actions, W a set of weights, and $\longrightarrow \subseteq S \times \mathsf{A} \times W \times \mathcal{D}(S)$. We normally write $s \xrightarrow{\alpha}_w \Delta$ to mean $(s, \alpha, w, \Delta) \in \longrightarrow$. $\Box$

In this paper we set W to be $\mathbb{R}_{\geq 0}$, the set of non-negative real numbers, and we assume A has the structure $\mathsf{Act}_\tau = \mathsf{Act} \cup \{\tau\}$ where each a in Act has an inverse $\overline{a}$ satisfying $\overline{\overline{a}} = a$. We write $s \stackrel{\alpha}{\not\rightarrow}$ if there is no w, Δ such that $s \xrightarrow{\alpha}_w \Delta$. We also use the following terminology. A wMDP is

- *finite-state* if S is a finite set;
- *finitely branching* if for each $s \in S$, the set $\{(\alpha, w, \Delta) \mid s \xrightarrow{\alpha}_w \Delta\}$ is finite;
- *finitary* if it is both finite-state and finitely branching,
- *deterministic* if from every $s \in S$ there is at most one outgoing transition.

In the Introduction we have used a straightforward graphical representation for wMDPs; a state s is represented by a node $\boxed{s}$ while darkened circular nodes are used for distributions, and arrows between nodes and distributions are annotated with their weights. Often a point distribution is represented by the unique state in its support; see the first series of examples with initial states s_0, t_0 and u_0.

The simplest approach to discussing compositionality is, as in [Her02], to introduce a process calculus-like syntax for wMDPs. Our calculus, called CCMDP, is based on CCS:

$$P \quad ::= \quad \alpha_w.(\oplus_{i \in I} p_i \cdot P_i) \;\mid\; P \mid P \;\mid\; P + P \;\mid\; \mathbf{0} \;\mid\; P \backslash a \;\mid\; A \tag{1}$$

The main operator is prefixing, $\alpha_w.(\oplus_{i \in I} p_i \cdot P_i)$. Here α is taken from Act_τ, w from $\mathbb{R}_{\geq 0}$, I is a finite index set and p_i are probabilities satisfying $\sum_{i \in I} p_i = 1$. We also assume a set of definitional constants, ranged over by A, and we assume that each such A has a definition associated with it, a process term P_A. We often write these definitions as

$$A \Leftarrow P_A$$

We will use the auxiliary operator $||$, letting $P \mid\mid Q$ stand for the process $(P \mid Q)\backslash\mathsf{Act}$.

Let $\mathcal{P}$ denote the set of all terms P definable in this language. Intuitively, we view each such term as describing a wMDP. Formally we describe one overarching wMDP where the states are all terms in $\mathcal{P}$ and the weighted actions $P \xrightarrow{\alpha}_w \Delta$ are those which can be derived by the rules in Figure 4; obvious symmetric counterparts to the rules (L-ALT) (L-PAR) are omitted. In rule (L-ACT) we use the obvious notation $\mathcal{D}ist(\{\,(p_i, P_i) \mid i \in I\,\})$ for constructing a distribution from the formal term $\oplus_{i \in I} p_i \cdot P_i$. In rules (L-COMM) and (L-PAR) we take advantage of the fact that

(L-ACT)

$$\alpha_w.(\oplus_{i\in I} p_i \cdot P_i) \xrightarrow{\alpha}_w \mathcal{D}ist(\{\, (p_i, P_i) \mid i \in I \,\})$$

(L-ALT)

$$\frac{P_1 \xrightarrow{\alpha}_w \Delta}{P_1 + P_2 \xrightarrow{\alpha}_w \Delta}$$

(L-PAR)

$$\frac{P_1 \xrightarrow{\alpha}_w \Delta}{P_1 \mid P_2 \xrightarrow{\alpha}_w \Delta \mid P_2}$$

(L-COMM)

$$\frac{P_1 \xrightarrow{a}_{w_1} \Delta_1, \quad P_2 \xrightarrow{\overline{a}}_{w_2} \Delta_2}{P_1 \mid P_2 \xrightarrow{\tau}_w \Delta_1 \mid \Delta_2} \quad w = w_1 + w_2$$

(L-HIDE)

$$\frac{P \xrightarrow{\alpha}_w \Delta}{P\backslash a \xrightarrow{\alpha}_w \Delta\backslash a} \quad \alpha \neq a, \overline{a}$$

(L-DEF)

$$\frac{P_A \xrightarrow{\alpha}_w \Delta}{A \xrightarrow{\alpha}_w \Delta} \quad A \Leftarrow P_A$$

Figure 4: Weighted actions

parallel composition can be viewed as a binary operator over process terms $|\colon \mathcal{P} \times \mathcal{P} \to \mathcal{P}$, and therefore can be lifted to distributions of processes as explained above: $|^\dagger\colon \mathcal{P} \times \mathcal{P} \to \mathcal{P}$. An equivalent definition is given by

$$(\Delta_1 \mid^\dagger \Delta_2)(Q) \;=\; \begin{cases} \Delta_1(P_1) \cdot \Delta_2(P_2) & \text{if } Q = P_1 \mid P_2, \\ 0 & \text{otherwise} \end{cases}$$

and in the sequel we drop the annotation †. The hiding operator is treated in a similar manner.

Note that all of the wMDPs described graphically in the Introduction can be described in CCMDP. In the sequel we will not distinguish between the syntactic term P, its interpretation as a state in the wMDP defined in Figure 4, and the wMDP it induces by considering only those states, that is process terms, accessible from it.

2.2 Lifted relations

In a wMDP actions are only performed by states, in that actions are given by relations from states to distributions. But formal systems or 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.

We first recall some standard terminology. For any subset X of $\mathbb{R}_{\geq} \times \mathcal{D}_{sub}(S)$, with S a set, let $\updownarrow X$, the *convex closure* of X, be the least set satisfying $\langle r, \Theta \rangle \in \updownarrow X$ if and only if $\langle r, \Theta \rangle = \sum_{i \in I} p_i \cdot \langle r_i, \Theta_i \rangle$, where $\langle r_i, \Theta_i \rangle \in X$ and $p_i \in [0,1]$, for some index set I such that $\sum_{i \in I} p_i = 1$. We say a set X is *convex* if $\updownarrow X = X$. Let $\mathcal{R}$ be a relation in $Y \times (\mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S))$. It is

1. *convex* whenever the set $\{\langle r, \Theta \rangle \mid y \, \mathcal{R} \, \langle r, \Theta \rangle\}$ is convex for every y in Y; $\updownarrow \mathcal{R}$ denotes the smallest convex relation containing $\mathcal{R}$
2. *linear* whenever $\Delta_i \, \mathcal{R} \, \langle r_i, \Theta_i \rangle$ for $i \in I$ implies $(\sum_{i \in I} p_i \cdot \Delta_i) \;\; \mathcal{R} \;\; (\sum_{i \in I} p_i \cdot \langle r_i, \Theta_i \rangle)$ for any $p_i \in [0,1]$ $(i \in I)$ with $\sum_{i \in I} p_i \leq 1$
3. *decomposable* whenever $(\sum_{i \in I} p_i \cdot \Delta_i) \, \mathcal{R} \, \langle w, \Theta \rangle$ implies $\langle w, \Theta \rangle = \sum_{i \in I} p_i \cdot \langle w_i, \Theta_i \rangle$ for some weights w_i and subdistributions Θ_i such that $\Delta_i \, \mathcal{R} \, \langle w_i, \Theta_i \rangle$ for $i \in I$.

Note that if $\mathcal{R}$ is linear it is automatically convex.

Definition 2.2 Let $\mathcal{R} \subseteq S \times (\mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S))$ be a relation from states to pairs of weights and subdistributions. Then $\overline{\mathcal{R}} \subseteq \mathcal{D}_{sub}(S) \times (\mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S))$ is the smallest linear relation that satisfies $s \, \mathcal{R} \, \langle r, \Theta \rangle$ implies $\overline{s} \, \overline{\mathcal{R}} \, \langle r, \Theta \rangle$. □

By construction $\overline{\mathcal{R}}$ is both linear and convex. Moreover the lifting operation is monotonic, in that $\mathcal{R}_1 \subseteq \mathcal{R}_2$ implies $\overline{\mathcal{R}_1} \subseteq \overline{\mathcal{R}_2}$. Also, because $s \, (\updownarrow \mathcal{R}) \, \Theta$ implies $\overline{s} \, \overline{\mathcal{R}} \, \Theta$ we have $\overline{\mathcal{R}} = \overline{\updownarrow \mathcal{R}}$. Finally note that if $\mathcal{R}$ itself is 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.

Lemma 2.3 $\Delta \, \overline{\mathcal{R}} \, \langle r, \Theta \rangle$ if and only if

1. $\Delta = \sum_{i \in I} p_i \cdot \overline{s_i}$, where I is an index set and $\sum_{i \in I} p_i \leq 1$,
2. For each $i \in I$ there is a pair $\langle r_i, \Theta_i \rangle$ such that $s_i \, \mathcal{R} \, \langle r_i, \Theta_i \rangle$,
3. $r = \sum_{i \in I} p_i r_i$ and $\Theta = \sum_{i \in I} p_i \cdot \Theta_i$.

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.

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

Definition 2.4 Let $\mathcal{R} \subseteq S \times (\mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S))$ be a relation. Then $f : S \to (\mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S))$ is a *choice function for* $\mathcal{R}$, written $f \in \mathbf{Ch}(\mathcal{R})$, if $s \, \mathcal{R} \, f(s)$ for every $s \in \mathrm{dom}(\mathcal{R})$. □

Note that if f is a choice function of $\mathcal{R}$ then f behaves properly at each state s in the domain of $\mathcal{R}$, but for each state s' outside the domain of $\mathcal{R}$, the value $f(s')$ can be arbitrarily chosen.

Proposition 2.5 Suppose $\mathcal{R} \subseteq S \times (\mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S))$ is a convex relation. Then for any $\Delta \in \mathcal{D}_{sub}(S)$, $\Delta \, \overline{\mathcal{R}} \, \langle w, \Theta \rangle$ if and only if there is some choice function $f \in \mathbf{Ch}(\mathcal{R})$ such that $\langle w, \Theta \rangle = \mathrm{Exp}_\Delta(f)$.

Proof. First suppose $\langle w, \Theta \rangle = \mathrm{Exp}_\Delta(f)$ for some choice function $f \in \mathbf{Ch}(\mathcal{R})$, that is $\langle w, \Theta \rangle = \sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot f(s)$. It now follows from Lemma 2.3 that $\Delta\ \overline{\mathcal{R}}\ \langle w, \Theta \rangle$ since $s\ \mathcal{R}\ f(s)$ for each $s \in \mathrm{dom}(\mathcal{R})$.

Conversely suppose $\Delta\ \overline{\mathcal{R}}\ \langle w, \Theta \rangle$; we have to find a choice function $f \in \mathbf{Ch}(\mathcal{R})$ such that $\langle w, \Theta \rangle = \mathrm{Exp}_\Delta(f)$. Applying Lemma 2.3 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) $\langle w, \Theta \rangle = \sum_{i \in I} p_i \cdot \langle w_i, \Theta_i \rangle$ for some $\langle w_i, \Theta_i \rangle$ satisfying $s_i\ \mathcal{R}\ \langle w_i, \Theta_i \rangle$.

Now define the function $f : S \to (\mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S))$ as follows:

- if $s \in \lceil \Delta \rceil$ then $f(s) = \sum_{\{\, i \in I \,|\, s_i = s\,\}} (\frac{p_i}{\Delta(s)}) \cdot \langle w_i, \Theta_i \rangle$;
- if $s \in \mathrm{dom}(\mathcal{R}) \backslash \lceil \Delta \rceil$ then $f(s) = \langle w', \Theta' \rangle$ for any $\langle w', \Theta' \rangle$ with $s\ \mathcal{R}\ \langle w', \Theta' \rangle$;
- otherwise, $f(s) = \langle 0, \varepsilon \rangle$, where ε is the empty subdistribution.

Note that if $s \in \lceil \Delta \rceil$ then $\Delta(s) = \sum_{\{\, i \in I \,|\, s_i = s\,\}} p_i$ and therefore by convexity $s\ \mathcal{R}\ f(s)$; so f is a choice function for $\mathcal{R}$ as $s\ \mathcal{R}\ f(s)$ for each $s \in \mathrm{dom}(\mathcal{R})$. Moreover, a simple calculation shows that $\mathrm{Exp}_\Delta(f) = \sum_{i \in I} p_i \cdot \langle w_i, \Theta_i \rangle$, which by (ii) above is $\langle w, \Theta \rangle$. □

By Definition 2.2, a lifted relation is linear and convex; we now show that it is also decomposable.

Proposition 2.6 Let $\mathcal{R} \subseteq \mathcal{D}_{sub}(S) \times (\mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S))$ be a relation. Then $\overline{\mathcal{R}}$ is decomposable.

Proof. Let $\Delta\ \overline{\mathcal{R}}\ \langle w, \Theta \rangle$ where $\Delta = \sum_{i \in I} p_i \cdot \Delta_i$. By Proposition 2.5, using that $\overline{\mathcal{R}} = \overline{\updownarrow \mathcal{R}}$, there is a choice function $f \in \mathbf{Ch}(\updownarrow \mathcal{R})$ such that $\langle w, \Theta \rangle = \mathrm{Exp}_\Delta(f)$. Take $\langle w_i, \Theta_i \rangle := \mathrm{Exp}_{\Delta_i}(f)$ for $i \in I$. Using that $\lceil \Delta_i \rceil \subseteq \lceil \Delta \rceil$, Proposition 2.5 yields $\Delta_i\ \overline{\mathcal{R}}\ \langle w_i, \Theta_i \rangle$ for $i \in I$. Finally,

$$\sum_{i \in I} p_i \cdot \langle w_i, \Theta_i \rangle = \sum_{i \in I} p_i \cdot \sum_{s \in \lceil \Delta_i \rceil} \Delta_i(s) \cdot f(s) = \sum_{s \in \lceil \Delta \rceil} \sum_{i \in I} p_i \cdot \Delta_i(s) \cdot f(s) = \sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot f(s) = \mathrm{Exp}_\Delta(f) = \langle w, \Theta \rangle.$$
□

The converse to the above is not true in general: from $\Delta\ \overline{\mathcal{R}}\ (\sum_{i \in I} p_i \cdot \langle w_i, \Theta_i \rangle)$ it does not follow that Δ can correspondingly be decomposed. For example, we have

$$\overline{a_0.(b_0.\,\mathbf{0}\ {}_{\frac{1}{2}}\!\oplus c_0.\,\mathbf{0})} \xrightarrow{a}_0 \frac{1}{2} \cdot \overline{b_0.\,\mathbf{0}} + \frac{1}{2} \cdot \overline{c_0.\,\mathbf{0}},$$

yet $\overline{a.(b_0.\,\mathbf{0}\ {}_{\frac{1}{2}}\!\oplus c_0.\,\mathbf{0})}$ cannot be written as $\frac{1}{2} \cdot \Delta_1 + \frac{1}{2} \cdot \Delta_2$ such that $\Delta_1 \xrightarrow{a}_0 \overline{b_0.\,\mathbf{0}}$ and $\Delta_2 \xrightarrow{a}_0 \overline{c_0.\,\mathbf{0}}$.

In fact a simplified form of Proposition 2.6 holds for un-lifted relations, provided they are convex:

Corollary 2.7 If $(\sum_{i \in I} p_i \cdot \overline{s_i})\ \overline{\mathcal{R}}\ \langle w, \Theta \rangle$ and $\mathcal{R}$ is convex, then $\langle w, \Theta \rangle = \sum_{i \in I} p_i \cdot \langle w_i, \Theta_i \rangle$ for weights w_i and subdistributions Θ_i with $s_i\ \mathcal{R}\ \langle w_i, \Theta_i \rangle$ for $i \in I$.

Proof. Take Δ_i to be $\overline{s_i}$ in Proposition 2.6, whence $\langle w, \Theta \rangle = \sum_{i \in I} p_i \cdot \langle w_i, \Theta_i \rangle$ for some weights w_i and subdistributions Θ_i such that $\overline{s_i}\ \overline{\mathcal{R}}\ \langle w_i, \Theta_i \rangle$ for $i \in I$. Because $\mathcal{R}$ is convex, we then have $s_i\ \mathcal{R}\ \langle w_i, \Theta_i \rangle$. □

2.3 Hyper-derivations

As we have seen in the Introduction, when reasoning informally that t_1 can simulate s_0, the limiting behaviour of internal computations must be taken into account. We formalise this by extending the approach originally given in [DvGHM09]. By employing the lifting operation defined in Section 2.2, we now formally define weak actions performed by subdistributions.

Definition 2.8 [Hyper-derivations] A hyper-derivation consists of a collection of subdistributions $\Delta, \Delta_k^{\rightarrow}, \Delta_k^{\times}$, for $k \geq 0$, with the following properties:

$$\begin{array}{rcl} \Delta & = & \Delta_0^{\rightarrow} + \Delta_0^{\times} \\ \Delta_0^{\rightarrow} & \xrightarrow{\tau}_{w_0} & \Delta_1^{\rightarrow} + \Delta_1^{\times} \\ & \vdots & \\ \Delta_k^{\rightarrow} & \xrightarrow{\tau}_{w_k} & \Delta_{k+1}^{\rightarrow} + \Delta_{k+1}^{\times} \\ & \vdots & \\ \Delta' & = & \sum_{k=0}^{\infty} \Delta_k^{\times} \end{array} \tag{2}$$

Then we call $\Delta' = \sum_{k=0}^{\infty} \Delta_k^{\times}$ a *hyper-derivative* of Δ, and write $\Delta \stackrel{\tau}{\Longrightarrow}_w \Delta'$, where $w = \sum_{k=0}^{\infty} w_i$, to mean that Δ can make a *(weak) hyper-move* to its derivative Δ' with weight w. Note that in general $w \in \mathbb{R}_{\geq 0} \cup \{\infty\}$; that is there is no guarantee that the sum $\sum_{k=0}^{\infty} w_i$ has a finite limit. □

One question to answer is when can we ensure that this sum does indeed have a limit. This will be studied in Section 2.5.

Example 2.9 Consider the wMDP with initial state t_1 discussed in the Introduction. Then we have the following hyper-derivation:

$$\begin{array}{rcl} \overline{U} & = & \overline{U} + \varepsilon \\ \overline{U} & \xrightarrow{\tau}_0 & \frac{3}{4} \cdot \overline{R} + \frac{1}{4} \cdot \overline{D} \\ \frac{3}{4} \cdot \overline{R} & \xrightarrow{\tau}_{\frac{3}{4}} & \frac{3}{4} \cdot \overline{U} + \varepsilon \\ \frac{3}{4} \cdot \overline{U} & \xrightarrow{\tau}_0 & (\frac{3}{4})^2 \cdot \overline{R} + (\frac{3}{4})\frac{1}{4} \cdot \overline{D} \\ (\frac{3}{4})^2 \cdot \overline{R} & \xrightarrow{\tau}_{(\frac{3}{4})^2} & (\frac{3}{4})^2 \cdot \overline{U} + \varepsilon \\ & \vdots & \\ (\frac{3}{4})^k \cdot \overline{U} & \xrightarrow{\tau}_0 & (\frac{3}{4})^{(k+1)} \cdot \overline{R} + (\frac{3}{4})^k \frac{1}{4} \cdot \overline{D} \\ (\frac{3}{4})^{(k+1)} \cdot \overline{R} & \xrightarrow{\tau}_{(\frac{3}{4})^{(k+1)}} & (\frac{3}{4})^{(k+1)} \cdot \overline{U} + \varepsilon \\ & \vdots & \end{array}$$

That is, $\overline{U} \stackrel{\tau}{\Longrightarrow}_w \sum_{k=0}^{\infty} (\frac{3}{4})^k(\frac{1}{4} \cdot \overline{D})$ where $w = \sum_{k=1}^{\infty}(\frac{3}{4})^k$. However this weight evaluates to 3, while the sum of the subdistributions is the full point distribution $\overline{D}$. In other words $\overline{U} \stackrel{\tau}{\Longrightarrow}_3 \overline{D}$. $\square$

Definition 2.10 [Weak actions] In a wMDP $\langle S, \mathsf{Act}_\tau, \mathbb{R}_{\geq 0}, \longrightarrow \rangle$ for $\Delta, \Theta \in \mathcal{D}_{sub}(S)$ we write $\Delta \stackrel{a}{\Longrightarrow}_w \Delta$ whenever $\Delta \stackrel{\tau}{\Longrightarrow}_{w_1} \Delta' \stackrel{a}{\longrightarrow}_{w_2} \Theta' \stackrel{\tau}{\Longrightarrow}_{w_3} \Theta$ and $w = w_1 + w_2 + w_3$. $\square$

We complete this subsection by enumerating some elementary properties of hyper-derivations; their proofs are relegated to Appendix A.

Proposition 2.11

1. If $\Delta \stackrel{\tau}{\Longrightarrow}_v \Theta$ then $|\Delta| \geq |\Theta|$.
2. If $\Delta \stackrel{\tau}{\Longrightarrow}_v \Theta$ and $p \in \mathbb{R}_{\geq 0}$ such that $|p \cdot \Delta| \leq 1$, then $p \cdot \Delta \stackrel{\tau}{\Longrightarrow}_{pv} p \cdot \Theta$.
3. (Binary decomposition) If $\Gamma + \Lambda \stackrel{\tau}{\Longrightarrow}_v \Pi$ then $\Pi = \Pi^\Gamma + \Pi^\Lambda$ with $\Gamma \stackrel{\tau}{\Longrightarrow}_{v^\Gamma} \Pi^\Gamma$, $\Lambda \stackrel{\tau}{\Longrightarrow}_{v^\Lambda} \Pi^\Lambda$, and $v = v^\Gamma + v^\Lambda$.
4. (Linearity) Let $p_i \in [0,1]$ for $i \in I$ where $\sum_{i \in I} p_i \leq 1$. Then $\Delta_i \stackrel{\tau}{\Longrightarrow}_{w_i} \Theta_i$ for all $i \in I$ implies $\sum_{i \in I} p_i \cdot \Delta_i \stackrel{\tau}{\Longrightarrow}_{(\sum_{i \in I} p_i \cdot w_i)} \sum_{i \in I} p_i \cdot \Theta_i$.
5. (Decomposability) suppose $\sum_{i \in I} p_i \cdot \Delta_i \stackrel{\tau}{\Longrightarrow}_w \Theta$, where $p_i \in [0,1]$ and $\sum_{i \in I} p_i \leq 1$. Then $w = \sum_{i \in I} p_i \cdot w_i$ and $\Theta = \sum_{i \in I} p_i \cdot \Theta_i$ for weights w_i and subdistributions Θ_i such that $\Delta_i \stackrel{\tau}{\Longrightarrow}_{w_i} \Theta_i$ for all $i \in I$.

Proof. See Appendix A. $\square$

With these results the relation $\stackrel{\tau}{\Longrightarrow} \subseteq \mathcal{D}_{sub}(S) \times (\mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S))$ can be obtained as the lifting of a relation $\stackrel{\tau}{\Longrightarrow}_S$ from S to $\mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S)$, which is defined by writing $s \stackrel{\tau}{\Longrightarrow}_S \langle w, \Theta \rangle$ just when $\overline{s} \stackrel{\tau}{\Longrightarrow}_w \Theta$.

Corollary 2.12 $\overline{(\stackrel{\tau}{\Longrightarrow}_S)} = (\stackrel{\tau}{\Longrightarrow})$.

Proof. That $\Delta \; \overline{(\stackrel{\tau}{\Longrightarrow}_S)} \; \langle w, \Theta \rangle$ implies $\Delta \stackrel{\tau}{\Longrightarrow}_w \Theta$ is a simple application of Part 4 followed by Part 3 of Proposition 2.11. For the other direction, suppose $\Delta \stackrel{\tau}{\Longrightarrow}_w \Theta$. Given that $\Delta = \sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot \overline{s}$, Part 5 of the same proposition enables us to decompose Θ into $\sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot \Theta_s$ and w into $\sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot w_s$, where $\overline{s} \stackrel{\tau}{\Longrightarrow}_{w_s} \Theta_s$ for each s in $\lceil \Delta \rceil$. But the latter actually means that $s \stackrel{\tau}{\Longrightarrow}_S \langle w_s, \Theta_s \rangle$, and so by definition this implies $\Delta \; \overline{(\stackrel{\tau}{\Longrightarrow}_S)} \; \langle w, \Theta \rangle$. $\square$

Corollary 2.12 implies that the hyper-derivation relation $\stackrel{\tau}{\Longrightarrow}$ is convex. It is trivial to check that $\stackrel{\tau}{\Longrightarrow}$ is also reflexive because $\Delta \stackrel{\tau}{\Longrightarrow}_0 \Delta$ for any $\Delta \in \mathcal{D}_{sub}(S)$. But transitivity is less obvious.

Theorem 2.13 [Transitivity of $\stackrel{\tau}{\Longrightarrow}$] If $\Delta \stackrel{\tau}{\Longrightarrow}_u \Theta$ and $\Theta \stackrel{\tau}{\Longrightarrow}_v \Lambda$ then $\Delta \stackrel{\tau}{\Longrightarrow}_{u+v} \Lambda$.

Proof. See Appendix A. $\square$

2.4 Finite generability

We aim to establish that, for any subdistribution Δ in a finitary wMDP where hyper-derivations only yield finite weights, the set $\{\langle w, \Delta' \rangle \mid \Delta \stackrel{\tau}{\Longrightarrow}_w \Delta'\}$ can be generated by taking the convex closure of a finite set of pairs $\{\langle w_1, \Delta_1 \rangle, ..., \langle w_n, \Delta_n \rangle\}$. The proof is non-trivial and requires a digression into the world of payoff functions and policies.

Let us fix a finite-state space $S = \{s_1, ..., s_n\}$ with $n \geq 1$ and define an extended state space $S \cup \{s_0\}$. This allows us to deal with vectors and in particular to use vector arithmetic. For example, a subdistribution $\Delta \in \mathcal{D}_{sub}(S)$ can be viewed as the n-dimensional vector $\langle \Delta(s_1), ..., \Delta(s_n) \rangle$, and a pair $\langle w, \Delta \rangle$ consisted of weight w and subdistribution Δ may be viewed as the $(n+1)$-dimensional vector $\langle w, \Delta(s_1), ..., \Delta(s_n) \rangle$ in some contexts.

Definition 2.14 [Weight functions] A *weight function* is a function $\mathbf{w} : S \cup \{s_0\} \to [-1, 1]$ from the extended state space into the real interval [-1,1]. □

This notion of *weight function* is not to be confused with the weights associated with actions in a wMDP; instead they will be applied to the results of executing hyper-derivations. We often consider a weight function as the $(n+1)$-dimensional vector $\langle \mathbf{w}(s_0), ..., \mathbf{w}(s_n) \rangle$. Therefore the the result of applying the weight function $\mathbf{w}$ to $\langle w, \Delta \rangle$ is given by the inner product of the two vectors $\mathbf{w} \centerdot \langle w, \Delta \rangle$.

Definition 2.15 [Payoff functions] Given a weight function $\mathbf{w}$, the *payoff function* $\mathbb{P}^{\mathbf{w}}_{\max} : S \to \mathbb{R}$ is defined by

$$\mathbb{P}^{\mathbf{w}}_{\max}(s) \;=\; \sup\{\mathbf{w} \centerdot \langle w, \Delta' \rangle \mid \overline{s} \stackrel{\tau}{\Longrightarrow}_w \Delta'\}$$

and we will generalise it to be of type $\mathcal{D}_{sub}(S) \to \mathbb{R}$ by letting $\mathbb{P}^{\mathbf{w}}_{\max}(\Delta) = \sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot \mathbb{P}^{\mathbf{w}}_{\max}(s)$. □

A priori these payoff functions for a given state s are determined by its set of hyper-derivatives. However they can also be calculated by using derivative policies, decision mechanisms for guiding a computation through a wMDP.

Definition 2.16 A *static (derivative) policy* (SP) for a wMDP is a partial function $\mathsf{pp} : S \rightharpoonup \mathbb{R}_{\geq 0} \times \mathcal{D}(S)$ such that if $\mathsf{pp}(s) = \langle w, \Delta \rangle$ then $s \stackrel{\tau}{\longrightarrow}_w \Delta$.

If pp is undefined at s, we write $\mathsf{pp}(s)\uparrow$. Otherwise, we write $\mathsf{pp}(s)\downarrow$. □

A derivative policy pp, as its name suggests, can be used to guide the derivation of a weak derivative. Suppose $\overline{s} \stackrel{\tau}{\Longrightarrow}_w \Delta$, using a derivation as given in Definition 2.8; for convenience we abbreviate $(\Delta_k^{\rightarrow} + \Delta_k^{\times})$ to Δ^k. Then we write $\overline{s} \stackrel{\tau}{\Longrightarrow}_{\mathsf{pp},w} \Delta$ whenever $\Delta_0 = \overline{s}$ and, for all $k \geq 0$,

(a) $\langle w_{k+1}, \Delta_{k+1} \rangle = \sum\{\Delta_k(s) \cdot \mathsf{pp}(s) \mid s \in \lceil \Delta_k \rceil \text{ and } \mathsf{pp}(s)\downarrow\}$

(b) $\Delta_k^{\times}(s) \;=\; \left\{ \begin{array}{ll} 0 & \text{if } \mathsf{pp}(s)\downarrow \\ \Delta_k(s) & \text{otherwise} \end{array} \right\}$

We refer to $\overline{s} \stackrel{\tau}{\Longrightarrow}_{\mathsf{pp},w} \Delta$ as a hyper-SP-derivation from s. Intuitively the conditions mean that the derivation of Δ from s, and the accumulation of weights, is guided at each stage by the policy pp; the division of Δ_k into $\Delta_k^{\rightarrow}$, the subdistribution which will continue marching, and $\Delta_k^{\times}$, the subdistribution which will stop, is determined by the domain of the derivative policy pp.

Lemma 2.17 Let pp be derivative policy in a pLTS. Then

(1) If $\overline{s} \stackrel{\tau}{\Longrightarrow}_{\mathsf{pp},v} \Delta$ and $\overline{s} \stackrel{\tau}{\Longrightarrow}_{\mathsf{pp},w} \Theta$ then $v = w$ and $\Delta = \Theta$.

(2) For every state s there exists some w, Δ such that $\overline{s} \stackrel{\tau}{\Longrightarrow}_{\mathsf{pp},w} \Delta$.

Proof. To prove part (1) consider the derivation of $\overline{s} \stackrel{\tau}{\Longrightarrow}_v \Delta$ and $\overline{s} \stackrel{\tau}{\Longrightarrow}_w \Theta$ as in Definition 2.8, via the subdistributions $\Delta_k,\ \Delta_k^{\rightarrow},\ \Delta_k^{\times}$ and $\Theta_k,\ \Theta_k^{\rightarrow},\ \Theta_k^{\times}$ respectively, and the weights v_k, w_k. Because both derivations are guided by the same derivative policy pp it is easy to show by induction on k that

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

from which $\Delta = \Theta$ and $v = w$ follow immediately.

To prove (2) generate subdistributions $\Delta_k,\ \Delta_k^{\rightarrow},\ \Delta_k^{\times}$ and weights w_k for each $k \geq 0$ satisfying the constraints of Definition 2.8 by applying (a) and (b) above to pp. The result will then follow by letting Δ be $\sum_{k\geq 0} \Delta_k^{\times}$ and w to be $\sum_{k\geq 0} w_k$. □

The net effect of this lemma is that a derivative policy pp determines a *total* function over states. Moreover a policy can used as an alternative to the method used in Definition 2.15 to calculate weighted payoffs.

Definition 2.18 [Policy-following payoffs] Given a weight function $\mathbf{w}$, and static policy pp, the *policy-following payoff function* $\mathbb{P}^{\mathsf{pp},\mathbf{w}} : S \to \mathbb{R}^{\infty}$ is defined by

$$\mathbb{P}^{\mathsf{pp},\mathbf{w}}(s) \;=\; \mathbf{w} \centerdot \langle\, w, \Delta' \,\rangle$$

where w, Δ are determined uniquely by $\overline{s} \stackrel{\tau}{\Longrightarrow}_{\mathsf{pp},w} \Delta'$. □

It should be clear that the use of derivative policies limits considerably the scope for calculating weighted payoffs. Each particular policy can only derive one weak derivative, and moreover in finitary pLTS there are only a finite number of derivative policies. Nevertheless this limitation is more apparent than real.

Theorem 2.19 In a finitary wMDP, for any weight function $\mathbf{w}$ there exists a static policy pp such that $\mathbb{P}^{\mathbf{w}}_{\max} = \mathbb{P}^{\mathsf{pp},\mathbf{w}}$.

The proof of this theorem is non-trivial, requiring the use of discounted policies and payoffs. It is relegated to Appendix B.

Theorem 2.20 [Finite generability] Let $\mathsf{pp}_1, ..., \mathsf{pp}_n$ $(n \geq 1)$ be all the static policies in a finitary wMDP. Suppose $\Delta \stackrel{\tau}{\Longrightarrow}_{\mathsf{pp}_i,w_i} \Delta_i'$ and $w_i < \infty$ for all $1 \leq i \leq n$. If $\Delta \stackrel{\tau}{\Longrightarrow}_w \Delta'$ then there are probabilities p_i for all $1 \leq i \leq n$ with $\sum_{i=1}^n p_i = 1$ such that $\langle\, w, \Delta' \,\rangle = \sum_{i=1}^n p_i \cdot \langle\, w_i, \Delta_i' \,\rangle$.

Proof. Let X be the convex closure of the finite set $\{\langle\, w_i, \Delta_i' \,\rangle \mid 1 \leq i \leq n\}$. It suffices to show that whenever $\Delta \stackrel{\tau}{\Longrightarrow}_w \Delta'$ then $\langle\, w, \Delta' \,\rangle$ belongs to X. Suppose for a contradiction that $\langle\, w, \Delta' \,\rangle$ is not in X. Since X is convex, Cauchy closed and bounded, by the Hyperplane separation theorem, Theorem 1.2.4 in [Mat02], $\langle\, w, \Delta' \,\rangle$ can be separated from X by a hyperplane H whose normal can be scaled into $[-1,1]$ because we are in finitely many dimensions. The scaled normal induces a weight function $\mathbf{w}_H$ such that, for some $c \in \mathbb{R}$, we have $\mathbf{w}_H \cdot \langle\, w, \Delta' \,\rangle > c$ but $\mathbf{w}_H \cdot x < c$ for all $x \in X$. Then we have $\mathbb{P}^{\mathbf{w}_H}_{\max}(\Delta) > c$ but $\mathbb{P}^{\mathsf{pp}_i, \mathbf{w}_H}(\Delta) < c$ for all $0 \leq i \leq n$, contradicting Theorem 2.19. Therefore, $\langle\, w, \Delta' \,\rangle$ must be in X, and is a convex combination of $\{\langle\, w_i, \Delta'_i \,\rangle \mid 1 \leq i \leq n\}$. □

Remark 2.21 It is important that in Theorem 2.20 the weight given by every static policy is finite. Consider a wMDP consisted of two states s_1, s_2 and two transitions $s_1 \xrightarrow{\tau}_1 \overline{s_2}$, $s_1 \xrightarrow{\tau}_1 \overline{s_1}$. It can only have two static policies. The first one, say pp_1, is given by $\mathsf{pp}_1(s_1) = \langle\, 1, \overline{s_2} \,\rangle$ and $\mathsf{pp}_1(s_2)\uparrow$. The second one, say pp_2 is given by $\mathsf{pp}_2(s_1) = \langle\, 1, \overline{s_1} \,\rangle$ and $\mathsf{pp}_2(s_2)\uparrow$. They determine two hyper-derivations from s_1, namely $\overline{s_1} \overset{\tau}{\Longrightarrow}_{\mathsf{pp}_1,1} \overline{s_2}$ and $\overline{s_1} \overset{\tau}{\Longrightarrow}_{\mathsf{pp}_2,\infty} \varepsilon$. Now consider the hyper-derivation $\overline{s_1} \overset{\tau}{\Longrightarrow}_2 \overline{s_1}$. Clearly, $\langle\, 2, \overline{s_1} \,\rangle$ is not a convex combination of $\langle\, 1, \overline{s_2} \,\rangle$ and $\langle\, \infty, \varepsilon \,\rangle$.

Here the culprit is pp_2 which gives an infinite weight. In fact, the convex closure of the set $\{\langle\, 1, \overline{s_2} \,\rangle,\ \langle\, \infty, \varepsilon \,\rangle\}$ is unbounded, thus the Hyperplane separation theorem does not apply, and as a matter of fact it is impossible to separate $\langle\, 2, \overline{s_1} \,\rangle$ from that set.

2.5 Bounded wMDPs

Definition 2.22 A *bounded wMDP* is a finitary wMDP such that if Δ is a subdistribution over it and

$$\Delta \xrightarrow{\tau}_{w_1} \Delta_1 \xrightarrow{\tau}_{w_2} \Delta_2 \xrightarrow{\tau}_{w_3} \cdots$$

then $\sum_{i=1}^{\infty} w_i < \infty$.

In other words, a bounded wMDP is a finitary wMDP that might diverge, but with bounded weights. □

In this section we give an alternative characterisation of boundedness (Theorem 2.27), followed by a useful criteria which ensures boundedness (Theorem 2.29).

Definition 2.23 A wMDP is *convergent* if no state is wholly divergent, i.e. $\overline{s} \overset{\tau}{\Longrightarrow}_w \varepsilon$ for no state $s \in S$ and weight w. □

We will show that this condition is sufficient to ensure that a finitary wMDP is bounded.

Lemma 2.24 Let Δ be a subdistribution in a *finite-state, convergent* and *deterministic* wMDP. If $\Delta \overset{\tau}{\Longrightarrow}_w \Delta'$ then

1. w is a finite real number and
2. $|\Delta| = |\Delta'|$.

Proof. Since the wMDP is convergent, then $\overline{s} \overset{\tau}{\Longrightarrow}_w \varepsilon$ for no state $s \in S$ and weight w. In other words, each τ sequence from a state s is finite and ends with a distribution Δ_{n_s} which cannot enable a τ transition.

$$\overline{s} \xrightarrow{\tau}_{w_1} \Delta_1 \xrightarrow{\tau}_{w_2} \Delta_2 \xrightarrow{\tau}_{w_3} \cdots \xrightarrow{\tau}_{w_{n_s}} \Delta_{n_s} \overset{\tau}{\not\rightarrow}$$

In a deterministic wMDP, each state has at most one outgoing transition. So from each s there is a unique τ sequence with length $n_s \geq 0$. Let p_s be $\Delta_{n_s}(s')$ where s' is any state in the support of Δ_{n_s}. We set

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

Note that since we are considering a finite-state wMDP both n and p are well defined. Now let $\Delta \stackrel{\tau}{\Longrightarrow}_w \Delta'$ be any hyper-derivation constructed by a collection of $\Delta_k^{\rightarrow}, \Delta_k^{\times}, w_k$ such that

$$\begin{array}{rcl} \Delta & = & \Delta_0^{\rightarrow} + \Delta_0^{\times} \\ \Delta_0^{\rightarrow} & \stackrel{\tau}{\longrightarrow}_{w_0} & \Delta_1^{\rightarrow} + \Delta_1^{\times} \\ & \vdots & \\ \Delta_k^{\rightarrow} & \stackrel{\tau}{\longrightarrow}_{w_k} & \Delta_{k+1}^{\rightarrow} + \Delta_{k+1}^{\times} \\ & \vdots & \end{array}$$

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

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

Since the wMDP is finite-state and deterministic, it is finitely branching. Therefore, there exists a maximum weight $w_{\max}$ such that whenever $s \stackrel{\tau}{\longrightarrow}_v \Theta$ then $v \leq w_{\max}$. It follows that

$$w \;=\; \sum_{i=0}^{\infty} w_i \;\leq\; \sum_{i=0}^{\infty} |\Delta_i^{\rightarrow}| w_{\max} \;\leq\; \frac{|\Delta_0^{\rightarrow}| n w_{\max}}{p}$$

which means that the weight w is finite.

From above, $\sum_{j=0}^{\infty} |\Delta_j^{\rightarrow}|$ is bounded (by $|\Delta_0^{\rightarrow}|\frac{n}{p}$). It follows that $\lim_{k\to\infty} \Delta_k^{\rightarrow} = 0$, which in turn means that $|\Delta'| = |\Delta|$. □

Example 2.25 In Lemma 2.24 it is important to require the wMDP to be convergent. In a finite-state deterministic but divergent system, a hyper-derivation $\Delta \stackrel{\tau}{\Longrightarrow}_w \Delta'$ may yield an infinite weight w, even in the case that both Δ and Δ' are full distributions. For example, consider a system consisting of one state s together with a self τ loop $s \stackrel{\tau}{\longrightarrow}_1 \overline{s}$. We construct a hyper-derivation as follows.

$$\begin{array}{rcl} \overline{s} & = & \frac{1}{2}\overline{s} + \frac{1}{2}\overline{s} \\ \frac{1}{2}\overline{s} & \stackrel{\tau}{\longrightarrow}_{\frac{1}{2}} & \frac{1}{3}\overline{s} + (\frac{1}{2} - \frac{1}{3})\overline{s} \\ \frac{1}{3}\overline{s} & \stackrel{\tau}{\longrightarrow}_{\frac{1}{3}} & \frac{1}{4}\overline{s} + (\frac{1}{3} - \frac{1}{4})\overline{s} \\ & \vdots & \\ & \Delta' & = \overline{s} \end{array}$$

So $\overline{s}$ makes a hyper-derivation to itself, but with weight $\sum_{k=2}^{\infty} \frac{1}{k} = \infty$. □

Lemma 2.26 [Distillation of divergence - static case] In a finite-state wMDP if there is a hyper-SP-derivation $\Delta \overset{\tau}{\Longrightarrow}_{\mathsf{pp},w} \Delta'$, there exists subdistribution Δ'_ε such that $\Delta \overset{\tau}{\Longrightarrow}_{w_1} (\Delta' + \Delta'_\varepsilon)$, $|\Delta| = |\Delta' + \Delta'_\varepsilon|$, $\Delta'_\varepsilon \overset{\tau}{\Longrightarrow}_{w_2} \varepsilon$, w_1 is finite and $w_1 + w_2 = w$.

Proof. (*Schema*) We modify pp so as to obtain a static policy pp' by setting $\mathsf{pp}'(s) = \mathsf{pp}(s)$ except when $\overline{s} \overset{\tau}{\Longrightarrow}_{\mathsf{pp},w_s} \varepsilon$ for some weight w_s, in which case we set $\mathsf{pp}'(s)\uparrow$. Intuitively, for any state s which can potentially leads to total divergence under policy pp, the new policy pp' requires it to stop marching at the very beginning. The new policy determines a unique hyper-SP-derivation $\Delta \overset{\tau}{\Longrightarrow}_{\mathsf{pp}',w_1} \Delta''$ for some w_1 and Δ'', and induces a sub-wMDP from the wMDP induced by pp. Note that the sub-wMDP is deterministic, and convergent too because all divergent states in the original wMDP do not contribute any τ move in the sub-wMDP. By Lemma 2.24, we know that w_1 is finite and $|\Delta| = |\Delta''|$. We split Δ'' up into $\Delta''_1 + \Delta''_\varepsilon$ so that each state in $\lceil \Delta''_\varepsilon \rceil$ is wholly divergent under policy pp and Δ''_1 is supported by all other states. From Δ'_ε the policy pp determines the hyper-SP-derivation $\Delta'_\varepsilon \overset{\tau}{\Longrightarrow}_{\mathsf{pp},w_2} \varepsilon$ for some w_2. Combining the two hyper-SP-derivations we have $\overline{s} \overset{\tau}{\Longrightarrow}_{\mathsf{pp}',w_1} \Delta''_1 + \Delta''_\varepsilon \overset{\tau}{\Longrightarrow}_{\mathsf{pp},w_2} \Delta''_1$.

In the above analysis, we divide the original hyper-SP-derivation into two stages by letting the subdistribution Δ''_ε pause in the first stage and then resume marching in the second stage. Note that the two-staged hyper-SP-derivation consists of the same τ transitions from the original hyper-SP-derivation, which means that the overall weight and the final subdistribution remain the same as before, thus we have $w_1 + w_2 = w$ and $\Delta''_1 = \Delta'$. □

Theorem 2.27 A finitary wMDP is bounded if and only if for any subdistribution Δ, $\Delta \overset{\tau}{\Longrightarrow}_w \Delta'$ implies w is a finite real number.

Proof. ($\Leftarrow$) First consider a finitary wMDP where we are assured that for any hyper-derivation from any distribution $\Delta \overset{\tau}{\Longrightarrow}_w \Delta'$, the weight w is finite. It is straightforward to see that the wMDP is bounded: if $\Delta \overset{\tau}{\Longrightarrow}_w \varepsilon$, then by the hypothesis we know that w is finite.

($\Rightarrow$) In a finitary wMDP, there are only finitely many static policies, say pp_i for $i \in I$ where I is a finite index set. For each pp_i we have the unique hyper-SP-derivation $\Delta \overset{\tau}{\Longrightarrow}_{\mathsf{pp}_i,w_i} \Delta'_i$. By Lemma 2.26 there exists subdistribution $\Delta'_{i\,\varepsilon}$ such that $\Delta \overset{\tau}{\Longrightarrow}_{w_{i1}} (\Delta'_i + \Delta'_{i\,\varepsilon})$, $|\Delta| = |\Delta'_i + \Delta'_{i\,\varepsilon}|$, $\Delta'_{i\,\varepsilon} \overset{\tau}{\Longrightarrow}_{w_{i2}} \varepsilon$, w_{i1} is finite and $w_{i1} + w_{i2} = w_i$. If the wMDP is bounded, then w_{i2} is finite. It follows that w_i is also finite as it is the sum of two finite real numbers. Now we can apply Theorem 2.20 to obtain that whenever $\Delta \overset{\tau}{\Longrightarrow}_w \Delta'$ then w is a convex combination of $\{w_i \mid i \in I\}$ which must be finite. □

This theorem enables us to generalise Lemma 2.26 to arbitrary hyper-derivations, provided we restrict attention to bounded wMDPs.

Corollary 2.28 [Distillation of divergence - general case] In a bounded wMDP if $\Delta \overset{\tau}{\Longrightarrow}_w \Delta'$ then there exists subdistribution Δ'_ε such that $\Delta \overset{\tau}{\Longrightarrow}_{w_1} (\Delta' + \Delta'_\varepsilon)$, $|\Delta| = |\Delta' + \Delta'_\varepsilon|$, $\Delta'_\varepsilon \overset{\tau}{\Longrightarrow}_{w_2} \varepsilon$ and $w_1 + w_2 = w$.

Proof. Let $\{\mathsf{pp}_i \mid i \in I\}$ (I is a finite index set) be all the static policies in the bounded wMDP. Each policy determines a hyper-SP-derivation $\Delta \overset{\tau}{\Longrightarrow}_{\mathsf{pp}_i,w_i} \Delta'_i$. By Theorem 2.27, we know that $w_i < \infty$ for all $i \in I$. From Theorem 2.20 we know that if $\Delta \overset{\tau}{\Longrightarrow}_w \Delta'$ then $\langle\, w, \Delta' \,\rangle = \sum_{i \in I} p_i \cdot \langle\, w_i, \Delta'_i \,\rangle$

for some p_i with $\sum_{i\in I} p_i = 1$ and $\Delta \stackrel{\tau}{\Longrightarrow}_{w_i} \Delta'_i$. By Lemma 2.26, for each $i \in I$, there is some $\Delta'_{i,\varepsilon}$ such that $\Delta \stackrel{\tau}{\Longrightarrow}_{w_{i1}} (\Delta'_i + \Delta'_{i,\varepsilon})$, $\Delta = |\Delta'_i + \Delta'_{i,\varepsilon}|$, $\Delta'_{i,\varepsilon} \stackrel{\tau}{\Longrightarrow}_{w_{i2}} \varepsilon$ and $w_{i1} + w_{i2} = w_i$. Let $w_1 = \sum_{i\in I} p_i w_{i1}$, $w_2 = \sum_{i\in I} p_i w_{i2}$, $\Delta'_\varepsilon = \sum_{i\in I} p_i \cdot \Delta'_{i,\varepsilon}$. By Proposition 2.11(4), it can be seen that $\Delta \stackrel{\tau}{\Longrightarrow}_{w_1} (\Delta' + \Delta'_\varepsilon)$, $|\Delta| = |\Delta' + \Delta'_\varepsilon|$, $\Delta'_\varepsilon \stackrel{\tau}{\Longrightarrow}_{w_2} \varepsilon$ and $w_1 + w_2 = w$. □

Theorem 2.27 gives a useful property of bounded wMDPs, but there is a simpler criteria which ensures boundedness.

Theorem 2.29 Every *finitary* and *convergent* wMDP is also bounded.

Proof. In a finitary and convergent wMDP, suppose $\Delta \stackrel{\tau}{\Longrightarrow}_w \Delta'$. We show that the weight w is finite. Let $\mathsf{pp}_1, ..., \mathsf{pp}_n$ $(n \geq 1)$ be all the static policies in a finitary wMDP. Each static policy pp_i induces a deterministic sub-wMDP from the original wMDP, and determines a hyper-derivation $\Delta \stackrel{\tau}{\Longrightarrow}_{\mathsf{pp}_i, w_i} \Delta'_i$ from Δ. Clearly, the sub-wMDP is also convergent. By Lemma 2.24, we know that $w_i < \infty$ and $|\Delta| = |\Delta_i|$ for each i. Suppose $\Delta \stackrel{\tau}{\Longrightarrow}_w \Delta'$. It follows from Theorem 2.20 that $\langle w, \Delta' \rangle$ is an interpolation of $\langle w_1, \Delta'_1 \rangle, ..., \langle w_n, \Delta'_n \rangle$. Therefore, we have $|\Delta| = |\Delta'|$ and $w < \infty$. □

The final result of this section concerns closure with respect to parallel composition. This will be useful in Section 4, where we define a testing preorder between processes (Definition 4.10).

Theorem 2.30 If P is a bounded wMDP and Q is a finite wMDP, then their parallel composition $P \mid Q$ is bounded.

Proof. (*Schema*) We use the simple syntax to represent finite wMDPs.

$$Q \;:=\; \mathbf{0} \mid \bigoplus_{i\in I} p_i \cdot Q_i \mid \sum_{i\in I} \langle \alpha_i, w_i \rangle . Q_i$$

where $\mathbf{0}$ is the deadlock state, $\bigoplus_{i\in I} p_i \cdot Q_i$ represents a distribution that gives probability p_i to state Q_i, and $\sum_{i\in I} \langle \alpha_i, w_i \rangle . Q_i$ is a state that can nondeterministically evolve into state Q_i by performing action α_i with weight w_i. We prove by induction on the size of Q that if $P \mid Q \stackrel{\tau}{\Longrightarrow}_w \varepsilon$ then w is finite.

- $Q \equiv \mathbf{0}$. This is the base case. If $P \mid \mathbf{0} \stackrel{\tau}{\Longrightarrow}_w \varepsilon$ then obviously we have $P \stackrel{\tau}{\Longrightarrow}_w \varepsilon$. Since P is a bounded wMDP, we know that w is finite.
- $Q \equiv \bigoplus_{i\in I} p_i \cdot Q_i$. If $(P \mid \bigoplus_{i\in I} p_i \cdot Q_i) \stackrel{\tau}{\Longrightarrow}_w \varepsilon$, then we have $P \mid Q_i \stackrel{\tau}{\Longrightarrow}_{w_i} \varepsilon$ and $w = \sum_{i\in I} p_i w_i$. By induction hypothesis, each w_i is finite. It follows that w is also finite.
- $Q \equiv \sum_{i\in I} \langle \alpha_i, w_i \rangle . Q_i$. Note that it is easy to see Q generates a finitary wMDP. By Theorem 2.20 it suffices to show that, for each static policy pp which determines the hyper-SP-derivation $P \mid Q \stackrel{\tau}{\Longrightarrow}_{\mathsf{pp}, w} \varepsilon$, the weight w is finite, because the finite generability theorem ensures that the weight of a general hyper-derivation is the convex combination of the weights given by static policies. We prove this using a schema similar to that in the proof of Lemma 2.26.

  We call a state in the compound wMDP $P \mid Q$ *productive* if it is in the form $P' \mid Q$ and $\mathsf{pp}(P' \mid Q) = \langle w_i, P'' \mid Q_i \rangle$ for some $i \in I$ and P''. That is, Q has participated in the transition $P' \mid Q \xrightarrow{\tau}_{w_i} P'' \mid Q_i$. We modify pp so as to obtain a static policy pp' by setting $\mathsf{pp}'(s) = \mathsf{pp}(s)$ except when s is productive, in which case we set $\mathsf{pp}'(s)\uparrow$. The new policy determines a unique hyper-SP-derivation $P \mid Q \overset{\tau}{\Longrightarrow}_{\mathsf{pp}',w_1} \Delta$ for some w_1 and Δ, and induces a sub-wMDP from the wMDP induced by pp. The subdistribution Δ is in the form $P' \mid Q$ because Q does not participate in any τ-transition in order to derive Δ, and there is a hyper-derivation in P such that $P \overset{\tau}{\Longrightarrow}_{w_1} P'$. Since P is bounded, we know that w_1 is finite. We split Δ up into $\Delta_1 + \Delta_2$ so that each state in $\lceil\Delta_2\rceil$ is productive under policy pp and Δ_1 is supported by all other states, if there are any at all. From Δ_2 the policy pp determines the hyper-SP-derivation $\Delta_2 \overset{\tau}{\Longrightarrow}_{\mathsf{pp},w_2} \varepsilon$ for some w_2. Then there are some w_{2s} such that $w_2 = \sum_{s\in\lceil\Delta_2\rceil} \Delta_2(s) \cdot w_{2s}$ and $\overline{s} \overset{\tau}{\Longrightarrow}_{\mathsf{pp},w_{2s}} \varepsilon$ for each $s \in \lceil\Delta_2\rceil$. Since each state s in $\lceil\Delta_2\rceil$ is productive, it must be in the form $P_s \mid Q$ and make the transitions $P_s \mid Q \xrightarrow{\tau}_{w_s} P'' \mid Q_i \overset{\tau}{\Longrightarrow}_{\mathsf{pp},w'_s} \varepsilon$ with $w_s + w'_s = w_{2s}$. By induction hypothesis, the weight w'_s is finite. Then w_{2s} is finite because w_s trivially is. It follows that w_2 is finite. Combining the two hyper-SP-derivations $P \mid Q \overset{\tau}{\Longrightarrow}_{\mathsf{pp}',w_1} \Delta_1 + \Delta_2$ and $\Delta_2 \overset{\tau}{\Longrightarrow}_{\mathsf{pp},w_2} \varepsilon$ we have $P \mid Q \overset{\tau}{\Longrightarrow}_{\mathsf{pp}',w_1} \Delta_1 + \Delta_2 \overset{\tau}{\Longrightarrow}_{\mathsf{pp},w_2} \Delta_1$. As we only divide the original hyper-SP-derivation into two stages, and does not change the τ transition from each state, the overall weight and the final subdistribution will not change, thus we have $w_1 + w_2 = w$ and $\Delta_1 = \varepsilon$. Since both w_1 and w_2 are shown to be finite, it follows that w is finite as well.

□

3 Amortised weighted simulations

3.1 Introduction

Here we assume some wMDP $\langle S, \mathsf{Act}_\tau, \mathbb{R}_{\geq 0}, \longrightarrow\rangle$. Weighed simulations can be defined either at the distribution level or at the state level. We choose the latter.

Definition 3.1 Given a relation $\mathcal{R} \subseteq S \times (\mathbb{R}_{\geq 0} \times \mathcal{D}(S))$, let $\mathcal{S}(\mathcal{R}) \subseteq S \times (\mathbb{R}_{\geq 0} \times \mathcal{D}(S))$ be the relation defined by letting $s\, \mathcal{S}(\mathcal{R})\, \langle r, \Theta \rangle$ whenever

$$s \xrightarrow{\alpha}_v \Delta \text{ implies the existence of some } w \text{ and } \Theta' \text{ such that } \Theta \overset{\alpha}{\Longrightarrow}_w \Theta' \text{ and } \Delta\ \overline{\mathcal{R}}\ \langle r + w - v, \Theta' \rangle$$

The operator $\mathcal{S}(-)$ is monotonic and so it has a maximal fixed point, which we denote by $\lhd$. We often write $s \lhd_r \Theta$ for $s \lhd \langle r, \Theta \rangle$ and use $\Delta \sqsubseteq_{sim} \Theta$ to mean that there is some initial investment r such that $\Delta\ \overline{\lhd}_r\ \Theta$. □

The basic idea here is that $s \lhd_r \Theta$ intuitively means that Θ can simulate the actions of s but with *more benefit*, or at least not less benefit. The parameter r should be viewed as compensation which Θ has accumulated which can be used in local comparisons between the benefits of individual actions. Thus when we simulate $s \xrightarrow{\alpha}_v \Delta$ with $\Theta \overset{\alpha}{\Longrightarrow}_w \Theta'$ there are two possibilities:

(i) $w > v$; here the accumulated compensation is increased from r to $r + (w - v)$. In subsequent rounds this extra compensation may be used to successfully simulate a heavier action with a lighter one.

(ii) $w \leq v$; here the compensation is decreased from r to $r - (v - w)$.

Finally it is important that $r \geq 0$, and remains greater than or equal to zero, or otherwise the presence of weights would have no effect. Thus in case (ii) if $(v - w) > r$ then the simulation is not successful.

We now show that with this formal definition of the relation $\sqsubseteq_{sim}$ the various statements asserted in the Introduction are true:

Example 3.2 Consider the first two systems, s_0 and t_0, viewed as wMDPs. Then the relation $\mathcal{R}$ given by

$$\mathcal{R} \;=\; \{(s_0,\; \langle\, r, \overline{t_0}\,\rangle) \mid r \geq 1\} \;\cup\; \{(s_d,\; \langle\, r, \overline{t_d}\,\rangle) \mid r \geq 0\}$$

is a simulation. Thus $s_0 \lhd_r \overline{t_0}$ for any $r \geq 1$. As pointed out in [KAK05] this example shows the need for the parametrisation with respect to initial investments r; Because of the weights associated with the action up an initial investment of at least one is required in order for $\overline{t_0}$ to be able to match s_0.

We also have $s_0 \lhd_r \overline{s_1}$ for any $r \geq 1$ because of the following simulation:

$$\mathcal{R} \;=\; \{(s_0,\; \langle\, r, \overline{s_1}\,\rangle) \mid r \geq 1\} \;\cup\; \{(s_d,\; \langle\, r, \Delta\,\rangle) \mid r \geq 0\}$$

where Δ is the distribution $\frac{1}{4} \cdot \overline{O} + \frac{3}{4} \cdot \overline{T}$. Note that this is indeed a simulation because $\Delta \stackrel{\mathsf{down}}{\Longrightarrow}_{2.5} \overline{s_1}$. Incidently this example shows why it is necessary to relate states to distributions, rather than states; there is no individual state accessible from s_1 which can simulate s_d.

Similarly $s_1 \lhd_r \overline{t_1}$ for every $r \geq 0$ because of the simulation:

$$\mathcal{R} \;=\; \{(s_1,\; \langle\, r, \overline{t_1}\,\rangle) \mid r \geq 0\} \;\cup\; \{(O,\; \langle\, r, \overline{U}\,\rangle) \mid r \geq 0\} \;\cup\; \{(T,\; \langle\, r, \overline{U}\,\rangle) \mid r \geq 0\}$$

Note that from Example 2.9 we have seen that $\overline{U} \stackrel{\tau}{\Longrightarrow}_3 \overline{D}$ and therefore by transitivity $\overline{U} \stackrel{\mathsf{down}}{\Longrightarrow}_4 \overline{t_1}$.

Finally $s_0 \lhd_2 s_2$ because of the following simulation:

$$\mathcal{R} \;=\; \{(s_0,\; \langle\, r, \overline{s_2}\,\rangle) \mid r \geq 2\} \;\cup\; \{(s_d,\; \langle\, r, \Delta\,\rangle) \mid r \geq 0\}$$

where Δ is the distribution $\frac{1}{4} \cdot \overline{S} + \frac{3}{4} \cdot \overline{T}$. Note that $\Delta \stackrel{\mathsf{down}}{\Longrightarrow}_3 \overline{s_2}$ although it is also possible for it to do the down action for much less benefit. □

Our first result about the simulation preorder $\lhd$ is that its lifting $\overline{\lhd}$ is a precongruence relation for the language CCMDP.

Lemma 3.3 1. If $\Delta \stackrel{\alpha}{\Longrightarrow}_r \Delta'$ then $\Delta \mid \Gamma \stackrel{\alpha}{\Longrightarrow}_r \Delta' \mid \Gamma$ and $\Gamma \mid \Delta \stackrel{\alpha}{\Longrightarrow}_r \Gamma \mid \Delta'$.

2. If $\Delta \stackrel{a}{\longrightarrow}_{r_1} \Delta'$ and $\Gamma \stackrel{\bar{a}}{\longrightarrow}_{r_2} \Gamma'$ then $\Delta \mid \Gamma \stackrel{\tau}{\longrightarrow}_{r_1+r_2} \Delta' \mid \Gamma'$.

Proof. Straightforward calculations. □

Theorem 3.4 The relation $\overline{\lhd}$ is a precongruence.

Proof. It is easy to verify that $\overline{\lhd}$ is closed under prefixing, nondeterministic choice, and hiding operators. Here we only show that the closure under parallel composition is also preserved, namely, if $\Delta \mathrel{\overline{\lhd}_r} \Theta$ then $(\Delta \mid \Gamma) \mathrel{\overline{\lhd}_r} (\Theta \mid \Gamma)$. We first construct the following relation

$$\mathcal{R} := \{(s \mid t, \langle r,\ \Theta \mid t\rangle) \mid s \lhd_r \Theta\}$$

and check that $\mathcal{R} \subseteq \lhd$. Suppose that $(s \mid t)\ \mathcal{R}_r\ (\Theta \mid t)$.

- If $s \mid t \xrightarrow{\alpha}_v \Delta \mid t$ because of the transition $s \xrightarrow{\alpha}_v \Delta$, then $\Theta \xRightarrow{\alpha}_w \Theta'$ and $\Delta \mathrel{\overline{\lhd}_{r+w-v}} \Theta'$. By Lemma 3.3 we have $\Theta \mid t \xRightarrow{\alpha}_w \Theta' \mid t$. It also holds that $(\Delta \mid t)\ \overline{\mathcal{R}}_{r+w-v}\ (\Theta' \mid t)$.
- If $s \mid t \xrightarrow{\alpha}_v s \mid \Gamma$ because of the transition $t \xrightarrow{\alpha}_v \Gamma$, then $\Theta \mid t \xrightarrow{\alpha}_v \Theta \mid \Gamma$ and we have that $(s \mid \Gamma)\ \overline{\mathcal{R}}_r\ (\Theta \mid \Gamma)$.
- If $s \mid t \xrightarrow{\tau}_v \Delta \mid \Gamma$ because of the transitions $s \xrightarrow{a}_{v_1} \Delta$ and $t \xrightarrow{\bar{a}}_{v_2} \Gamma$ with $v = v_1 + v_2$, then $\Theta \xRightarrow{a}_{w_1} \Theta'$ and $\Delta \mathrel{\overline{\lhd}_{r+w_1-v_1}} \Theta'$. By Lemma 3.3 we derive that $\Theta \mid t \xRightarrow{\tau}_{w_1+v_2} \Theta' \mid \Gamma$. Note that $r + (w_1 + v_2) - (v_1 + v_2) = r + w_1 - v_1$ and $(\Delta \mid \Gamma)\ \overline{\mathcal{R}}_{r+w_1-v_1}\ (\Theta' \mid \Gamma)$.

So we have shown that $\mathcal{R}$ is a simulation relation. It follows that $\Delta \mathrel{\overline{\lhd}_r} \Theta$ implies $(\Delta \mid \Gamma)\ \overline{\mathcal{R}}_r\ (\Theta \mid \Gamma)$, thus $(\Delta \mid \Gamma) \mathrel{\overline{\lhd}_r} (\Theta \mid \Gamma)$. □

Example 3.5 Let P, Q be two processes with $P \lhd_0 \overline{Q}$. Consider the following processes:

$$\begin{array}{lcl} U & \Leftarrow & \tau_0.(\tau_1.U \,{}_{\frac{3}{4}}\!\oplus \mathsf{down}_1.Q) \\ P' & \equiv & \mathsf{up}_2.(\mathsf{down}_1.P \,{}_{\frac{1}{4}}\!\oplus \mathsf{down}_3.P) \\ Q' & \equiv & \mathsf{up}_2.U \end{array}$$

By the analysis in Example 2.9 we know that $\overline{U} \xRightarrow{\tau}_3 \overline{\mathsf{down}_1.Q}$, thus $\overline{U} \xRightarrow{\mathsf{down}}_4 \overline{Q}$. Then it is easy to see that $\mathsf{down}_1.P \lhd_0 \overline{U}$ and $\mathsf{down}_3.P \lhd_0 \overline{U}$. It follows from the compositionality of $\overline{\lhd}_0$ that $(\mathsf{down}_1.P \,{}_{\frac{1}{4}}\!\oplus \mathsf{down}_3.P) \mathrel{\overline{\lhd}_0} \overline{U}$ and furthermore $\overline{P'} \mathrel{\overline{\lhd}_0} \overline{Q'}$. □

Note that in Definition 3.1 for $s \lhd_r \Theta$ to be true we only require that strong moves from s be matched by weak moves from Θ; this restriction makes the proof of the congruence result, Theorem 3.4, relatively straightforward. But later, in particular when giving a logical characterisation of the simulation preorder, it will be useful to know that this transfer property is also true for weak moves from s. We end this section with a proof of this result, which first requires a lemma.

Lemma 3.6 Let Δ and Θ be two subdistributions in a bounded wMDP. Suppose $\Delta \mathrel{\overline{\lhd}_r} \Theta$ for some $r \in \mathbb{R}_{\geq 0}$. If $\Delta \xrightarrow{\alpha}_v \Delta'$ then $\Theta \xRightarrow{\alpha}_w \Theta'$ for some w and Θ' such that $\Delta' \mathrel{\overline{\lhd}_{r+w-v}} \Theta'$.

Proof. Note that in the statement of the lemma both Δ and Θ are in general subdistributions. Although the relations $\lhd_r$ only relate states to full distributions, the lifted relations $\overline{\lhd}_r$ are relations over subdistributions.

Suppose $\Delta \mathrel{\overline{\lhd}_r} \Theta$ and $\Delta \xrightarrow{\alpha}_v \Delta'$. By Lemma 2.3 there is an index set I such that (i) $\Delta = \sum_{i\in I} p_i \cdot \overline{s_i}$, (ii) $r = \sum_{i\in I} p_i r_i$, (iii) $\Theta = \sum_{i\in I} p_i \cdot \Theta_i$, and (iv) $s_i \lhd_{r_i} \Theta_i$ for each $i \in I$ with $\sum_{i\in I} p_i \leq 1$. By the condition $\Delta \xrightarrow{\alpha}_v \Delta'$, (i) and Proposition 2.6, there are some weights v_i

and subdistributions Δ'_i such that $v = \sum_{i\in I} p_i v_i$, $\Delta' = \sum_{i\in I} p_i \cdot \Delta'_i$, and $\overline{s_i} \xrightarrow{\alpha}_{v_i} \Delta'_i$ for each $i \in I$. By Lemma 2.3 again, for each $i \in I$, there is an index set J_i such that $v_i = \sum_{j\in J_i} q_{ij} v_{ij}$, $\Delta'_i = \sum_{j\in J_i} q_{ij} \cdot \Delta'_{ij}$ and $s_i \xrightarrow{\alpha}_{v_{ij}} \Delta'_{ij}$ for each $j \in J_i$ and $\sum_{j\in J_i} q_{ij} = 1$. By (iv) there is some w_{ij} and Θ'_{ij} such that $\Theta_i \stackrel{\alpha}{\Longrightarrow}_{w_{ij}} \Theta'_{ij}$ and $\Delta'_{ij} \,\overline{\lhd}_{r_i + w_{ij} - v_{ij}}\, \Theta'_{ij}$. Let $w = \sum_{i\in I, j\in J_i} p_i q_{ij} w_{ij}$ and $\Theta' = \sum_{i\in I, j\in J_i} p_i q_{ij} \cdot \Theta'_{ij}$. By Proposition 2.11 the relation $\stackrel{\tau}{\Longrightarrow}$ is linear, from which it follows that $\stackrel{\alpha}{\Longrightarrow}$ is also linear for an arbitrary α. It follows that $\Theta = \sum_{i\in I} p_i \sum_{j\in J_i} q_{ij} \cdot \Theta_i \stackrel{\alpha}{\Longrightarrow}_w \Theta'$. By the linearity of $\overline{\lhd}$, we conclude that $\Delta' = (\sum_{i\in I} p_i \sum_{j\in J_i} q_{ij} \cdot \Delta'_{ij}) \,\overline{\lhd}_{r+w-v}\, \Theta'$. □

Proposition 3.7 [Weak transfer property] Let s be a state and Θ a distribution in a bounded wMDP such that $s \lhd_r \Theta$ for some $r \in \mathbb{R}_{\geq 0}$. Suppose $\overline{s} \stackrel{\alpha}{\Longrightarrow}_v \Delta'$ where Δ' is again a distribution. Then $\Theta \stackrel{\alpha}{\Longrightarrow}_w \Theta'$ for some w and Θ' such that $\Delta' \,\overline{\lhd}_{r+w-v}\, \Theta'$.

Proof. Before embarking on the proof first note that we are assured that the matching Θ' in the statement of the lemma is also a distribution. Using the characterisation in Lemma 2.3, it is easy to check that if $\Delta \,\overline{\mathcal{R}}\, \langle r, \Theta \rangle$ for any relation $\mathcal{R} \subseteq S \times (\mathbb{R}_{\geq 0} \times \mathcal{D}(S))$ then $|\Delta| = |\Theta|$. Since $\Delta' \,\overline{\lhd}_{r+w-v}\, \Theta'$ and Δ' is a distribution it follows that Θ' must also be a distribution.

We give the proof in the case when α is τ; the case for $a \in \mathsf{Act}$ follows from this in a straightforward manner. Suppose $s \lhd_r \Theta$ and $\overline{s} \stackrel{\tau}{\Longrightarrow}_v \Delta'$ with $|\Delta'| = 1$. So there are Δ_k, $\Delta_k^{\rightarrow}$ and $\Delta_k^{\times}$ for $k \geq 0$ such that $\overline{s} = \Delta_0$, $\Delta_k = \Delta_k^{\rightarrow} + \Delta_k^{\times}$, $\Delta_k^{\rightarrow} \xrightarrow{\tau}_{v_{k+1}} \Delta_{k+1}$, $v = \sum_{k=1}^{\infty} v_k$ and $\Delta' = \sum_{k=0}^{\infty} \Delta_k^{\times}$. Since $\Delta_0^{\rightarrow} + \Delta_0^{\times} = \overline{s} \,\overline{\lhd}_r\, \Theta$, by Proposition 2.6 we can make the decomposition $\Theta = \Theta_0^{\rightarrow} + \Theta_0^{\times}$ so that $\Delta_0^{\rightarrow} \,\overline{\lhd}_{r_0^{\rightarrow}}\, \Theta_0^{\rightarrow}$ and $\Delta_0^{\times} \,\overline{\lhd}_{r_0^{\times}}\, \Theta_0^{\times}$ for some $r_0^{\rightarrow}, r_0^{\times}$ with $r_0^{\rightarrow} + r_0^{\times} = r$. Since $\Delta_0^{\rightarrow} \xrightarrow{\tau}_{v_1} \Delta_1$ and $\Delta_0^{\rightarrow} \,\overline{\lhd}_{r_0^{\rightarrow}}\, \Theta_0^{\rightarrow}$, by Lemma 3.6 we have $\Theta_0^{\rightarrow} \stackrel{\tau}{\Longrightarrow}_{w_1} \Theta_1$ with $\Delta_1 \,\overline{\lhd}_{(r_0^{\rightarrow} + w_1 - v_1)}\, \Theta_1$.

Repeating the above procedure gives us inductively a series $\Theta_k, \Theta_k^{\rightarrow}, \Theta_k^{\times}$ of subdistributions, for $k \geq 0$, and weights $r_k^{\rightarrow}, r_k^{\times}$, for $k \geq 1$, such that $\Theta = \Theta_0$, $\Delta_k \,\overline{\lhd}_{(r_{k-1}^{\rightarrow} + w_k - v_k)}\, \Theta_k$, $\Theta_k = \Theta_k^{\rightarrow} + \Theta_k^{\times}$, $\Delta_k^{\rightarrow} \,\overline{\lhd}_{r_k^{\rightarrow}}\, \Theta_k^{\rightarrow}$, $\Delta_k^{\times} \,\overline{\lhd}_{r_k^{\times}}\, \Theta_k^{\times}$, $\Theta_k^{\rightarrow} \stackrel{\tau}{\Longrightarrow}_{w_{k+1}} \Theta_{k+1}$ and $r_{k-1}^{\rightarrow} + w_k - v_k = r_k^{\rightarrow} + r_k^{\times}$. We define $\Theta' = \sum_{k=0}^{\infty} \Theta_k^{\times}$, $w = \sum_{k=1}^{\infty} w_k$ and $r' = \sum_{k=0}^{\infty} r_k^{\times}$. It follows from Definition 2.2 that $\Delta' \,\overline{\lhd}_{r'}\, \Theta'$. Below we show that $\Theta \stackrel{\tau}{\Longrightarrow}_w \Theta'$ and $r' = r + w - v$.

By the transitivity of hyper-derivations, Theorem 2.13, it can be established that $\Theta \stackrel{\tau}{\Longrightarrow}_{\sum_{k\leq i} w_k} (\Theta_i^{\rightarrow} + \sum_{k\leq i} \Theta_k^{\times})$ for each $i \geq 0$. Since $|\Delta'| = 1$, we must have $\lim_{i\to\infty} |\Delta_i^{\rightarrow}| = 0$. Again using the characterisation in Lemma 2.3 we know that $|\Theta_i^{\rightarrow}| = |\Delta_i^{\rightarrow}|$ for each i. Therefore, since $\Delta_i^{\rightarrow} \,\overline{\lhd}_{r_i^{\rightarrow}}\, \Theta_i^{\rightarrow}$, we then have $\lim_{i\to\infty} |\Theta_i^{\rightarrow}| = 0$. Thus, $\lim_{i\to\infty} (\Theta_i^{\rightarrow} + \sum_{k\leq i} \Theta_k^{\times}) = \sum_{k=0}^{\infty} \Theta_k^{\times} = \Theta'$. We also have $\lim_{i\to\infty} \sum_{k\leq i} w_k = \sum_{k=1}^{\infty} w_k = w$. In Appendix C, specifically in Corollary C.1, we show that the set $\{\langle v, \Gamma \rangle \mid \Theta \stackrel{\tau}{\Longrightarrow}_v \Gamma\}$ is compact. From this it follows that $\Theta \stackrel{\tau}{\Longrightarrow}_w \Theta'$.

By an easy inductive proof it can be seen that $r = r_i^{\rightarrow} + \sum_{k\leq i} r_k^{\times} + \sum_{k<i} v_k - \sum_{k<i} w_k$ for each $i \geq 0$. From $\lim_{i\to\infty} |\Delta_i^{\rightarrow}| = 0$ and $\Delta_i^{\rightarrow} \,\overline{\lhd}_{r_i^{\rightarrow}}\, \Theta_i^{\rightarrow}$ it follows that $\lim_{i\to\infty} r_i^{\rightarrow} = 0$. Therefore, $r = \sum_{k\geq 0} r_k^{\times} + \sum_{k\geq 0} v_k - \sum_{k\geq 0} w_k = r' + v - w$, i.e. $r' = r + w - v$. □

This weak transfer property is easily generalised to distributions:

Corollary 3.8 Suppose $\Delta \,\overline{\lhd}_r\, \Theta$ for some $r \in \mathbb{R}_{\geq 0}$, where Δ, Θ are two distributions in a bounded wMDP. If $\Delta \stackrel{\alpha}{\Longrightarrow}_v \Delta'$, where Δ' is also a distribution, then there exists a distribution Θ' such that $\Theta \stackrel{\alpha}{\Longrightarrow}_w \Theta'$ and $\Delta' \,\overline{\lhd}_{r+w-v}\, \Theta'$.

Proof. Combining Proposition 2.11 and Proposition 3.7. □

3.2 Infinite approximation

The simulation relations $\lhd_r$ are defined coinductively. But in bounded wMDPs they can also be characterised inductively.

Definition 3.9 For every $k \geq 0$ we define the relation $\lhd^k \subseteq S \times (\mathbb{R}_{\geq 0} \times \mathcal{D}(S))$ as follows:

(i) $\lhd^0 = S \times (\mathbb{R}_{\geq 0} \times \mathcal{D}(S))$

(ii) $\lhd^{k+1} = \mathcal{S}(\lhd^k)$.

Finally we let $\lhd^\infty$ be $\bigcap_{k=0}^{\infty} \lhd^k$. □

Standard arguments ensure that $\lhd_r \subseteq \lhd_r^k$ for every $k \geq 0$ and therefore that $\lhd_r \subseteq \lhd_r^\infty$. The converse is also true in bounded wMDPs, as we now demonstrate, using compactness arguments.

We note that the metric space $(\mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S), d_1)$ equipped with the distance function

$$d_1(\langle v, \Delta \rangle, \langle w, \Theta \rangle) = \max(\{|w - v|\} \cup \{|\Delta(s) - \Theta(s)| \quad | \; s \in S\})$$

is complete. Provided the set S is finite, the distance function on the space $(S \to \mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S), d_2)$ given by $d_2(f, g) = \max_{s \in S} d_1(f(s), g(s))$ is well-defined, and the resulting metric space is also complete.

Proposition 3.10 For any subdistribution Δ in a bounded wMDP the set $\{\langle w, \Delta' \rangle \mid \Delta \stackrel{\alpha}{\Longrightarrow}_w \Delta'\}$ is closed.

Proof. See Appendix C; a more general result is given in Lemma C.6. □

Definition 3.11 A relation $\mathcal{R} \subseteq S \times (\mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S))$ is *closed* (resp. *bounded*) if for every $s \in S$ the set $s \cdot \mathcal{R} = \{ \langle w, \Delta \rangle \mid s \; \mathcal{R} \; \langle w, \Delta \rangle \}$ is closed (resp. bounded). It is *compact* if it is both closed and bounded. □

The main technical result we require is the following:

Proposition 3.12 In a bounded wMDP, for every $k \in \mathbb{N}$, the relation $\lhd^k$ is closed and convex.

Proof. Because of the style of argument required this proof is also relegated to Appendix C. □

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

Lemma 3.13 Let S be a finite set of states. Suppose $\mathcal{R}^k \subseteq S \times (\mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S))$ is a sequence of closed and convex relations such that $\mathcal{R}^{(k+1)} \subseteq \mathcal{R}^k$. Then it holds that

$$(\cap_{k=0}^{\infty} \overline{\mathcal{R}^k}) \subseteq \overline{(\cap_{k=0}^{\infty} \mathcal{R}^k)} \; .$$

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

Since $\mathcal{R}^k$ is closed and convex for each k, it follows that $\mathcal{R}^\infty$ is also closed and convex. Moreover it is easy to check that the set of choice functions $\mathbf{Ch}(\mathcal{R})$ is also closed. Therefore, $\mathbf{Ch}(\mathcal{R}^k)$ for each $k \in \mathbb{N}$ and $\mathbf{Ch}(\mathcal{R}^\infty)$ are closed.

Now consider

$$G = \{\, f : S \to \mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S) \mid \langle w, \Theta \rangle = \mathrm{Exp}_\Delta(f) \,\}$$

which is easily seen to be a closed set. Consider the collection of closed sets $H^k = \mathbf{Ch}(\mathcal{R}^k) \cap G$; since $\Delta\ \overline{\mathcal{R}^k}\ \langle w, \Theta \rangle$, Proposition 2.5 assures us that all of these are non-empty. Also $H^{(k+1)} \subseteq H^k$ and therefore by the finite-intersection property [Lip65] $\cap_{k=0}^\infty H^k$ is also non-empty.

Let f be an arbitrary element of this intersection. For any state $s \in \mathrm{dom}(\mathcal{R}^\infty)$, and for every $k \geq 0$, we have $s \in \mathrm{dom}(\mathcal{R}^k)$ because $\mathrm{dom}(\mathcal{R}^\infty) \subseteq \mathrm{dom}(\mathcal{R}^k)$. Therefore, $s\ \mathcal{R}^k\ f(s)$ as $f \in \mathbf{Ch}(\mathcal{R}^k)$. It follows that $s\ \mathcal{R}^\infty\ f(s)$. So f is a choice function for $\mathcal{R}^\infty$, $f \in \mathbf{Ch}(\mathcal{R}^\infty)$. From Proposition 2.5 it follows that $\Delta\ \overline{\mathcal{R}^\infty}\ \mathrm{Exp}_\Delta(f)$. But from the definition of the G we know that $\langle w, \Theta \rangle = \mathrm{Exp}_\Delta(f)$, and the required result follows. □

Theorem 3.14 In a bounded wMDP, $s \lhd_r \Theta$ if and only if $s \lhd_r^\infty \Theta$.

Proof. Since $\lhd \subseteq \lhd^\infty$ it is sufficient to show the opposite inclusion, which by definition holds if $\lhd^\infty$ is a simulation, viz. if $\lhd^\infty \subseteq \mathcal{S}(\lhd^\infty)$. Suppose $s \lhd_r^\infty \Theta$, which means that $s \lhd_r^k \Theta$ for every $k \geq 0$. In order to show $s\ \mathcal{S}(\lhd^\infty)_r\ \Theta$ we have to establish that if $s \xrightarrow{\alpha}_v \Delta'$ then $\Theta \stackrel{\alpha}{\Longrightarrow}_w \Theta'$ for some Θ' such that $\Delta'\ \overline{\lhd^\infty}_{(r+w-v)}\ \Theta'$.

For every $k \geq 0$ there exists some w_k, Θ'_k such that $\Theta \stackrel{\alpha}{\Longrightarrow}_{w_k} \Theta'_k$ and $\Delta'\ \overline{\lhd^k}_{(r+w_k-v)}\ \Theta'_k$. Now construct the sets

$$D^k = \{\, \langle w, \Theta' \rangle \mid \Theta \stackrel{\alpha}{\Longrightarrow}_w \Theta' \text{ and } \Delta'\ \overline{\lhd^k}_{(r+w-v)}\ \Theta' \,\}.$$

We now argue that each D^k is a closed set.

To do so we rewrite it to a form which makes the fact that it is closed obvious. First define the function $\mathcal{E} : \mathcal{D}_{sub}(S) \times (S \to \mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S)) \to \mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S)$ by $\mathcal{E}(\Theta, f) = \mathrm{Exp}_\Theta(f)$, which is obviously continuous. It is also a closed function, meaning that the image of every closed set under $\mathcal{E}$ is closed, because positive scaling and sum are operations that preserve closedness of functions. Because of Proposition 2.5 it can be shown that

$$D^k = (\Theta \stackrel{\alpha}{\Longrightarrow}) \cap G^{-1} \circ \mathcal{E}(\{\Delta'\} \times \mathbf{Ch}(\lhd^k)),$$

where $G(\langle w, \Theta' \rangle) = \langle r + w - v, \Theta' \rangle$. Thus by Proposition 3.10 and Proposition 3.12 they 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 $\langle w, \Theta' \rangle$ in it we know $\Theta \stackrel{\alpha}{\Longrightarrow}_w \Theta'$ and $\Delta'\ \overline{\lhd^k}_{(r+w-v)}\ \Theta'$ for every $k \geq 0$; that is $\Delta'\,(\cap_{k=0}^\infty \overline{\lhd^k}_{(r+w-v)})\,\Theta'$. By Proposition 3.12, the relations $\lhd^k$ are all closed and convex. Therefore, Lemma 3.13 may be applied to them, which enables us to conclude $\Delta'\ \overline{\lhd^\infty}_{(r+w-v)}\ \Theta'$.

□

3.3 Modal logic

As part of our argument in favour of amortised simulations for wMDPs we show that it has associated with it a natural property or modal logic. We show two results. The first is that a finitary version characterises the behavioural preorders $\lhd_r$. Secondly we show that with the addition of fixed points we can capture logically the behaviour of any state. Both results are restricted to bounded wMDPs.

Here we develop a modal logic which characterises the relations $\lhd_r^\infty$ in an arbitrary wMDP, and thus $\lhd_r$ in bounded wMDPs.

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

- $\mathsf{tt} \in \mathcal{L}$
- $\langle\alpha\rangle_w(\phi_1\,{}_p\!\oplus\,\phi_2) \in \mathcal{L}$ when $\phi_i \in \mathcal{L}$, $\alpha \in \mathsf{Act}_\tau$, $w \in \mathbb{R}_{\geq 0}$ and $p \in [0,1]$
- $\phi_1 \wedge \phi_2 \in \mathcal{L}$ when $\phi_1,\ \phi_2 \in \mathcal{L}$

Let $\mathcal{C}on$ denote the set of all pairs $\langle r, \Delta\rangle$, called *configurations*, where $r \in \mathbb{R}_{\geq 0}$ and $\Delta \in \mathcal{D}(S)$, with S denoting the state space of some wMDP. Intuitively this represents a probabilistic system which has accumulated compensation r which it can use to satisfy formulae in the future. The satisfaction relation $\models\ \subseteq \mathcal{C}on \times \mathcal{L}$ is now given by:

(i) $\langle r, \Delta\rangle \models \mathsf{tt}$ for every configuration

(ii) $\langle r, \Delta\rangle \models \phi_1 \wedge \phi_2$ whenever $\langle r, \Delta\rangle \models \phi_1$ and $\langle r, \Delta\rangle \models \phi_2$

(iii) $\langle r, \Delta\rangle \models \langle\alpha\rangle_v(\phi_1\,{}_p\!\oplus\,\phi_2)$ whenever $\Delta \stackrel{\alpha}{\Longrightarrow}_w \Delta'$, $\langle r+w-v, \Delta'\rangle = \langle r_1, \Delta'_1\rangle\,{}_p\!\oplus\,\langle r_2, \Delta'_2\rangle$, and $\langle r_i, \Delta'_i\rangle \models \phi_i$.

Let $\mathcal{L}(r,\Delta) = \{\, \phi \in \mathcal{L} \mid \langle r, \Delta\rangle \models \phi \,\}$.

The idea here is that $\langle r, \Delta\rangle$ represents a process which has built up compensation r which it can use to help satisfy a formula. The principal formula is $\langle\alpha\rangle_v\phi$ which represents the ability to do an α action with benefit at least v and then satisfy ϕ. In (iii) above when this is satisfied by $\langle r, \Delta\rangle$ because $\Delta \stackrel{\alpha}{\Longrightarrow}_w \Delta'$ there are two possibilities:

(i) $v > w$: here the compensation comes into play. The action may be accepted despite being too heavy but the compensation is reduced from r to $r-(v-w)$; note this is only possible if this sum $r-(v-w) \geq 0$.

(ii) $v \leq w$: The action is accepted and the compensation is increased from r to $r+(w-v)$.

For convenience of presentation, we generalise binary probabilistic choice to be n-ary and often write $\langle\alpha\rangle_w \bigoplus_{i\in I} p_i \cdot \phi_i$ for finite index set I. It is easy to see that, for instance,

$$\langle r, \Delta\rangle \models \langle\alpha\rangle_w \bigoplus_{i=1..3} p_i \cdot \phi_i \quad \text{if and only if} \quad \langle r, \Delta\rangle \models \langle\alpha\rangle_w(\phi_1\,{}_{p_1}\!\oplus\,(\langle\tau\rangle_0(\phi_2\,{}_{\frac{p_2}{1-p_1}}\!\oplus\,\phi_3)))$$

for any configuration $\langle\, r, \Delta \,\rangle$.

The modal logic $\mathcal{L}$ has a limited number of operators, and for this reason the satisfaction relation is in some sense impervious to hyper-derivations:

Lemma 3.15 Suppose $\Delta \stackrel{\tau}{\Longrightarrow}_w \Delta'$ and $r + w \geq r'$. Then $\langle r', \Delta' \rangle \models \phi$ implies $\langle r, \Delta \rangle \models \phi$.

Proof. By structural induction on ϕ. □

Proposition 3.16 Suppose in a bounded wMDP that $\Delta \mathrel{\overline{\lhd}_r} \Theta$. Then for every $\phi \in \mathcal{L}$, $\langle r_\Delta, \Delta \rangle \models \phi$ implies $\langle r_\Delta + r, \Theta \rangle \models \phi$.

Proof. By structural induction on ϕ.

- $\phi \equiv \mathsf{tt}$. This case is trivial.
- $\phi \equiv \phi_1 \wedge \phi_2 \in \mathcal{L}$ we appeal to induction hypothesis directly.
- $\phi \equiv \langle \alpha \rangle_c (\phi_1 \,{}_p\!\oplus\, \phi_2)$. Suppose $\langle r_\Delta, \Delta \rangle \models \phi$. This means $\Delta \stackrel{\alpha}{\Longrightarrow}_v \Delta'$, $\langle r_\Delta + v - c, \Delta' \rangle = \langle r_1, \Delta'_1 \rangle \,{}_p\!\oplus\, \langle r_2, \Delta'_2 \rangle$, and $\langle r_i, \Delta'_i \rangle \models \phi_i$ for $i = 1, 2$. Note that by the definition of the satisfaction relation $\models$ we know that $\Delta, \Delta' \in \mathcal{D}(S)$, i.e. $|\Delta| = |\Delta'| = 1$. In a bounded wMDP we know from Corollary 3.8 that $\overline{\lhd_r}$ also satisfies the transfer property for weak moves and therefore $\Theta \stackrel{\alpha}{\Longrightarrow}_w \Theta'$ such that $\Delta' \mathrel{\overline{\lhd}_{(r+w-v)}} \Theta'$. By Proposition 2.6, $\langle\, r + w - v, \Theta' \,\rangle = \langle\, t_1, \Theta'_1 \,\rangle \,{}_p\!\oplus\, \langle\, t_2, \Theta'_2 \,\rangle$ so that $\Delta'_i \mathrel{\overline{\lhd}_{t_i}} \Theta'_i$ for $i = 1, 2$. By induction hypothesis, we have $\langle r_i + t_i, \Theta'_i \rangle \models \phi_i$ for $i = 1, 2$. Since

  $$(r_1 + t_1)\,{}_p\!\oplus (r_2 + t_2) \;=\; (r_\Delta + v - c) + (r + w - v) \;=\; (r_\Delta + r) + w - c$$

  it follows that $\langle r_\Delta + r, \Theta \rangle \models \phi$.

□

Theorem 3.17 In a bounded wMDP, $\mathcal{L}(0, \overline{s}) \subseteq \mathcal{L}(r, \Theta)$ implies $\overline{s} \mathrel{\overline{\lhd}_r} \Theta$.

Proof. Since we assume the wMDP is bounded, by Theorem 3.14 it is sufficient to prove the result for $\lhd^\infty$ rather than $\lhd$. Thus we have to show that for every $k \geq 0$, $\mathcal{L}(0, \overline{s}) \subseteq \mathcal{L}(r, \Theta)$ implies $\overline{s} \mathrel{\overline{\lhd^k}_r} \Theta$. This will follow immediately if for every state s and every index k we can construct the k-th *characteristic formulae* ϕ^k_s satisfying:

(a) $\langle 0, \overline{s} \rangle \models \phi^k_s$

(b) $\langle r, \Theta \rangle \models \phi^k_s$ implies $s \lhd^k_r \Theta$.

The construction is by induction on k:

(i) $\phi^0_s = \mathsf{tt}$

(ii) $\phi^{(k+1)}_s = \bigwedge_{s \stackrel{\alpha}{\longrightarrow}_w \Delta} \langle \alpha \rangle_w \phi^k_\Delta$

(iii) $\phi^k_\Delta = \bigoplus_{s \in \lceil \Delta \rceil} \Delta(s) \cdot \phi^k_s$.

The proof that properties (a) and (b) are satisfied proceeds by induction on k, with the case $k = 0$ being trivial. As an example of the inductive case we first show $\langle r, \Theta\rangle \models \phi_s^{(k+1)}$ implies $s \lhd_r^{(k+1)} \Theta$.

So let us assume $\langle r, \Theta\rangle \models \phi_s^{(k+1)}$. Let $s \xrightarrow{\alpha}_v \Delta$ be an arbitrary move from s; because of the construction of the characteristic formula we have that $\langle r, \Theta\rangle \models \langle\alpha\rangle_v \phi_\Delta^k$. By definition this means $\Theta \stackrel{\alpha}{\Longrightarrow}_w \Theta'$, where $\langle\, (r + w - v), \Theta' \,\rangle = \sum_{s\in\lceil\Delta\rceil} \Delta(s) \cdot \langle\, r_s, \Theta'_s \,\rangle$ and $\langle r_s, \Theta'_s\rangle \models \phi_s^k$. At this point we invoke induction to obtain $s \lhd_{r_s}^k \Theta'_s$ from which it follows by the definition of lifting that $\Delta \overline{\lhd_{(r+w-v)}^k} \Theta'$. Therefore, we have verified that $s \lhd_r^{(k+1)} \Theta$. □

As an immediate corollary we have a logical characterisation of our simulation preorder.

Corollary 3.18 In a bounded wMDP, $\overline{s} \overline{\lhd}_r \Theta$ if and only if $\mathcal{L}(0, \overline{s}) \subseteq \mathcal{L}(r, \Theta)$.

Proof. Combining Proposition 3.16 and Theorem 3.17. □

We now turn our attention to **characteristic formulae.** To this end we extend the modal logic $\mathcal{L}$ with a fixed point operator.

Let *Var* be a countable set of variables. We define a set $\mathcal{L}_{\texttt{fix}}$ of modal formulae by the following grammar:

$$\phi \quad := \quad \mathsf{tt} \mid \langle\alpha\rangle_w(\phi_1\,{}_p\!\oplus\,\phi_2) \mid \phi_1 \wedge \phi_2 \mid X \mid \mathsf{max}\; X.\phi$$

where $\alpha \in \mathsf{Act}_\tau$ $w \in \mathbb{R}_{\geq 0}$ and $p \in [0, 1]$. Sometimes we also use the finite conjunction $\bigwedge_{i\in I} \phi_i$. As usual, we have $\bigwedge_{i\in\emptyset} \phi_i = \mathsf{tt}$. The fixed point operator $\mathsf{max}\; X$ binds the variable X. We apply the usual terminology of free and bound variables in a formula and write $\mathit{fv}(\phi)$ for the set of free variables in ϕ.

We use *environments*, which binds free variables to sets of distributions, in order to give semantics to formulae. We fix a bound wMDP and let S be its state set. Let

$$\mathit{Env} = \{\, \rho \;\mid\; \rho : \mathit{Var} \to \mathcal{P}(\mathbb{R}_{\geq 0} \times \mathcal{D}(S)) \,\}$$

be the set of all environments and ranged over by ρ. For a set $V \subseteq \mathbb{R}_{\geq 0} \times \mathcal{D}(S)$ and a variable $X \in \mathit{Var}$, we write $\rho[X \mapsto V]$ for the environment that maps X to V and Y to $\rho(Y)$ for all $Y \neq X$.

The semantics of a formula ϕ can be given as the set of configurations satisfying it. This entails a semantic functional $[\![\ ]\!] : \mathcal{L}_{\texttt{fix}} \to \mathit{Env} \to \mathcal{P}(\mathbb{R}_{\geq 0} \times \mathcal{D}(S))$ defined inductively in Figure 5. As the meaning of a closed formula ϕ does not depend on the environment, we write $[\![\phi]\!]$ for $[\![\phi]\!]_\rho$ where ρ is an arbitrary environment.

The semantics of the fixed point logic is similar to that of the modal mu-calculus [Koz83], but formulae are now satisfied by configurations. The characterisation of *greatest fixed point formula* $\mathsf{max}\; X.\phi$ follows from the well-known Knaster-Tarski fixed point theorem [Tar55].

We shall consider (closed) *equation systems* of formulae of the form

$$\begin{aligned} E : X_1 &= \phi_1 \\ &\vdots \\ X_n &= \phi_n \end{aligned}$$

$$
\begin{array}{rcl}
[\![\mathsf{tt}]\!]_\rho & = & \mathbb{R}_{\geq 0} \times \mathcal{D}(S) \\
[\![\phi_1 \wedge \phi_2]\!]_\rho & = & [\![\phi_1]\!]_\rho \cap [\![\phi_2]\!]_\rho \\
[\![\langle\alpha\rangle_v(\phi_1\, {}_p\!\oplus\, \phi_2)]\!]_\rho & = & \{\, \langle\, r, \Delta\,\rangle \in \mathbb{R}_{\geq 0} \times \mathcal{D}(S) \mid \exists \Delta' : \Delta \stackrel{\alpha}{\Longrightarrow}_w \Delta' \text{ and} \\
 & & \langle\, r + w - v, \Delta'\,\rangle = \langle\, r_1, \Delta_1\,\rangle\, {}_p\!\oplus \langle\, r_2, \Delta_2\,\rangle \text{ with } \langle\, r_i, \Delta_i\,\rangle \in [\![\phi_i]\!]_\rho \,\} \\
[\![X]\!]_\rho & = & \rho(X) \\
[\![\mathsf{max}\; X.\phi]\!]_\rho & = & \bigcup\{\, V \subseteq \mathbb{R}_{\geq 0} \times \mathcal{D}(S) \mid V \subseteq [\![\phi]\!]_{\rho[X \mapsto V]}\,\}
\end{array}
$$

Figure 5: Semantics of the fixed point logic

where $X_1, ..., X_n$ are mutually distinct variables and $\phi_1, ..., \phi_n$ are formulae having at most $X_1, ..., X_n$ as free variables. Here E can be viewed as a function $E : \mathit{Var} \to \mathcal{L}_{\mathtt{fix}}$ defined by $E(X_i) = \phi_i$ for $i = 1, ..., n$ and $E(Y) = Y$ for other variables $Y \in \mathit{Var}$.

An environment ρ is a *solution* of an equation system E if $\forall i : \rho(X_i) = [\![\phi_i]\!]_\rho$. The existence of solutions for an equation system can be seen from the following arguments. The set *Env*, which includes all candidates for solutions, together with the partial order $\leq$ defined by

$$\rho \leq \rho' \text{ iff } \forall X \in \mathit{Var} : \rho(X) \subseteq \rho'(X)$$

forms a complete lattice. The *equation functional* $\mathcal{E} : \mathit{Env} \to \mathit{Env}$ given in the λ-calculus notation by

$$\mathcal{E} := \lambda\rho.\lambda X.[\![E(X)]\!]_\rho$$

is monotonic. Thus, the Knaster-Tarski fixed point theorem guarantees existence of solutions, and the largest solution

$$\rho_E := \bigsqcup\{\, \rho \mid \rho \leq \mathcal{E}(\rho)\,\}.$$

We first observe that Proposition 3.16 can be generalised to this fixed point logic $\mathcal{L}_{\mathtt{fix}}$.

Let $f : L \to L$ be a monotonic function over a complete lattice L. For every ordinal λ define f^λ by:

- $f^0 = \top_L$, where $\top_L$ is the greatest element of the lattice
- $f^{\lambda+1} = f(f^\lambda)$
- if λ is a limit ordinal let $f^\lambda = \bigsqcap\{f^\beta \mid \beta < \lambda\}$.

Theorem 3.19 [Tarski] There exists an ordinal λ such that f^λ is the greatest fixed point of f.

A subset C of $\mathcal{C}on$ is *upper-closed (UC)* if $\langle r_\Delta, \Delta\rangle \in C$ and $\Delta \mathrel{\overline{\lhd}_r} \Theta$ implies $\langle r_\Delta + r, \Theta\rangle \in C$. An environment ρ is UC if $\rho(X)$ is UC for every variable $X \in \mathit{Var}$.

Theorem 3.20 If ρ is UC then so is $[\![\phi]\!]_\rho$ for every formula $\phi \in \mathcal{L}_{\mathtt{fix}}$.

Proof. We proceed by structural induction on the formula ϕ. The case for $\langle\alpha\rangle_r\phi'$ is similar to the proof in Proposition 3.16. All other cases are straightforward except for the greast fixed point.

Let $\phi = \mathsf{max}\; X.\phi'$. Note that by structural induction we can assume that the result holds for ϕ'. For every ordinal λ we define the set C^λ as follows:

(i) $C^0 = \mathbb{R}_{\geq 0} \times \mathcal{D}(S)$

(ii) $C^{\lambda+1} = [\![\phi']\!]_{\rho[X \mapsto C^\lambda]}$

(iii) $C^\lambda = \bigcap\{C^\beta \mid \beta < \lambda\}$ if λ is a limit ordinal.

By Tarski's theorem there is some ordinal λ such that $C^\lambda = [\![\phi]\!]_\rho$. So it is sufficient to prove, by induction over the ordinals, that C^λ is UC for every λ.

Case (i) is trivial. Case (ii) follows by structural induction, since by the inner induction the environment $\rho[x \mapsto C^\lambda]$ is UC. Case (iii) is trivial since the collection of UC sets are closed under intersection. □

Corollary 3.21 Suppose in a bounded wMDP that $\Delta \mathrel{\overline{\lhd}_r} \Theta$. Then for every closed formula $\phi \in \mathcal{L}_{\texttt{fix}}$, $\langle r_\Delta, \Delta\rangle \in [\![\phi]\!]$ implies $\langle r_\Delta + r, \Theta\rangle \in [\![\phi]\!]$.

Let $\mathcal{L}_{\texttt{fix}}(r, \Delta) = \{\, \phi \in \mathcal{L}_{\texttt{fix}} \mid \mathit{fv}(\phi) = \emptyset \wedge \langle r, \Delta\rangle \models \phi \,\}$. Then we have the extension of Corollary 3.18 from $\mathcal{L}$ to $\mathcal{L}_{\texttt{fix}}$.

Corollary 3.22 In a bounded wMDP, $\overline{s} \mathrel{\overline{\lhd}_r} \Theta$ if and only if $\mathcal{L}_{\texttt{fix}}(0, \overline{s}) \subseteq \mathcal{L}_{\texttt{fix}}(r, \Theta)$.

Proof. It follows from Corollary 3.21 and Theorem 3.17. □

Below we characterise the behaviour of a process by an equation system of modal formulae. To do so it will be convenient to use a generalised modality operator of the form $\langle\alpha\rangle_w \bigoplus_{i\in I} p_i \cdot \phi_i$ where I is a finite index set I. The satisfaction relation can be extended to these formulae so that they become derived operators in the language $\mathcal{L}_{\texttt{fix}}$, as we did in $\mathcal{L}$.

Definition 3.23 Given a bounded wMDP, its *characteristic equation system* consists of one equation for each state $s_1, ..., s_n \in S$.

$$\begin{aligned} E : X_{s_1} &= \phi_{s_1} \\ &\vdots \\ X_{s_n} &= \phi_{s_n} \end{aligned}$$

where

$$\phi_s := \bigwedge_{s \xrightarrow{\alpha}_v \Delta} \langle\alpha\rangle_v X_\Delta \tag{3}$$

with $X_\Delta := \bigoplus_{s\in\lceil\Delta\rceil} \Delta(s) \cdot X_s$. □

Theorem 3.24 Suppose E is a characteristic equation system. Then $s \lhd_r \Theta$ if and only if $\langle\, r, \Theta \,\rangle \in \rho_E(X_s)$.

Proof. ($\Leftarrow$) Let $\mathcal{R} := \{\, (s, \langle\, r, \Theta \,\rangle) \mid \langle\, r, \Theta \,\rangle \in \rho_E(X_s) \,\}$. We first show that

$$\langle\, r, \Theta \,\rangle \in [\![X_\Delta]\!]_{\rho_E} \text{ implies } \Delta \mathrel{\overline{\mathcal{R}}} \langle\, r, \Theta \,\rangle. \tag{4}$$

Let $\Delta = \bigoplus_{i\in I} p_i \cdot \overline{s_i}$, then $X_\Delta = \bigoplus_{i\in I} p_i \cdot X_{s_i}$. Suppose $\langle r, \Theta \rangle \in [\![X_\Delta]\!]_{\rho_E}$. We have that $\langle r, \Theta \rangle = \sum_{i\in I} p_i \cdot \langle r_i, \Theta_i \rangle$ and, for all $i \in I$, $\langle r_i, \Theta_i \rangle \in [\![X_{s_i}]\!]_{\rho_E}$, i.e. $s_i \mathcal{R} \langle r_i, \Theta_i \rangle$. It follows that $\Delta \overline{\mathcal{R}} \langle r, \Theta \rangle$.

Now we show that $\mathcal{R}$ is an amortised weighted simulation. Suppose $s \mathcal{R} \langle r, \Theta \rangle$ and $s \xrightarrow{\alpha}_v \Delta$. Then $\langle r, \Theta \rangle \in \rho_E(X_s) = [\![\phi_s]\!]_{\rho_E}$. It follows from (3) that $\langle r, \Theta \rangle \in [\![\langle\alpha\rangle_v X_\Delta]\!]_{\rho_E}$. So there exists some Θ' such that $\Theta \stackrel{\alpha}{\Longrightarrow}_w \Theta'$ and $\langle r + w - v, \Theta' \rangle \in [\![X_\Delta]\!]_{\rho_E}$. Now we apply (4).

($\Rightarrow$) We define the environment ρ by

$$\rho(X_s) := \{ \langle r, \Theta \rangle \mid s \lhd_r \Theta \}.$$

It suffices to show that ρ is a post-fixed point of $\mathcal{E}$, i.e.

$$\rho \leq \mathcal{E}(\rho) \tag{5}$$

because in that case we have $\rho \leq \rho_E$, thus $s \lhd \langle r, \Theta \rangle$ implies $\langle r, \Theta \rangle \in \rho(X_s)$ which in turn implies $\langle r, \Theta \rangle \in \rho_E(X_s)$.

We first show that

$$\Delta \mathrel{\overline{\lhd}} \langle r, \Theta \rangle \text{ implies } \langle r, \Theta \rangle \in [\![X_\Delta]\!]_\rho. \tag{6}$$

Suppose $\Delta \mathrel{\overline{\lhd}} \langle r, \Theta \rangle$. Then we have that (i) $\Delta = \sum_{i\in I} p_i \cdot \overline{s_i}$, (ii) $\langle r, \Theta \rangle = \sum_{i\in I} p_i \cdot \langle r_i, \Theta_i \rangle$, (iii) $s_i \lhd \langle r_i, \Theta_i \rangle$ for all $i \in I$. We know from (iii) that $\langle r_i, \Theta_i \rangle \in [\![X_{s_i}]\!]_\rho$. Using (ii) we have that $\langle r, \Theta \rangle \in [\![\bigoplus_{i\in I} p_i \cdot X_{s_i}]\!]_\rho$. Using (i) we obtain $\langle r, \Theta \rangle \in [\![X_\Delta]\!]_\rho$.

Now we are in a position to show (5). Suppose $\langle r, \Theta \rangle \in \rho(X_s)$. We must prove that $\langle r, \Theta \rangle \in [\![\phi_s]\!]_\rho$, i.e.

$$\langle r, \Theta \rangle \in \bigcap_{s \xrightarrow{\alpha}_v \Delta} [\![\langle\alpha\rangle_v X_\Delta]\!]_\rho$$

by (3).

We assume that $s \xrightarrow{\alpha}_v \Delta$. Since $s \lhd_r \Theta$, there exists some Θ' such that $\Theta \stackrel{\alpha}{\Longrightarrow}_w \Theta'$ and $\Delta \mathrel{\overline{\lhd}} \langle r + w - v, \Theta' \rangle$. By (6), we get $\langle r + w - v, \Theta' \rangle \in [\![X_\Delta]\!]_\rho$. It follows that $\langle r, \Theta \rangle \in [\![\langle a\rangle_v X_\Delta]\!]_\rho$.

□

So far we know how to construct the characteristic equation system for a bounded wMDP. As introduced in [MO98], the three transformation rules in Figure 6 can be used to obtain from an equation system E a formula whose interpretation coincides with the interpretation of X_1 in the greatest solution of E. The formula thus obtained from a characteristic equation system is called a *characteristic formula*.

Theorem 3.25 Given a characteristic equation system E, there is a characteristic formula ϕ_s such that $\rho_E(X_s) = [\![\phi_s]\!]$ for any state s. □

The above theorem, together with Theorem 3.24, gives rise to the following corollary.

Corollary 3.26 For each state s in a bounded wMDP, there is a characteristic formula ϕ_s such that $s \lhd \langle r, \Theta \rangle$ iff $\langle r, \Theta \rangle \in [\![\phi_s]\!]$. □

1. Rule 1: $E \to F$
2. Rule 2: $E \to G$
3. Rule 3: $E \to H$ if $X_n \notin fv(\phi_1, ..., \phi_n)$

$$
\begin{array}{rclrclrclrcl}
E : X_1 & = & \phi_1 & F : X_1 & = & \phi_1 & G : X_1 & = & \phi_1[\phi_n/X_n] & H : X_1 & = & \phi_1 \\
 & \vdots & & & \vdots & & & \vdots & & & \vdots & \\
X_{n-1} & = & \phi_{n-1} & X_{n-1} & = & \phi_{n-1} & X_{n-1} & = & \phi_{n-1}[\phi_n/X_n] & X_{n-1} & = & \phi_{n-1} \\
X_n & = & \phi_n & X_n & = & \mathsf{max}\, X_n.\phi_n & X_n & = & \phi_n & & &
\end{array}
$$

Figure 6: Transformation rules

4 Testing

4.1 Benefits testing

Standard theories of testing involve the idea of applying tests to processes and seeing if the result is a *success*. With the presence of weights on wMDPs we have a more elementary way of testing; we run a test in parallel with the process being tested and calculate the possible benefits which can be accrued. Then two wMDPs can be compared via the resulting sets of possible benefits.

Definition 4.1 A wMDP of the form $\langle S, \{\tau\}, W, \longrightarrow\rangle$ is referred to as a *(weighted) computation structure.* □

An arbitrary wMDP can be viewed as a weighted computation structure by ignoring all the actions $s \xrightarrow{a}_w \Delta$ other than $s \xrightarrow{\tau}_w \Delta$; indeed weighted computation structures correspond more or less directly with the more standard notion of *Markov decision processes.* Here we are interested in the computation structures generated by wMDPs of the form

$$[\![P]\!] \mid\mid [\![T]\!]$$

where P is a wMDP which we wish to investigate and T is a finite wMDP, representing the investigation. The question now is how do we associate a set of possible rewards with a distribution over the set of states of a weighted computation structure?

Consider the simple fully probabilistic wMDP in Figure 7(a), which results from running the test $T = \overline{\mathsf{up}_1}.\overline{\mathsf{down}_4}.\mathbf{0}$ in parallel with the system s_1 from the Introduction. Formally this is the sub-wMDP of the wMDP $(\overline{s_1} \mid T)$ obtained by concentrating on the internal actions τ_w, which is just the wMDP represented by $(\overline{s_1} \mid T)\backslash\mathsf{Act}$ that we denote by $\overline{s_1} \mid\mid T$. Every time the experiment runs we get the initial benefit 3; three-quarters of the time we also get the benefit 7 while a quarter of time we get 5. So the total benefit is

$$3 + \frac{3}{4} \cdot 7 + \frac{1}{4} \cdot 5 = 9.5.$$

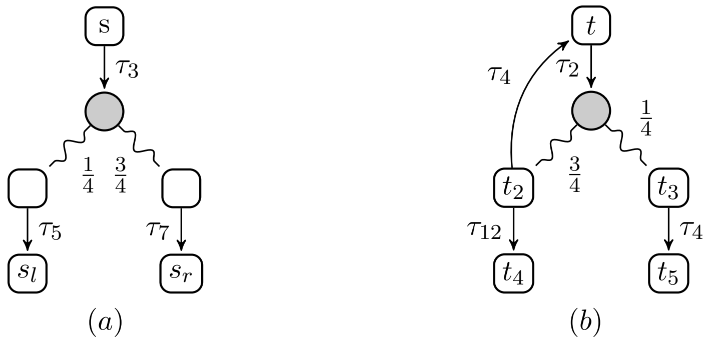

Figure 7: Testing systems

In the presence of nondeterminism there will in general be a set of possible benefits, depending on the way in which the nondeterminism is resolved. Traditionally this resolution is expressed in terms of a scheduler, or adversary, which for each state decides which of its successors is chosen for execution, with the resulting set of benefits consequently depending on the choice of scheduler. Here we take a more abstract approach, following [DvGHM09], and essentially allow arbitrary schedulers.

Definition 4.2 [Extreme derivatives] For any Δ in a computation structure we write $\Delta \Longrightarrow\!\!\succ_w \Phi$ if

- $\Delta \stackrel{\tau}{\Longrightarrow}_w \Phi$, that is Φ is a hyper-derivative of Δ
- Φ is *stable*, that is $s \stackrel{\tau}{\not\rightarrow}$ for every s in $\lceil \Phi \rceil$.

We say Φ is an *extreme derivative* of Δ, with weight w. □

Intuitively every extreme derivation $\Delta \Longrightarrow\!\!\succ_w \Phi$ represents a computation from the initial distribution Δ guided by some implicit scheduler. For example, consider the hyper-derivation:

$$
\begin{array}{ccl}
\Delta & = & \Delta_0^{\rightarrow} + \Delta_0^{\times} \\
\Delta_0^{\rightarrow} & \stackrel{\tau}{\longrightarrow}_{w_0} & \Delta_1^{\rightarrow} + \Delta_1^{\times} \\
 & \vdots & \\
\Delta_k^{\rightarrow} & \stackrel{\tau}{\longrightarrow}_{w_k} & \Delta_{k+1}^{\rightarrow} + \Delta_{k+1}^{\times} \\
 & \vdots & \\
 & \Phi & = \sum_{k=0}^{\infty} \Delta_k^{\times}
\end{array}
\tag{7}
$$

where $w = \sum_{k \geq 0} w_k$. Initially, since $\Delta_0^{\times}$ is stable, $\Delta_0^{\rightarrow}$ contains (in its support) all states which can proceed with the computation. The implicit scheduler decides for each of these states which step to take, cumulating in the first move, $\Delta_0^{\rightarrow} \stackrel{\tau}{\longrightarrow}_{w_0} \Delta_1^{\rightarrow} + \Delta_1^{\times}$. At an arbitrary stage $\Delta_k^{\rightarrow}$ contains all states which can continue; the scheduler decides which step to take for each individual state and the overall result of the schedulers decision for this stage is captured in the step $\Delta_k^{\rightarrow} \stackrel{\tau}{\longrightarrow}_{w_k} \Delta_{k+1}^{\rightarrow} + \Delta_{k+1}^{\times}$.

Example 4.3 Referring to Figure 7(a) it is easy to see that $\overline{s}$ has a unique (degenerate) extreme derivative, $\overline{s_1} \Longrightarrow\!\!\!\!\succ_{9.5} (\frac{1}{4}\overline{s_l} + \frac{3}{4}\overline{s_r})$, intuitively representing the unique weighted computation from $\overline{s_1}$. However, consider the wMDP in Figure 7(b), in which there is a nondeterministic choice from state t_2; here the extreme derivatives generated from $\overline{t}$, and their associated weights, will depend on the choices made during the computation by the implicit scheduler.

First suppose that the scheduler uses the static policy which maps t_2 to $\langle 12, \overline{t_4} \rangle$. Then it is easy to see that the generated extreme derivative, which is degenerate, is $\overline{t} \Longrightarrow\!\!\!\!\succ_{12} (\frac{3}{4}\overline{t_4} + \frac{1}{4}\overline{t_5})$. However using the static policy which maps t_2 to $\langle 4, \overline{t} \rangle$ we generate, using (7), a non-degenerate extreme derivative; after some calculations this can be seen to be $\overline{t_1} \Longrightarrow\!\!\!\!\succ_{24} \overline{t_5}$.

However there are many other possible implicit schedulers, for example at different times in the computations employing either of these static policies, or even choosing nondeterministically between them. But these are the only static policies and therefore we know from Theorem 2.20 that if $\overline{t_1} \Longrightarrow\!\!\!\!\succ_w \Delta$ then w must take the form $p \cdot 12 + (1-p) \cdot 24$ for some $0 \leq p \leq 1$. That is the set of benefits which can be generated from $\overline{t_1}$ is $\{ 24 - 12 \cdot p \mid 0 \leq p \leq 1 \}$. □

Definition 4.4 In a wMDP, for any $\Delta \in \mathcal{D}(S)$, let

$$\mathbf{Benefits}(\Delta) = \{ w \in W \mid \Delta \Longrightarrow\!\!\!\!\succ_w \Phi, \text{ for some } \Phi \in \mathcal{D}_{sub}(S) \}.$$

□

Note that in general $\mathbf{Benefits}(\Delta)$ may contain ∞, although by Theorem 2.27 this cannot be the case if the wMDP is bounded.

We compare Benefit sets as follows:

$$B_1 \leq^r_{\text{Ho}} B_2 \text{ if for every } r_1 \in B_1 \text{ there exists some } r_2 \in B_2 \text{ such that } r_1 \leq r + r_2.$$

Definition 4.5 [May testing] For any two distributions Δ, Θ we write $\Delta \sqsubseteq^r_{\text{may}} \Theta$ if for every finite (testing) process T, $\mathbf{Benefits}(\Delta \mid\mid T) \leq^r_{\text{Ho}} \mathbf{Benefits}(\Theta \mid\mid T)$. We write $\Delta \sqsubseteq_{\text{may}} \Theta$ to mean that there is some $r \in \mathbb{R}_{\geq 0}$ such that $\Delta \sqsubseteq^r_{\text{may}} \Theta$. □

This interpretation of processes is inherently optimistic; $\Delta \sqsubseteq^r_{\text{may}} \Theta$ means that, given the investment r, every possible benefit produced by Δ can in principle be improved upon by Θ. Note that if we confine ourselves to bounded wMDPs then by Theorem 2.27 and Theorem 2.30 no benefits set used in this definition will contain ∞.

Our first result shows that simulations can be used as a sound proof technique for this semantics. In order to prove that result, we need the following technical lemmas.

Lemma 4.6 Let Δ, Θ be two distributions in a bounded wMDP. Suppose $\Delta \mathrel{\overline{\lhd}_r} \Theta$ for some $r \in \mathbb{R}_{\geq 0}$. If $\Delta \stackrel{\tau}{\Longrightarrow}_v \varepsilon$ then $\Theta \stackrel{\tau}{\Longrightarrow}_w \Theta'$ for some Θ' such that $r + w - v \geq 0$.

Proof. If $\Delta \stackrel{\tau}{\Longrightarrow}_v \varepsilon$ then there is a sequence of τ transitions

$$\Delta \stackrel{\tau}{\longrightarrow}_{v_1} \Delta_1 \stackrel{\tau}{\longrightarrow}_{v_2} \Delta_2 \stackrel{\tau}{\longrightarrow}_{v_3} \cdots$$

such that $\sum_{k\geq 1} v_k = v$. Since $\Delta \mathrel{\overline{\lhd}_r} \Theta$, it can be shown by induction on i that there are weights w_i and subdistributions Θ_i with

$$\begin{array}{l} \Theta \stackrel{\tau}{\Longrightarrow}_{(\sum_{1\leq k\leq i} w_k)} \Theta_i \\ \Delta_i \mathrel{\overline{\lhd}_{(r+\sum_{1\leq k\leq i} w_k - \sum_{1\leq k\leq i} v_k)}} \Theta_i \end{array}$$

for all $i \geq 1$. The compactness arguments in Appendix C (Corollary C.1) ensures that the set $\{\langle w', \Theta' \rangle \mid \Theta \stackrel{\tau}{\Longrightarrow}_{w'} \Theta'\}$ is closed. As the sequence $\{\sum_{1\leq k\leq i} w_k\}_{i=1}^{\infty}$ has limit $\sum_{k\geq 1} w_k$, there exists some subdistribution Θ' such that $\Theta \stackrel{\tau}{\Longrightarrow}_{(\sum_{k\geq 1} w_k)} \Theta'$. Since for each $i \geq 1$, we have that $r + \sum_{1\leq k\leq i} w_k - \sum_{1\leq k\leq i} v_k \geq 0$. It follows that $r + \sum_{k\geq 1} w_k - \sum_{k\geq 1} v_k \geq 0$. □

Lemma 4.7 Let Δ, Θ be two distributions in a bounded computation structure. If $\Delta \mathrel{\overline{\lhd}_r} \Theta$ then $\mathbf{Benefits}(\Delta) \leq_{\mathrm{Ho}}^{r} \mathbf{Benefits}(\Theta)$.

Proof. For any $v \in \mathbf{Benefits}(\Delta)$, there is some subdistribution Δ' such that $\Delta \Longrightarrow\!\!\!\!\succ_v \Delta'$. By Corollary 2.28 there is some subdistribution Δ'_ε such that $\Delta \stackrel{\tau}{\Longrightarrow}_{v_1} (\Delta' + \Delta'_\varepsilon)$, $|\Delta| = |\Delta' + \Delta'_\varepsilon|$, $\Delta'_\varepsilon \stackrel{\tau}{\Longrightarrow}_{v_2} \varepsilon$ and $v_1 + v_2 = v$. By Corollary 3.8 there is some Θ'' such that $\Theta \stackrel{\tau}{\Longrightarrow}_{w_1} \Theta''$ and $(\Delta' + \Delta'_\varepsilon) \mathrel{\overline{\lhd}_{r+w_1-v_1}} \Theta''$. By Proposition 2.6 we can decompose Θ'' such that $\Theta'' = \Theta' + \Theta'_\varepsilon$, $\Delta' \mathrel{\overline{\lhd}_{r_1}} \Theta'$, $\Delta'_\varepsilon \mathrel{\overline{\lhd}_{r_2}} \Theta'_\varepsilon$, and

$$r_1 + r_2 = r + w_1 - v_1. \tag{8}$$

By Lemma 4.6 there is some Θ''_ε such that $\Theta'_\varepsilon \stackrel{\tau}{\Longrightarrow}_{w_2} \Theta''_\varepsilon$ and

$$r_2 + w_2 - v_2 \geq 0. \tag{9}$$

By the transitivity of hyper-derivations, Theorem 2.13, we obtain that $\Theta \stackrel{\tau}{\Longrightarrow}_{w_1+w_2} \Theta' + \Theta''_\varepsilon$. It follows that there is some extreme derivation $\Theta \Longrightarrow\!\!\!\!\succ_w \Theta'''$ for some w, Θ''' with

$$w \geq w_1 + w_2. \tag{10}$$

By (8), (9) and (10) we derive that

$$w \geq (r_1 + r_2 - r + v_1) + (v_2 - r_2) = v - r + r_1 \geq v - r.$$

Therefore, we have found some $w \in \mathbf{Benefits}(\Theta)$ with $v \leq r + w$. Since this holds for any $v \in \mathbf{Benefits}(\Delta)$, we have that $\mathbf{Benefits}(\Delta) \leq_{\mathrm{Ho}}^{r} \mathbf{Benefits}(\Theta)$. □

Theorem 4.8 [Soundness] In a bounded wMDP, $P \mathrel{\overline{\lhd}_r} Q$ implies $P \sqsubseteq_{\mathrm{may}}^{r} Q$.

Proof. For any finite test T, we can infer that

$$\begin{array}{lll} & P \mathrel{\overline{\lhd}_r} Q & \\ \Rightarrow & (P \mid\mid T) \mathrel{\overline{\lhd}_r} (Q \mid\mid T) & \text{by Theorem 3.4} \\ \Rightarrow & \mathbf{Benefits}(P \mid\mid T) \leq_{\mathrm{Ho}}^{r} \mathbf{Benefits}(Q \mid\mid T) & \text{by Lemma 4.7} \\ \Leftrightarrow & P \sqsubseteq_{\mathrm{may}}^{r} Q & \text{by definition} \end{array}$$

□

In the next section we will see a partial converse to this result, in Corollary 4.13.

4.2 Success based testing

We follow our earlier approach [DvGHM09] of testing nondeterministic and probabilistic processes. A *test* is simply a process from the language CCMDP except that it may use special actions for reporting success. Thus we assume a countable set Ω of fresh success actions not already in Act_τ; intuitively each ω in Ω can be viewed as a particular way in which success can be achieved. We call CCMDP$^\Omega$ the language CCMDP extended with the new actions in Ω. Its operational semantics is as in Figure 4 except that the rules (L-ALT) and (L-PAR) are modified as follows, where α ranges over Act_τ.

$$\text{(L-ALT1)}\quad \frac{P_1 \xrightarrow{\alpha}_w Q \qquad P_2 \overset{\omega}{\nrightarrow} \text{ for all } \omega \in \Omega}{P_1 + P_2 \xrightarrow{\alpha}_w Q}$$

$$\text{(L-ALT2)}\quad \frac{P_1 \xrightarrow{\omega}_w Q \qquad P_2 \overset{\omega'}{\nrightarrow} \text{ for all } \omega' \in \Omega \backslash \{\omega\}}{P_1 + P_2 \xrightarrow{\omega}_w Q}$$

$$\text{(L-PAR1)}\quad \frac{P_1 \xrightarrow{\alpha}_w Q \qquad P_2 \overset{\omega}{\nrightarrow} \text{ for all } \omega \in \Omega}{P_1 \mid P_2 \xrightarrow{\alpha}_w Q \mid P_2}$$

$$\text{(L-PAR2)}\quad \frac{P_1 \xrightarrow{\omega}_w Q \qquad P_2 \overset{\omega'}{\nrightarrow} \text{ for all } \omega' \in \Omega \backslash \{\omega\}}{P_1 \mid P_2 \xrightarrow{\omega}_w Q \mid P_2}$$

These rules guarantee that if a process P can report success via action ω, i.e. $P \xrightarrow{\omega}_w \Delta$ for some w and Δ, then no other actions are enabled at P – neither a normal action in Act_τ nor another success action in Ω is allowed. For this reason, we say that the wMDPs generated by the processes in CCMDP$^\Omega$ are *ω-respecting.*

Definition 4.9 Let $\Phi \in \mathcal{D}_{sub}(S)$, we write $\mathbf{Success}(\Phi)$ for the function (viewed as a vector) in $[0,1]^\Omega$ such that $\mathbf{Success}(\Phi)(\omega) = \sum\{\Phi(s) \mid s \in \lceil\Phi\rceil \text{ and } s \xrightarrow{\omega}\}$. We let

$$\mathbf{Outcomes}(\Delta) = \{\langle\, w, \mathbf{Success}(\Phi)\,\rangle \mid \Delta \Longrightarrow\!\!\!\!\!\succ_w \Phi \text{ for some } \Phi \in \mathcal{D}_{sub}(S)\}$$

□

Thus, intuitively, $\mathbf{Outcomes}(\Delta)$ tabulates the rewards associated with vectors of successes, each particular vector obtained by an execution to completion of Δ.

Let $B_1, B_2 \in \mathbb{R}_{\geq 0} \times [0,1]^\Omega$. We write $B_1 \leq^r_{\mathrm{Ho}} B_2$ if for each $\langle\, r_1, f_1\,\rangle \in B_1$ there exists some $\langle\, r_2, f_2\,\rangle \in B_2$ such that $r_1 \leq r + r_2$ and $f_1(\omega) \leq f_2(\omega)$ for all $\omega \in \Omega$.

Definition 4.10 [Multi-success testing] For any two processes P, Q we write $P \sqsubseteq^r_{\mathrm{mmay}} Q$ if for every finite (testing) process T, $\mathbf{Outcomes}(P \mid\mid T) \leq^r_{\mathrm{Ho}} \mathbf{Outcomes}(Q \mid\mid T)$. □

Theorem 4.11 [Multi-success testing coincides with benefits testing] For any $r \in \mathbb{R}_{\geq 0}$ and two processes P, Q whose operational semantics only give rise to bounded wMDPs,

$$P \sqsubseteq^r_{\mathrm{mmay}} Q \ \text{ iff } \ P \sqsubseteq^r_{\mathrm{may}} Q.$$

Proof. The general schema of the proof follows from [DvGMZ07] where it is shown that multi-success testing coincides with uni-success testing for finitary probabilistic automat.

We first define the function **Outcomes**$'$ which is the same as **Outcomes** except that we allow any derivation instead of just extreme derivations.

$$\mathbf{Outcomes}'(\Delta) = \{\langle\, w, \mathbf{Success}(\Phi)\,\rangle \mid \Delta \stackrel{\tau}{\Longrightarrow}_w \Phi \text{ for some } \Phi \in \mathcal{D}_{sub}(S)\}$$

We claim that **Outcomes**$'$ satisfies the next two properties.

1. For any $\Delta \in \mathcal{D}_{sub}(S)$, we have $\mathbf{Outcomes}(\Delta) \leq^0_{\mathrm{Ho}} \mathbf{Outcomes}'(\Delta)$ and also conversely $\mathbf{Outcomes}'(\Delta) \leq^0_{\mathrm{Ho}} \mathbf{Outcomes}(\Delta)$.
2. For any $\Delta \in \mathcal{D}_{sub}(S)$ in a bounded wMDP, the set $\mathbf{Outcomes}'(\Delta)$ is compact and convex.

For the first claim, we observe that $\mathbf{Outcomes}(\Delta) \subseteq \mathbf{Outcomes}'(\Delta)$ from which it follows that $\mathbf{Outcomes}(\Delta) \leq^0_{\mathrm{Ho}} \mathbf{Outcomes}'(\Delta)$. Since the wMDPs that we are considering are "ω-respecting", we have that if state s can enable a τ-action then $\mathbf{Success}(\overline{s}) = \vec{0}$ where $\vec{0}$ is the empty vector with $\vec{0}(\omega) = 0$ for all $\omega \in \Omega$. It follows that $\Delta \stackrel{\tau}{\Longrightarrow}_r \Delta'$ implies $\mathbf{Success}(\Delta) \leq \mathbf{Success}(\Delta')$. So if $\Delta \stackrel{\tau}{\Longrightarrow}_{r_1} \Phi$ then $\Phi \stackrel{\tau}{\Longrightarrow}_{r_2} \Phi'$ for some extreme derivation Φ', i.e. $\Delta \Longrightarrow\!\!\!\succ_{r_1+r_2} \Phi'$, such that $\mathbf{Success}(\Phi) \leq \mathbf{Success}(\Phi')$. Hence, it is easy to show that $\mathbf{Outcomes}'(\Delta) \leq^0_{\mathrm{Ho}} \mathbf{Outcomes}(\Delta)$.

For the second claim, we use the fact that the function **Success** is continuous. Let $F_{\mathbf{Success}}$ be the function given by

$$F_{\mathbf{Success}}(w, \Phi) = \langle\, w, \mathbf{Success}(\Phi)\,\rangle$$

which is also continuous. Again we appeal to the arguments in Appendix C (specifically Corollary C.1) which guarantees that the set $\{\langle\, w, \Phi\,\rangle \mid \Delta \stackrel{\tau}{\Longrightarrow}_w \Phi \text{ for some } \Phi \in \mathcal{D}_{sub}(S)\}$ is compact and convex. Its image under $F_{\mathbf{Success}}$, i.e. $\mathbf{Outcomes}'(\Delta)$, is also compact and easily seen to be convex.

With these two properties at hand, we are ready to prove that $P \sqsubseteq^r_{\mathrm{mmay}} Q$ iff $P \sqsubseteq^r_{\mathrm{may}} Q$. The *only if* direction is straightforward, so we focus on the *if* direction. We prove it by contradiction. Suppose that $P \sqsubseteq^r_{\mathrm{may}} Q$ but $P \not\sqsubseteq^r_{\mathrm{mmay}} Q$. Then there is some multi-success test T such that $\mathbf{Outcomes}(P \mid\mid T) \not\leq^r_{\mathrm{Ho}} \mathbf{Outcomes}(Q \mid\mid T)$. From claim (1) above, we have that

$$\mathbf{Outcomes}'(P \mid\mid T) \not\leq^r_{\mathrm{Ho}} \mathbf{Outcomes}'(Q \mid\mid T).$$

Let m be the number of different success actions appearing in T. There is some vector $\langle\, v, p_1, ..., p_m\,\rangle$ in $\mathbf{Outcomes}'(P \mid\mid T)$ such that $\langle\, v, p_1, ..., p_m\,\rangle \not\leq \langle\, w+r, q_1, ..., q_m\,\rangle$ for all vectors $\langle\, w, q_1, ..., q_m\,\rangle$ in $\mathbf{Outcomes}'(Q \mid\mid T)$. Let O_1 and O_2 be the two sets defined as follows.

$$\begin{array}{rcl} O_1 & = & \{\langle\, v', p'_1, ..., p'_m\,\rangle \in \mathbb{R}_{\geq 0} \times [0,1]^m \mid \langle\, v, p_1, ..., p_m\,\rangle \leq \langle\, v', p'_1, ..., p'_m\,\rangle\} \\ O_2 & = & \{\langle\, w+r, q_1, ..., q_m\,\rangle \mid \langle\, w, q_1, ..., q_m\,\rangle \in \mathbf{Outcomes}'(Q \mid\mid T)\} \end{array}$$

It is obvious that O_1 is closed and convex. Using claim (2) above, we know that O_2 is compact and convex. Clearly, O_1 and O_2 are disjoint. By the Hyperplane separation theorem, Theorem 1.2.4 in

[Mat02], we can separate O_1 from O_2 by a hyperplane whose normal is $\langle\, h_0, h_1, ..., h_m \,\rangle$. That is, there is some $c \in \mathbb{R}$ such that, without loss of generality,

$$h_0 v' + \sum_{i=1}^{m} h_i p'_i \;>\; c \;>\; h_0(w+r) + \sum_{i=1}^{m} h_i q_i \tag{11}$$

for all $\langle\, v', p'_1, ..., p'_m \,\rangle \in O_1$ and $\langle\, w+r, q_1, ..., q_m \,\rangle \in O_2$.

We now argue that each h_i, for $0 \leq i \leq m$, is non-negative. Assume for a contradiction that $h_i < 0$. Choose some $d > 0$ large enough so that the vector $\langle\, v', ..., p'_i + d, ..., p'_m \,\rangle$ is still in O_1 but $h_0 v' + h_i(p'_i + d) + \sum\{h_j p'_j \mid 1 \leq j \leq m \text{ but } j \neq i\} \;<\; c$. This would contradict the separation.

Then we distinguish two cases.

- $h_0 = 0$. Then (11) can be simplified to

$$\sum_{i=1}^{m} h_i p'_i \;>\; c \;>\; \sum_{i=1}^{m} h_i q_i. \tag{12}$$

  Since O_2 is compact, i.e. closed and bounded, we can let

$$\begin{array}{rcl} c' & = & \max\{\sum_{i=1}^{m} h_i q_i \mid \langle\, w+r, q_1, ..., q_m \,\rangle \in O_2\} \\ w' & = & \max\{w + r \mid \langle\, w+r, q_1, ..., q_m \,\rangle \in O_2\}. \end{array}$$

  Note that we have $c > c'$. Let e be any real number such that $e > \frac{w'}{c-c'}$. We infer that

$$\begin{array}{rcl} v' + \sum_{i=1}^{m} h_i e p'_i & \geq & e \sum_{i=1}^{m} h_i p'_i \\ & > & ec \\ & > & w' + ec' \\ & \geq & (w+r) + e \sum_{i=1}^{m} h_i q_i \\ & = & (w+r) + \sum_{i=1}^{m} h_i e q_i \end{array}$$

  for any $\langle\, v', p'_1, ..., p'_m \,\rangle \in O_1$ and $\langle\, w+r, q_1, ..., q_m \,\rangle \in O_2$. This means that O_1 can also be separated from O_2 by a hyperplane with normal $\langle\, 1, h_1 e, ..., h_m e \,\rangle$.

  We now construct a benefits test T' from the multi-success test T by letting

$$T' \;=\; T \mid\mid (\omega_{1}0.\tau_{h_1 e}.\mathbf{0} + \cdots + \omega_{m}0.\tau_{h_m e}.\mathbf{0})$$

  In T' an occurrence of ω_i yields weight 0 but it is followed by a tau move which yields weight $h_i e$. If $\langle\, v, p_1, ..., p_m \,\rangle$ is an outcome of testing P with T, then $v + \sum_{i=1}^{m} h_i e p_i$ is an outcome of testing P with T'. Testing Q with T' is similar. The above separation shows that P and Q can be distinguished by the benefits test T' because

$$\mathbf{Benefits}(P \mid\mid T') \not\leq^{r}_{\mathrm{Ho}} \mathbf{Benefits}(Q \mid\mid T')$$

  which contradicts the assumption that $P \sqsubseteq^{r}_{\mathrm{may}} Q$.

- $h_0 > 0$. It follows from (11) that

$$v' + \sum_{i=1}^{m} \frac{h_i}{h_0} p'_i \;>\; \frac{c}{h_0} \;>\; (w+r) + \sum_{i=1}^{m} \frac{h_i}{h_0} q_i \tag{13}$$

  for all $\langle v', p'_1, ..., p'_m \rangle \in O_1$ and $\langle w+r, q_1, ..., q_m \rangle \in O_2$. This means that O_1 can also be separated from O_2 by a hyperplane with normal $\langle 1, \frac{h_1}{h_0}, ..., \frac{h_m}{h_0} \rangle$. Similar to the last case, we construct a benefits test T' from the multi-success test T by letting

$$T' \;=\; T \,||\, (\omega_{10}.\tau_{\frac{h_1}{h_0}}.\mathbf{0} + \cdots + \omega_{m0}.\tau_{\frac{h_m}{h_0}}.\mathbf{0})$$

  and it can be seen that P and Q are distinguished by the benefits test T'.

Thus in both cases we obtain $P \not\sqsubseteq^r_{\text{may}} Q$, a contradiction to our original assumption. □

One consequence of this result is that we can show that benefits testing is complete for amortised simulations. This is achieved by using multi-success testing as an intermediary:

Theorem 4.12 In a bounded wMDP, if $\Delta \sqsubseteq^r_{\text{mmay}} \Theta$ then there exists some r' such that $r' \geq r$ and $\mathcal{L}(0, \Delta) \subseteq \mathcal{L}(r', \Theta)$.

Proof. The proof relies on designing, for each formula ϕ, a characteristic test T_ϕ; that is satisfying the formula ϕ coincides with passing the corresponding test T_ϕ, relative to a *target value*. The construction of the tests is quite complex; however the details are quite similar to those used in the corresponding result in [DvGHM09] and are therefore relegated to Appendix E. □

Corollary 4.13 [Completeness] In a bounded wMDP, if $\overline{s} \sqsubseteq_{\text{may}} \Theta$ then $\overline{s} \sqsubseteq_{sim} \Theta$.

Proof. By combining Theorems 4.11, 4.12 and Corollary 3.18, we can show that $\overline{s} \sqsubseteq^r_{\text{may}} \Theta$ implies the existence of some compensation $r' \geq r$ such that $\overline{s} \mathrel{\overline{\lhd}_{r'}} \Theta$, from which the required result follows. □

It is tempting to sharpen the above property to state that in a bounded wMDP $\Delta \sqsubseteq^r_{\text{may}} \Theta$ implies $\Delta \mathrel{\overline{\lhd}_r} \Theta$. Unfortunately, this would not be a valid statement, as demonstrated by the following example.

Example 4.14 Consider the two distributions $\Delta := \mathbf{0}\ {}_{\frac{1}{2}}\!\oplus a_1.\mathbf{0}$ and $\Theta := \tau_2.\mathbf{0}\ {}_{\frac{1}{2}}\!\oplus a_0.\mathbf{0}$. It is easy to see that $\Delta \mathrel{\overline{\not\lhd}_0} \Theta$ because there is no way to decompose Θ into $\Theta_1\ {}_{\frac{1}{2}}\!\oplus \Theta_2$ for some Θ_1, Θ_2 such that $a_1.\mathbf{0} \lhd_0 \Theta_2$. However, one can show that $\Delta \sqsubseteq^0_{\text{may}} \Theta$. This follows from the observations below:

(i) For all weight w and test T, $\mathbf{Benefits}(\tau_w.\mathbf{0} \,||\, T) \;=\; \{v + w \mid v \in \mathbf{Benefits}(\mathbf{0} \,||\, T)\}$.

(ii) For all weight w and test T, $\mathbf{Benefits}(a_w.\mathbf{0} \,||\, T) \;\leq^w_{\text{Ho}}\; \mathbf{Benefits}(a_0.\mathbf{0} \,||\, T)$.

Both assertions can be proved by structural induction on T.

Now suppose $w \in \mathbf{Benefits}(\Delta \mid\mid T)$ for an arbitrary test T. There is some stable derivative Γ such that $\Delta \mid\mid T \stackrel{\tau}{\Longrightarrow}_w \Gamma$. By Proposition 2.11(3) there are some $w_1, w_2, \Gamma_1, \Gamma_2$ with $\mathbf{0} \mid\mid T \stackrel{\tau}{\Longrightarrow}_{w_1} \Gamma_1$, $a_1.\mathbf{0} \mid\mid T \stackrel{\tau}{\Longrightarrow}_{w_2} \Gamma_2$, $w = \frac{1}{2}w_1 + \frac{1}{2}w_2$, and $\Gamma = \frac{1}{2} \cdot \Gamma_1 + \frac{1}{2} \cdot \Gamma_2$, where both Γ_1 and Γ_2 are stable. In other words, $w_1 \in \mathbf{Benefits}(\mathbf{0} \mid\mid T)$ and $w_2 \in \mathbf{Benefits}(a_1.\mathbf{0} \mid\mid T)$. By (i) above, $w_1 + 2 \in \mathbf{Benefits}(\tau_2.\mathbf{0} \mid\mid T)$; by (ii) above, there exists some $w_2' \in \mathbf{Benefits}(a_0.\mathbf{0} \mid\mid T)$ with $w_2 \leq w_2' + 1$. Thus, we can infer that

$$\begin{array}{rcl} w & = & \frac{1}{2}w_1 + \frac{1}{2}w_2 \\ & < & \frac{1}{2}(w_1+2) + \frac{1}{2}(w_2 - 1) \\ & \leq & \frac{1}{2}(w_1+2) + \frac{1}{2}w_2' \end{array}$$

Using Proposition 2.11(4), it can be seen that $\frac{1}{2}(w_1+2) + \frac{1}{2}w_2' \in \mathbf{Benefits}(\Theta \mid\mid T)$. Therefore, we have $\mathbf{Benefits}(\Delta \mid\mid T) \leq^0_{\mathrm{Ho}} \mathbf{Benefits}(\Theta \mid\mid T)$. Since this reasoning is carried out for an arbitrary test T, it follows that $\Delta \sqsubseteq^0_{\mathrm{may}} \Theta$.

□

4.3 Expected benefits testing

The testing approach introduced in the previous two sections can be called *total benefits testing* because benefits are calculated via extreme derivations, and the benefit of an extreme derivation is obtained by adding up the weights appeared in all τ-steps. An alternative approach would be to use one special action ω (i.e. $\Omega = \{\omega\}$) in a test to report success and to take the weighted average of the weight of each path leading to an occurrence of the success action, which we refer to as *expected benefits testing.*

In this section we develop this idea, but show a negative result: amortised simulations are not sound for this form of testing.

Definition 4.15 Given a fully probabilistic computation structure, we define a function $\mathcal{F} : (\mathbb{R}_{\geq 0} \times S \to \mathbb{R}_{\geq 0}) \to (\mathbb{R}_{\geq 0} \times S \to \mathbb{R}_{\geq 0})$ as follows.

$$\mathcal{F}(f)(w,s) = \begin{cases} w & \text{if } s \stackrel{\omega}{\longrightarrow} \\ 0 & \text{if } s \not\longrightarrow \\ f(w+v,\Delta) & \text{if } s \stackrel{\tau}{\longrightarrow}_v \Delta \end{cases} \tag{14}$$

where $f(w,\Delta) = \sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot f(w,s)$. □

It is clear that the set of functions of type $\mathbb{R}_{\geq 0} \times S \to \mathbb{R}_{\geq 0}$ forms a complete lattice, with the ordering $f \leq g$ iff $f(w,s) \leq g(w,s)$ for all $w \in \mathbb{R}_{\geq 0}$ and $s \in S$. The function $\mathcal{F}$ defined above is monotonic. Therefore, it has a least fixed point which we denote by $f^\star$. Then $f^\star(0,s)$ is the expected benefits obtained by following all the paths starting from s.

Example 4.16 Consider the computation structure defined by

$$\begin{array}{rcl} s & = & \tau_1.(s \,{}_{\frac{1}{2}}\!\oplus t) \\ t & = & \omega_1.\mathbf{0} \end{array}$$

Then we have that

$$
\begin{array}{rl}
& f^\star(0,s) \\
= & \frac{1}{2}f^\star(1,s) + \frac{1}{2}f^\star(1,t) \\
= & \frac{1}{4}f^\star(2,s) + \frac{1}{4}f^\star(2,t) + \frac{1}{2}f^\star(1,t) \\
= & \frac{1}{8}f^\star(3,s) + \frac{1}{8}f^\star(3,t) + \frac{1}{4}f^\star(2,t) + \frac{1}{2}f^\star(1,t) \\
\vdots & \\
= & \sum_{k\geq 1} \frac{1}{2^k} f^\star(k,t) \\
= & \sum_{k\geq 1} \frac{k}{2^k} \\
= & 2
\end{array}
$$

□

A general probabilistic computation structure can be resolved into fully probabilistic computation structures by pruning away multiple action-choices until only single choices are left. We use the approach of [DvGMZ07] to formalise this idea:

Definition 4.17 A *resolution* of a computation structure $\langle S, \{\tau\}, W, \rightarrow\rangle$ is a fully probabilistic computation structure $\langle R, \{\tau\}, W, \rightarrow\rangle$ such that there is a resolving function $f : R \rightarrow S$ which satisfies:

1. if $r \xrightarrow{\alpha}_w \Theta$ then $f(r) \xrightarrow{\alpha}_w f(\Theta)$
2. if $r \not\longrightarrow$ then $f(r) \not\longrightarrow$

where $f(\Theta)$ is the distribution defined by $f(\Theta)(s) := \sum_{f(r)=s} \Theta(r)$. We often use the meta-variable R to refer to a resolution, with resolving function f_R. □

Definition 4.18 In a wMDP M, for any $\Delta \in \mathcal{D}(S)$, let

$$\mathbf{EBenefits}(\Delta) \;=\; \{f^\star(0,\Theta) \mid R \text{ is a resolution of } M \text{ and } f_R(\Theta) = \Delta.\}$$

For any two processes P, Q we write $P \leqq^r_{\text{may}} Q$ if for every test T,

$$\mathbf{EBenefits}(P \mid\mid T) \leq^r_{\text{Ho}} \mathbf{EBenefits}(Q \mid\mid T).$$

□

Example 4.19 [$\lhd$ is not sound for $\leqq_{\text{may}}$] Consider the following processes:

$$
\begin{array}{rcl}
P & = & \tau_2.(\mathbf{0}\;{}_{\frac{1}{4}}\!\oplus a_0.\mathbf{0}) \\
Q & = & \tau_1.(\tau_2.(\mathbf{0}\;{}_{\frac{1}{2}}\!\oplus a_0.\mathbf{0})\;{}_{\frac{1}{2}}\!\oplus a_0.\mathbf{0})
\end{array}
$$

It is easy to see that $P \lhd_0 Q$ since the transition $P \xrightarrow{\tau}_2 \mathbf{0}\;{}_{\frac{1}{4}}\!\oplus a_0.\mathbf{0}$ can be simulated by the hyper-transition $Q \overset{\tau}{\Longrightarrow}_2 \mathbf{0}\;{}_{\frac{1}{4}}\!\oplus a_0.\mathbf{0}$. Now let T be the test $\bar{a}_0.\omega$. Both $P \mid\mid T$ and $Q \mid\mid T$ give rise to fully probabilistic wMDPs. We calculate the values of $f^\star(0, P \mid\mid T)$ and $f^\star(0, Q \mid\mid T)$ as follows.

$$
\begin{array}{rcl}
f^\star(0, P \mid\mid T) & = & \frac{1}{4}\cdot 0 + \frac{3}{4}\cdot 2 = \frac{3}{2} \\
f^\star(0, Q \mid\mid T) & = & \frac{1}{2}\cdot 1 + \frac{1}{2}(\frac{1}{2}\cdot 0 + \frac{1}{2}\cdot 3) = \frac{5}{4}
\end{array}
$$

As $\mathbf{EBenefits}(P \mid\mid T) = \{\frac{3}{2}\} \not\leq^0_{\text{Ho}} \{\frac{5}{4}\} = \mathbf{EBenefits}(Q \mid\mid T)$, we have that $P \not\leqq^0_{\text{may}} Q$. Note that if we consider total benefits, then $\mathbf{Benefits}(P \mid\mid T) = \{2\} = \mathbf{Benefits}(Q \mid\mid T)$. □

5 Concluding remarks

We have proposed the model of weighted Markov decision processes for compositional reasoning about the behaviour of systems with uncertainty. Amortised weighted simulation is coinductively defined to be a behavioural preorder for comparing different wMDPs. It is shown to be a precongruence relation with respect to all structural operators for constructing wMDPs from components. For finitary convergent wMDPs, we have also given logical and testing characterisations of the simulation preorder: it can be completely determined by a quantitative probabilistic logic and for each system we can find a characteristic formula to capture its behaviour; the simulation preorder also coincides with a notion of may testing preorder.

In Section 4.2 we have shown that multi-success testing coincides with benefits testing. We can also show that multi-success testing coincides with uni-testing, where only one success action is used in tests. An analogous result is proved in [DvGMZ07] for probabilistic automata; the ideas from that proof can be adapted to the current setting, although we have one extra dimension to take into account, the weights of actions.

The dual of may testing is must testing. It would be interesting to investigate the must preorder given by our testing approach. We leave it as future work to provide a coinductive formulation of the preorder and study its logical characterisations.

There is a very limited literature on compositional theories of Markov decision processes particularly in the presence of weights. There is however an extensive literature on probabilistic variations of bisimulation equivalence for Markov chains; see Chapter 10 of [BK08] for an elementary introduction and [JLY01] for a survey. Bisimulation equivalence has also been defined in [Her02] for *Interactive Markov Chains (IMCs)*, and it is shown to be compositional, in the sense of our Theorem 3.4: it is preserved by the operators of a process calculus interpreted as IMCs. Bisimulation and testing equivalence for Markovian process algebras are also investigated in [Hil96, BC00], but the analysis was mainly restricted to models free of nondeterminism. Recently a combination of probabilistic automata and IMCs has been studied in [EHZ10], where a notation of weak bisimulation is proposed. Since time rates are treated essentially as action names, some intuitively equivalent processes are differentiated by the weak bisimulation. A variant of the weak bisimulation is proposed in [DH11]; it is justified by its coincidence with a natural extensional equivalence relation for finitary systems.

There is also an extensive literature on weighted automata [DKV09], and probabilistic variations have also been studied [CDH09]. However there the focus is on traditional language theoretic issues, rather than our primary concern, compositionality.

A Elementary properties of hyper-derivations

This appendix contains the details proofs of the properties of hyper-derivations announced in Section 2.3.

Lemma A.1

1. If $\Delta \stackrel{\tau}{\Longrightarrow}_v \Theta$ then $|\Delta| \geq |\Theta|$.

2. If $\Delta \stackrel{\tau}{\Longrightarrow}_v \Theta$ and $p \in \mathbb{R}$ such that $|p \cdot \Delta| \leq 1$, then $p \cdot \Delta \stackrel{\tau}{\Longrightarrow}_{pv} p \cdot \Theta$.

3. If $\Gamma + \Lambda \stackrel{\tau}{\Longrightarrow}_v \Pi$ then $\Pi = \Pi^\Gamma + \Pi^\Lambda$ with $\Gamma \stackrel{\tau}{\Longrightarrow}_{v^\Gamma} \Pi^\Gamma$, $\Lambda \stackrel{\tau}{\Longrightarrow}_{v^\Lambda} \Pi^\Lambda$, and $v = v^\Gamma + v^\Lambda$.

Proof.

1. By definition $\Delta \stackrel{\tau}{\Longrightarrow}_v \Theta$ means that some $\Delta_k, \Delta_k^\times, \Delta_k^\rightarrow, v_k$ exist for all $k \geq 0$ such that

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

A simple inductive proof shows that

$$|\Delta| \;=\; |\Delta_i^\rightarrow| + \sum_{k \leq i} |\Delta_k^\times| \text{ for any } i \geq 0. \tag{15}$$

The sequence $\{\sum_{k \leq i} |\Delta_k|\}_{i=0}^{\infty}$ is nondecreasing and by (15) each element of the sequence is not greater than $|\Delta|$. Therefore, the limit of this sequence is bounded by $|\Delta|$. That is,

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

2. Now suppose $p \in \mathbb{R}$ such that $|p \cdot \Delta| \leq 1$. From Definition 2.2 it follows that

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

Hence Definition 2.8 yields $p \cdot \Delta \stackrel{\tau}{\Longrightarrow}_{pv} p \cdot \Theta$.

3. Suppose $\Gamma + \Lambda \stackrel{\tau}{\Longrightarrow}_v \Pi$. From Definition 2.8 we have

$$\Gamma + \Lambda = \Pi_0 = \Pi_0^\rightarrow + \Pi_0^\times \tag{16}$$

for some $\Pi_0^\rightarrow, \Pi_0^\times$ with $\Pi_0^\rightarrow \stackrel{\tau}{\longrightarrow}_{v_0} \Pi_1$ for some Π_1. Let us define subdistributions $\Gamma^\rightarrow, \Gamma^\times, \Lambda^\rightarrow, \Lambda^\times$ as follows. For any $s \in S$,

$$\begin{array}{rcl} \Gamma^\rightarrow(s) &=& \min(\Gamma(s), \Pi_0^\rightarrow(s)) \\ \Gamma^\times(s) &=& \Gamma(s) - \Gamma^\rightarrow(s) \\ \Lambda^\times(s) &=& \min(\Lambda(s), \Pi_0^\times(s)) \\ \Lambda^\rightarrow(s) &=& \Lambda(s) - \Lambda^\times(s) \end{array} \tag{17}$$

Clearly, we have $\Gamma = \Gamma^\rightarrow + \Gamma^\times$ and $\Lambda = \Lambda^\rightarrow + \Lambda^\times$. Below we show that

$$\Pi_0^\rightarrow = \Gamma^\rightarrow + \Lambda^\rightarrow \text{ and } \Pi_0^\times = \Gamma^\times + \Lambda^\times. \tag{18}$$

For any $s \in S$, we distinguish two cases:

(a) $\Pi_0^\rightarrow(s) \geq \Gamma(s)$. In this case we have $\Pi^\times(s) \leq \Lambda(s)$ by (16). It follows from (17) that $\Gamma^\rightarrow(s) = \Gamma(s)$, $\Gamma^\times(s) = 0$, $\Lambda^\times(s) = \Pi_0^\times(s)$, and $\Lambda^\rightarrow(s) = \Lambda(s) - \Pi_0^\times(s)$. Therefore,

$$\begin{array}{rcll} \Gamma^\rightarrow(s) + \Lambda^\rightarrow(s) &=& \Gamma(s) + \Lambda(s) - \Pi_0^\times(s) & \\ &=& \Pi_0(s) - \Pi_0^\times(s) & \text{by (16)} \\ &=& \Pi_0^\rightarrow(s) & \end{array}$$

$$\begin{array}{rcl} \Gamma^\times(s) + \Lambda^\times(s) &=& 0 + \Pi^\times(s) \\ &=& \Pi^\times(s) \end{array}$$

(b) $\Pi_0^{\rightarrow}(s) < \Gamma(s)$. Similarly we can show that $\Gamma^{\rightarrow}(s)+\Lambda^{\rightarrow}(s) = \Pi_0^{\rightarrow}(s)$ and $\Gamma^{\times}(s)+\Lambda^{\times}(s) = \Pi_0^{\times}(s)$.

So we have verified (18). Since $\Pi_0^{\rightarrow} \xrightarrow{\tau}_{v_0} \Pi_1$, we use (18) and Proposition 2.6 to find $v_0', v_0'', \Gamma_1, \Lambda_1$ with $\Gamma^{\rightarrow} \xrightarrow{\tau}_{v_0'} \Gamma_1$, $\Lambda^{\rightarrow} \xrightarrow{\tau}_{v_0''} \Lambda_1$, $v_0 = v_0' + v_0''$, and $\Pi_1 = \Gamma_1 + \Lambda_1$. Now from Γ_1, Λ_1 we can continue the above procedure for Γ, Λ to induce Γ_2, Λ_2, and then Γ_3, Λ_3, etc. such that

$$\Gamma = \Gamma_0, \qquad \Gamma_k = \Gamma_k^{\rightarrow} + \Gamma_k^{\times}, \quad \Gamma_k^{\rightarrow} \xrightarrow{\tau}_{v_k'} \Gamma_{k+1},$$
$$\Lambda = \Lambda_0, \qquad \Lambda_k = \Lambda_k^{\rightarrow} + \Lambda_k^{\times}, \quad \Lambda_k^{\rightarrow} \xrightarrow{\tau}_{v_k''} \Lambda_{k+1},$$
$$\Gamma_k + \Lambda_k = \Pi_k, \quad \Gamma_k^{\rightarrow} + \Lambda_k^{\rightarrow} = \Pi_k^{\rightarrow}, \quad \Gamma_k^{\times} + \Lambda_k^{\times} = \Pi_k^{\times}.$$

Let $\Pi^{\Gamma} := \sum_k \Gamma_k^{\times}$, $\Pi^{\Lambda} := \sum_k \Lambda_k^{\times}$, $v' = \sum_k v_k'$, and $v'' = \sum_k v_k''$. Then $\Pi = \Pi^{\Gamma} + \Pi^{\Lambda}$ and Definition 2.8 yields $\Gamma \overset{\tau}{\Longrightarrow}_{v'} \Pi^{\Gamma}$ and $\Lambda \overset{\tau}{\Longrightarrow}_{v''} \Pi^{\Lambda}$.

□

We now generalise the above binary decomposition to infinite (but still countable) decomposition, and also establish linearity.

Lemma A.2 Let $p_i \in [0,1]$ for $i \in I$ where I is a countable index set with $\sum_{i\in I} p_i \leq 1$. Then

1. (Linearity) If $\Delta_i \overset{\tau}{\Longrightarrow}_{w_i} \Theta_i$ for all $i \in I$ then $\sum_{i\in I} p_i \cdot \Delta_i \overset{\tau}{\Longrightarrow}_{(\sum_{i\in I} p_i \cdot w_i)} \sum_{i\in I} p_i \cdot \Theta_i$.
2. (Decomposability) If $\sum_{i\in I} p_i \cdot \Delta_i \overset{\tau}{\Longrightarrow}_{w} \Theta$ then $w = \sum_{i\in I} p_i \cdot w_i$ and $\Theta = \sum_{i\in I} p_i \cdot \Theta_i$ for weights w_i and subdistributions Θ_i such that $\Delta_i \overset{\tau}{\Longrightarrow}_{w_i} \Theta_i$ for all $i \in I$.

Proof.

1. Suppose $\Delta_i \overset{\tau}{\Longrightarrow}_{w_i} \Theta_i$ for all $i \in I$. By Definition 2.8 there are subdistributions $\Delta_{ik}, \Delta_{ik}^{\rightarrow}, \Delta_{ik}^{\times}$ and weights w_{ik} such that

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

Therefore, we have that $\sum_{i\in I} p_i \cdot \Delta_i = \sum_{i\in I} p_i \cdot \Delta_{i0}$, $\sum_{i\in I} p_i \cdot \Delta_{ik} = \sum_{i\in I} p_i \cdot \Delta_{ik}^{\rightarrow} + \sum_{i\in I} p_i \cdot \Delta_{ik}^{\times}$, $\sum_{i\in I} p_i \cdot \Delta_{ik}^{\rightarrow} \xrightarrow{\tau}_{(\sum_{i\in I} p_i \cdot w_{ik})} \sum_{i\in I} p_i \cdot \Delta_{i(k+1)}$ by Clause (2) of Definition 2.2, $\sum_{i\in I} p_i \cdot \Theta_i = \sum_{i\in I} p_i \cdot \sum_k \Delta_{ik}^{\times} = \sum_k (\sum_{i\in I} p_i \cdot \Delta_{ik}^{\times})$, and $\sum_{i\in I} p_i \cdot w_i = \sum_{i\in I} p_i \cdot \sum_k w_{ik} = \sum_k (\sum_{i\in I} p_i \cdot w_{ik})$. By Definition 2.8 we obtain $\sum_{i\in I} p_i \cdot \Delta_i \overset{\tau}{\Longrightarrow}_{(\sum_{i\in I} p_i \cdot w_i)} \sum_{i\in I} p_i \cdot \Theta_i$.

2. In the light of Lemma A.1(ii) it suffices to show that if $\sum_{i=0}^{\infty} \Delta_i \overset{\tau}{\Longrightarrow}_{w} \Theta$ then $w = \sum_{i=0}^{\infty} w_i$ for weights w_i and $\Theta = \sum_{i=0}^{\infty} \Theta_i$ for subdistributions Θ_i such that $\Delta_i \overset{\tau}{\Longrightarrow}_{w_i} \Theta_i$ for all $i \geq 0$. Since $\sum_{i=0}^{\infty} \Delta_i = \Delta_0 + \sum_{k\geq 1} \Delta_k$ and $\sum_{i=0}^{\infty} \Delta_i \overset{\tau}{\Longrightarrow}_{w} \Theta$, by Lemma A.1(3) there are $\Theta_0, \Theta_1^{\geq}, w_0, w_{\geq 1}$ such that

$$\Delta_0 \overset{\tau}{\Longrightarrow}_{w_0} \Theta_0, \qquad \sum_{k\geq 1} \Delta_k \overset{\tau}{\Longrightarrow}_{w_{\geq 1}} \Theta_1^{\geq}, \qquad \Theta = \Theta_0 + \Theta_1^{\geq}, \qquad w = w_0 + w_{\geq 1}.$$

Using Lemma A.1(3) again, we have $\Theta_1, \Theta_2^{\geq}, w_1, w_{\geq 2}$ such that

$$\Delta_1 \overset{\tau}{\Longrightarrow}_{w_1} \Theta_1, \qquad \sum_{k\geq 2} \Delta_k \overset{\tau}{\Longrightarrow}_{w_{\geq 2}} \Theta_2^{\geq}, \qquad \Theta_1^{\geq} = \Theta_1 + \Theta_2^{\geq}, \qquad w_{\geq 1} = w_1 + w_{\geq 2}$$

thus in combination $\Theta = \Theta_0 + \Theta_1 + \Theta_2^{\geq}$ and $w = w_0 + w_1 + w_{\geq 2}$. Continuing this process we have that

$$\Delta_k \overset{\tau}{\Longrightarrow}_{w_k} \Theta_k, \quad \sum_{j\geq k} \Delta_j \overset{\tau}{\Longrightarrow}_{w_{\geq k+1}} \Theta_{k+1}^{\geq}, \quad \Theta = \sum_{j=0}^{k} \Theta_j + \Theta_{k+1}^{\geq}, \quad w = \sum_{j=0}^{k} w_j + w_{\geq k+1} \quad (19)$$

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

$$w \;=\; \sum_{k=0}^{\infty} w_k, \qquad \Theta \;=\; \sum_{k=0}^{\infty} \Theta_k \;. \qquad (20)$$

□

Corollary A.3 The relation $\overset{\tau}{\Longrightarrow}$ is convex.

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

Theorem A.4 (Theorem 2.13) If $\Delta \overset{\tau}{\Longrightarrow}_u \Theta$ and $\Theta \overset{\tau}{\Longrightarrow}_v \Lambda$ then $\Delta \overset{\tau}{\Longrightarrow}_{u+v} \Lambda$.

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

$$\Delta = \Delta_0, \qquad \Delta_k = \Delta_k^{\times} + \Delta_k^{\rightarrow}, \qquad \Delta_k^{\rightarrow} \overset{\tau}{\longrightarrow}_{u_k} \Delta_{k+1}, \qquad \Theta = \sum_{k=0}^{\infty} \Delta_k^{\times}, \qquad u = \sum_{k=0}^{\infty} u_k. \qquad (21)$$

Since $\Theta = \sum_{k=0}^{\infty} \Delta_k^{\times}$ and $\Theta \overset{\tau}{\Longrightarrow}_v \Lambda$, by Lemma A.2(2) there are Λ_k, w_k for $k \geq 0$ such that

$$v = \sum_{k=0}^{\infty} v_k, \qquad \Lambda = \sum_{k=0}^{\infty} \Lambda_k, \qquad \Delta_k^{\times} \overset{\tau}{\Longrightarrow}_{v_k} \Lambda_k \qquad (22)$$

for all $k \geq 0$. For each $k \geq 0$, we know from $\Delta_k^{\times} \overset{\tau}{\Longrightarrow}_{v_k} \Lambda_k$ that there are some v_{kl}, Δ_{kl}, $\Delta_{kl}^{\times}, \Delta_{kl}^{\rightarrow}$ for $l \geq 0$ such that

$$\Delta_k^{\times} = \Delta_{k0}, \qquad \Delta_{kl} = \Delta_{kl}^{\times} + \Delta_{kl}^{\rightarrow}, \qquad \Delta_{kl}^{\rightarrow} \overset{\tau}{\longrightarrow}_{v_{kl}} \Delta_{k,l+1} \qquad \Lambda_k = \sum_{l\geq 0} \Delta_{kl}^{\times}, \qquad v_k = \sum_{l\geq 0} v_{kl}. \quad (23)$$

Therefore we can put all this together with

$$\Lambda \;=\; \sum_{k=0}^{\infty} \Lambda_k \;=\; \sum_{k,l\geq 0} \Delta_{kl}^{\times} \;=\; \sum_{i\geq 0} \left(\sum_{k,l|k+l=i} \Delta_{kl}^{\times} \right) \;, \qquad (24)$$

where the last step is a straightforward diagonalisation. Similarly,

$$v \quad = \quad \sum_{k=0}^{\infty} v_k \quad = \quad \sum_{k,l\geq 0} v_{kl} \quad = \quad \sum_{i\geq 0}\left(\sum_{k,l|k+l=i} v_{kl}\right) \ , \tag{25}$$

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

$$\Delta'_i = \Delta'^{\times}_i + \Delta'^{\rightarrow}_i, \qquad \Delta'^{\times}_i = \sum_{k,l|k+l=i} \Delta^{\times}_{kl}, \qquad \Delta'^{\rightarrow}_i = (\sum_{k,l|k+l=i} \Delta^{\rightarrow}_{kl}) + \Delta^{\rightarrow}_i \ , \qquad w_i = (\sum_{k,l|k+l=i} v_{kl}) + u_i \tag{26}$$

so that from (24) we have immediately that

$$\Lambda \quad = \quad \sum_{i\geq 0} \Delta'^{\times}_i \ . \tag{27}$$

We now show that

1. $\Delta = \Delta'_0$
2. $\Delta'^{\rightarrow}_i \stackrel{\tau}{\longrightarrow}_{w_i} \Delta'_{i+1}$
3. $\sum_{i\geq 0} w_i = u + v$

from which, with (26) and (27), we will have $\Delta \stackrel{\tau}{\Longrightarrow}_{u+v} \Lambda$ as required. For (1) we observe that

$$\begin{array}{lll}
 & \Delta & \\
= & \Delta_0 & (21) \\
= & \Delta^{\times}_0 + \Delta^{\rightarrow}_0 & (21) \\
= & \Delta_{00} + \Delta^{\rightarrow}_0 & (23) \\
= & \Delta^{\times}_{00} + \Delta^{\rightarrow}_{00} + \Delta^{\rightarrow}_0 & (23) \\
= & (\sum_{k,l|k+l=0} \Delta^{\times}_{kl}) + (\sum_{k,l|k+l=0} \Delta^{\rightarrow}_{kl}) + \Delta^{\rightarrow}_0 & \text{index arithmetic} \\
= & \Delta'^{\times}_0 + \Delta'^{\rightarrow}_0 & (26) \\
= & \Delta'_0 \ . & (26)
\end{array}$$

For (2) we observe that

$$\begin{array}{lll}
 & \Delta'^{\rightarrow}_i & \\
= & (\sum_{k,l|k+l=i} \Delta^{\rightarrow}_{kl}) + \Delta^{\rightarrow}_i & (26) \\
\stackrel{\tau}{\longrightarrow}_{w_i} & (\sum_{k,l|k+l=i} \Delta_{k,l+1}) + \Delta_{i+1} & (21),\ (23),\ \text{Definition 2.8(2)} \\
= & (\sum_{k,l|k+l=i}(\Delta^{\times}_{k,l+1} + \Delta^{\rightarrow}_{k,l+1})) + \Delta^{\times}_{i+1} + \Delta^{\rightarrow}_{i+1} & (21),\ (23) \\
= & (\sum_{k,l|k+l=i} \Delta^{\times}_{k,l+1}) + \Delta^{\times}_{i+1} + (\sum_{k,l|k+l=i} \Delta^{\rightarrow}_{k,l+1}) + \Delta^{\rightarrow}_{i+1} & \text{rearrange} \\
= & (\sum_{k,l|k+l=i} \Delta^{\times}_{k,l+1}) + \Delta_{i+1,0} + (\sum_{k,l|k+l=i} \Delta^{\rightarrow}_{k,l+1}) + \Delta^{\rightarrow}_{i+1} & (23) \\
= & (\sum_{k,l|k+l=i} \Delta^{\times}_{k,l+1}) + \Delta^{\times}_{i+1,0} + \Delta^{\rightarrow}_{i+1,0} + (\sum_{k,l|k+l=i} \Delta^{\rightarrow}_{k,l+1}) + \Delta^{\rightarrow}_{i+1} & (23) \\
= & (\sum_{k,l|k+l=i+1} \Delta^{\times}_{kl}) + (\sum_{k,l|k+l=i+1} \Delta^{\rightarrow}_{kl}) + \Delta^{\rightarrow}_{i+1} & \text{index arithmetic} \\
= & \Delta'^{\times}_{i+1} + \Delta'^{\rightarrow}_{i+1} & (26) \\
= & \Delta'_{i+1} \ . & (26)
\end{array}$$

For (3) we observe that $\sum_{i\geq 0} w_i = \sum_{i\geq 0}(\sum_{k,l|k+l=i} v_{kl}) + \sum_{i\geq 0} u_i = v + u$ by (26) and (21-23), which concludes the proof. □

B Proof of Theorem 2.19

In this section we introduce the machinery used to prove Theorem 2.19, which directly leads to the finite generability theorem. The machinery employs some concepts such as discounted hyper-derivation, discounted payoff, max-seeking policy etc., because we need to first establish a discounted version of Theorem 2.19.

Definition B.1 [Discounted hyper-derivation] The *discounted hyper-derivation* $\Delta \stackrel{\tau}{\Longrightarrow}_{\delta,w} \Delta'$ for discount factor δ $(0 \leq \delta \leq 1)$ is obtained from a hyper-derivation by discounting each τ transition by δ. That is, there is a collection of $\Delta_k^{\rightarrow}, \Delta_k^{\times}, w_k$ satisfying

$$\begin{array}{ccc} \Delta & = & \Delta_0^{\rightarrow} + \Delta_0^{\times} \\ \Delta_0^{\rightarrow} & \stackrel{\tau}{\longrightarrow}_{w_1} & \Delta_1^{\rightarrow} + \Delta_1^{\times} \\ & \vdots & \\ \Delta_k^{\rightarrow} & \stackrel{\tau}{\longrightarrow}_{w_{k+1}} & \Delta_{k+1}^{\rightarrow} + \Delta_{k+1}^{\times} \\ & \vdots & \end{array}$$

such that $w = \sum_{k=1}^{\infty} \delta^k w_k$ and $\Delta' = \sum_{k=0}^{\infty} \delta^k \Delta_k^{\times}$. □

It is trivial that the relation $\stackrel{\tau}{\Longrightarrow}_{1,w}$ coincides with $\stackrel{\tau}{\Longrightarrow}_{w}$.

Definition B.2 [Discounted payoff] Given a discount δ and weight function $\mathbf{w}$, the *discounted payoff function* $\mathbb{P}^{\delta,\mathbf{w}}_{\max} : S \to \mathbb{R}$ is defined by

$$\mathbb{P}^{\delta,\mathbf{w}}_{\max}(s) \;=\; \sup\{\mathbf{w} \centerdot \langle w, \Delta' \rangle \mid \overline{s} \stackrel{\tau}{\Longrightarrow}_{\delta,w} \Delta'\}$$

and we will generalise it to be of type $\mathcal{D}_{sub}(S) \to \mathbb{R}$ by letting $\mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta) = \sum_{s\in\lceil\Delta\rceil} \Delta(s) \cdot \mathbb{P}^{\delta,\mathbf{w}}_{\max}(s)$. □

Definition B.3 [Max-seeking policy] Given a wMDP, discount δ and weighted function $\mathbf{w}$, we say a static policy pp is *max-seeking* with respect to δ and $\mathbf{w}$ if for all s the following requirements are met.

1. If $\mathsf{pp}(s)\!\uparrow$, then $\mathbf{w} \centerdot \langle 0, \overline{s} \rangle \;\geq\; \delta(\mathbf{w} \centerdot \langle w_1, \varepsilon \rangle + \mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta_1))$ for all $s \stackrel{\tau}{\longrightarrow}_{w_1} \Delta_1$.
2. If $\mathsf{pp}(s) = \langle w, \Delta \rangle$ then
   (a) $\delta(\mathbf{w} \centerdot \langle w, \varepsilon \rangle + \mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta)) \;\geq\; \mathbf{w} \centerdot \langle 0, \overline{s} \rangle$ and
   (b) $\mathbf{w} \centerdot \langle w, \varepsilon \rangle + \mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta) \;\geq\; \mathbf{w} \centerdot \langle w_1, \varepsilon \rangle + \mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta_1)$ for all $s \stackrel{\tau}{\longrightarrow}_{w_1} \Delta_1$.

□

Lemma B.4 Given a finitary wMDP, discount δ and weighted function $\mathbf{w}$, there always exists a max-seeking policy.

Proof. Given a wMDP, discount δ and weighted function $\mathbf{w}$, the discounted payoff $\mathbb{P}^{\delta,\mathbf{w}}_{\max}(s)$ can be calculated for each state s. Then we can define a static policy pp in the following way. For any state s, if $\mathbf{w} \centerdot \langle 0, \overline{s} \rangle \;\geq\; \delta(\mathbf{w} \centerdot \langle w_1, \varepsilon \rangle + \mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta_1))$ for all $s \xrightarrow{\tau}_{w_1} \Delta_1$, then we set pp undefined at s. Otherwise, we choose a transition $s \xrightarrow{\tau}_{w} \Delta$ among the finite number of outgoing transitions from s such that $\mathbf{w} \centerdot \langle w, \varepsilon \rangle + \mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta) \;\geq\; \mathbf{w} \centerdot \langle w_1, \varepsilon \rangle + \mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta_1)$ for all other transitions $s \xrightarrow{\tau}_{w_1} \Delta_1$, and we set $\mathsf{pp}(s) = \langle w, \Delta \rangle$. □

Given a wMDP, discount δ, weight function $\mathbf{w}$, and static policy pp, we define the function $F^{\delta,\mathsf{pp},\mathbf{w}} : (S \to \mathbb{R}) \to (S \to \mathbb{R})$ by

$$F^{\delta,\mathsf{pp},\mathbf{w}} \;:=\; \lambda f.\lambda s. \begin{cases} \mathbf{w} \centerdot \langle 0, \overline{s} \rangle & \text{if } \mathsf{pp}(s)\uparrow \\ \delta(\mathbf{w} \centerdot \langle w, \varepsilon \rangle + f(\Delta)) & \text{if } \mathsf{pp}(s) = \langle w, \Delta \rangle \end{cases} \tag{28}$$

where $f(\Delta) = \sum_{s \in \lceil \Delta \rceil} \Delta(s) \cdot f(s)$.

Lemma B.5 Given a wMDP, discount $\delta < 1$, weight function $\mathbf{w}$, and static policy pp, the function $F^{\delta,\mathsf{pp},\mathbf{w}}$ has a unique fixed point.

Proof. We first show that the function $F^{\delta,\mathsf{pp},\mathbf{w}}$ is a contraction mapping. Let f, g be any two functions of type $S \to \mathbb{R}$.

$$\begin{array}{ll} & |F^{\delta,\mathsf{pp},\mathbf{w}}(f) - F^{\delta,\mathsf{pp},\mathbf{w}}(g)| \\ = & \sup\{|F^{\delta,\mathsf{pp},\mathbf{w}}(f)(s) - F^{\delta,\mathsf{pp},\mathbf{w}}(g)(s)| \;\mid s \in S\} \\ = & \sup\{|F^{\delta,\mathsf{pp},\mathbf{w}}(f)(s) - F^{\delta,\mathsf{pp},\mathbf{w}}(g)(s) \mid \;\mid s \in S \text{ and } \mathsf{pp}(s)\downarrow\} \\ = & \delta \cdot \sup\{|f(\Delta) - g(\Delta)| \;\mid s \in S \text{ and } \mathsf{pp}(s) = \langle w, \Delta \rangle \text{ for some } \Delta\} \\ \leq & \delta \cdot \sup\{|f(s') - g(s')| \;\mid s' \in S\} \\ = & \delta \cdot |f - g| \\ < & |f - g| \end{array}$$

By Banach unique fixed point theorem, the function $F^{\delta,\mathsf{pp},\mathbf{w}}$ has a unique fixed point. □

Lemma B.6 Given a wMDP, discount δ, weight function $\mathbf{w}$, and max-seeking static policy pp, the function $\mathbb{P}^{\delta,\mathbf{w}}_{\max}$ is a fixed point of $F^{\delta,\mathsf{pp},\mathbf{w}}$.

Proof. We need to show that $F^{\delta,\mathsf{pp},\mathbf{w}}(\mathbb{P}^{\delta,\mathbf{w}}_{\max})(s) = \mathbb{P}^{\delta,\mathbf{w}}_{\max}(s)$ holds for any state s. We distinguish two cases.

1. If $\mathsf{pp}(s)\uparrow$, then $F^{\delta,\mathsf{pp},\mathbf{w}}(\mathbb{P}^{\delta,\mathbf{w}}_{\max})(s) = \mathbf{w} \centerdot \langle 0, \overline{s} \rangle = \mathbb{P}^{\delta,\mathbf{w}}_{\max}(s)$ as expected.
2. If $\mathsf{pp}(s) = \langle w, \Delta \rangle$, then the arguments are more involved. First note that if $\overline{s} \overset{\tau}{\Longrightarrow}_{\delta,w} \Delta''$, then by Definition B.1 there exist some $\Delta_0^{\rightarrow}, \Delta_0^{\times}, \Delta_1, \Delta'', w_1, w'$ such that $\overline{s} = \Delta_0^{\rightarrow} + \Delta_0^{\times}$,

$\Delta_0^{\rightarrow} \xrightarrow{\tau}_{w_1} \Delta_1$, $\Delta_1 \stackrel{\tau}{\Longrightarrow}_{\delta,w'} \Delta''$, $\Delta' = \Delta_0^{\times} + \delta\cdot\Delta''$ and $w = \delta(w_1 + w')$. So we can do the following calculation.

$$
\begin{array}{rl}
& \mathbb{P}^{\delta,\mathbf{w}}_{\max}(s) \\
= & \sup\{\mathbf{w}\centerdot\langle\, w, \Delta'\,\rangle \mid \overline{s} \stackrel{\tau}{\Longrightarrow}_{\delta,w} \Delta'\} \\
= & \sup\{\mathbf{w}\centerdot\langle\, \delta(w_1+w'), \Delta_0^{\times} + \delta\cdot\Delta''\,\rangle \mid \overline{s} = \Delta_0^{\rightarrow} + \Delta_0^{\times}, \Delta_0^{\rightarrow} \xrightarrow{\tau}_{w_1} \Delta_1, \text{ and } \Delta_1 \stackrel{\tau}{\Longrightarrow}_{\delta,w'} \Delta'' \\
& \qquad \text{for some } \Delta_0^{\rightarrow}, \Delta_0^{\times}, \Delta_1, \Delta'', w_1, w'\} \\
= & \sup\{\mathbf{w}\centerdot\langle\, 0, \Delta_0^{\times}\,\rangle + \delta(\mathbf{w}\centerdot\langle\, w_1, \varepsilon\,\rangle + \mathbf{w}\centerdot\langle\, w', \Delta''\,\rangle) \mid \overline{s} = \Delta_0^{\rightarrow} + \Delta_0^{\times}, \Delta_0^{\rightarrow} \xrightarrow{\tau}_{w_1} \Delta_1, \text{ and } \Delta_1 \stackrel{\tau}{\Longrightarrow}_{\delta,w'} \Delta'' \\
& \qquad \text{for some } \Delta_0^{\rightarrow}, \Delta_0^{\times}, \Delta_1, \Delta'', w_1, w'\} \\
= & \sup\{\mathbf{w}\centerdot\langle\, 0, \Delta_0^{\times}\,\rangle + \delta(\mathbf{w}\centerdot\langle\, w_1, \varepsilon\,\rangle + \sup\{\mathbf{w}\centerdot\langle\, w', \Delta''\,\rangle \mid \Delta_1 \stackrel{\tau}{\Longrightarrow}_{\delta,w'} \Delta'' \text{ for some } w', \Delta''\}) \\
& \qquad \mid \overline{s} = \Delta_0^{\rightarrow} + \Delta_0^{\times} \text{ and } \Delta_0^{\rightarrow} \xrightarrow{\tau}_{w_1} \Delta_1 \text{ for some } \Delta_0^{\rightarrow}, \Delta_0^{\times}, \Delta_1, w_1\} \\
= & \sup\{\mathbf{w}\centerdot\langle\, 0, \Delta_0^{\times}\,\rangle + \delta(\mathbf{w}\centerdot\langle\, w_1, \varepsilon\,\rangle + \mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta_1)) \mid \overline{s} = \Delta_0^{\rightarrow} + \Delta_0^{\times} \text{ and } \Delta_0^{\rightarrow} \xrightarrow{\tau}_{w_1} \Delta_1 \\
& \qquad \text{for some } \Delta_0^{\rightarrow}, \Delta_0^{\times}, \Delta_1, w_1\} \\
= & \sup\{\mathbf{w}\centerdot\langle\, 0, (1-p)\overline{s}\,\rangle + p\delta(\mathbf{w}\centerdot\langle\, w_1, \varepsilon\,\rangle + \mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta_1)) \mid p \in [0,1] \text{ and } \overline{s} \xrightarrow{\tau}_{w_1} \Delta_1 \\
& \qquad \text{for some } \Delta_1, w_1\} \qquad [\overline{s} \text{ can be split into } p\overline{s} + (1-p)\overline{s} \text{ only}] \\
= & \sup\{\mathbf{w}\centerdot\langle\, 0, (1-p)\overline{s}\,\rangle + p\delta(\mathbf{w}\centerdot\langle\, w_1, \varepsilon\,\rangle + \mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta_1)) \mid p \in [0,1] \text{ and } s \xrightarrow{\tau}_{w_1} \Delta_1 \\
& \qquad \text{for some } \Delta_1, w_1\} \\
= & \sup\{\mathbf{w}\centerdot\langle\, 0, (1-p)\overline{s}\,\rangle + p\delta\cdot\sup\{\mathbf{w}\centerdot\langle\, w_1, \varepsilon\,\rangle + \mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta_1) \mid s \xrightarrow{\tau}_{w_1} \Delta_1\} \mid p \in [0,1]\} \\
= & \max(\mathbf{w}\centerdot\langle\, 0, \overline{s}\,\rangle, \quad \delta\cdot\sup\{\mathbf{w}\centerdot\langle\, w_1, \varepsilon\,\rangle + \mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta_1) \mid s \xrightarrow{\tau}_{w_1} \Delta_1\}) \\
= & \delta(\mathbf{w}\centerdot\langle\, w, \varepsilon\,\rangle + \mathbb{P}^{\delta,\mathbf{w}}_{\max}(\Delta)) \qquad [\text{as } \mathsf{pp} \text{ is max-seeking}] \\
= & F^{\delta,\mathsf{pp},\mathbf{w}}(\mathbb{P}^{\delta,\mathbf{w}}_{\max})(s)
\end{array}
$$

□

Definition B.7 [Discounted hyper-SP-derivation] Let Δ be a subdistribution and pp a static policy. We define a collection of subdistributions Δ_k and weights w_k as follows.

$$
\begin{array}{rcll}
\Delta_0 & = & \Delta & \\
\langle\, w_{k+1}, \Delta_{k+1}\,\rangle & = & \sum\{\Delta_k(s)\cdot\mathsf{pp}(s) \mid s \in \lceil\Delta_k\rceil \text{ and } \mathsf{pp}(s)\!\downarrow\} & \quad \text{for all } k \geq 0.
\end{array}
$$

Then $\Delta_k^{\times}$ is obtained from Δ_k by letting

$$
\Delta_k^{\times}(s) \;=\; \begin{cases} 0 & \text{if } \mathsf{pp}(s)\!\downarrow \\ \Delta_k(s) & \text{otherwise} \end{cases}
$$

for all $k \geq 0$. Then the *discounted hyper-SP-derivation* $\Delta \stackrel{\tau}{\Longrightarrow}_{\delta,\mathsf{pp},w} \Delta'$ determines a unique weight w and subdistribution Δ' with $w = \sum_{k=1}^{\infty} \delta^k w_k$ and $\Delta' = \sum_{k=0}^{\infty} \delta^k \Delta_k^{\times}$. □

In other words, if $\Delta \stackrel{\tau}{\Longrightarrow}_{\delta,\mathsf{pp},w} \Delta'$ then w and Δ' come from the discounted hyper-derivation $\Delta \stackrel{\tau}{\Longrightarrow}_{\delta,w} \Delta'$ which is constructed by following the static policy pp when choosing τ transitions from each state. If the discount factor $\delta = 1$, we write $\stackrel{\tau}{\Longrightarrow}_{\mathsf{pp},w}$ in place of $\stackrel{\tau}{\Longrightarrow}_{1,\mathsf{pp},w}$.

Definition B.8 [Policy-following payoff] Given a discount δ, weight function $\mathbf{w}$, and static policy pp, the *policy-following payoff function* $\mathbb{P}^{\delta,\mathsf{pp},\mathbf{w}} : S \to \mathbb{R}$ is defined by

$$
\mathbb{P}^{\delta,\mathsf{pp},\mathbf{w}}(s) \;=\; \mathbf{w}\centerdot\langle\, w, \Delta'\,\rangle
$$

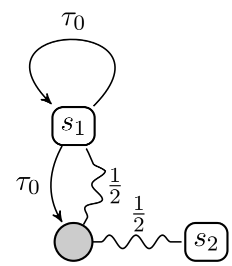

Figure 8: Max-seeking policies

where w, Δ are determined by the discounted hyper-SP-derivation $\overline{s} \overset{\tau}{\Longrightarrow}_{\delta,\mathsf{pp},w} \Delta'$. Note that for discount $\delta = 1$ this coincides with the function given in Definition 2.18; that is $\mathbb{P}^{1,\mathsf{pp},\mathbf{w}}(s) = \mathbb{P}^{\mathsf{pp},\mathbf{w}}(s)$. $\square$

Lemma B.9 For any discount δ, weight function $\mathbf{w}$, and static policy pp, the function $\mathbb{P}^{\delta,\mathsf{pp},\mathbf{w}}$ is a fixed point of $F^{\delta,\mathsf{pp},\mathbf{w}}$.

Proof. We need to show that $F^{\delta,\mathsf{pp},\mathbf{w}}(\mathbb{P}^{\delta,\mathsf{pp},\mathbf{w}})(s) = \mathbb{P}^{\delta,\mathsf{pp},\mathbf{w}}(s)$ holds for any state s. There are two cases.

1. If $\mathsf{pp}(s)\uparrow$, then $\overline{s} \overset{\tau}{\Longrightarrow}_{\delta,\mathsf{pp},w} \Delta'$ implies $w = 0$ and $\Delta' = \overline{s}$. Thus, $\mathbb{P}^{\delta,\mathsf{pp},\mathbf{w}}(s) = \mathbf{w} \centerdot \langle 0, \overline{s} \rangle = F^{\delta,\mathsf{pp},\mathbf{w}}(\mathbb{P}^{\delta,\mathsf{pp},\mathbf{w}})(s)$ as required.
2. Suppose $\mathsf{pp}(s) = \langle w_1, \Delta_1 \rangle$. If $\overline{s} \overset{\tau}{\Longrightarrow}_{\delta,\mathsf{pp},w} \Delta'$ then $s \overset{\tau}{\longrightarrow}_{w_1} \Delta_1$, $\Delta_1 \overset{\tau}{\Longrightarrow}_{\delta,\mathsf{pp},w'} \Delta''$, $\Delta' = \delta\Delta''$ and $w = \delta(w_1 + w')$ for some weight w' and subdistribution Δ''. Therefore,

$$
\begin{aligned}
& \mathbb{P}^{\delta,\mathsf{pp},\mathbf{w}}(s) \\
= \; & \mathbf{w} \centerdot \langle w, \Delta' \rangle \\
= \; & \mathbf{w} \centerdot \langle \delta(w_1 + w'), \delta\Delta'' \rangle \\
= \; & \delta(\mathbf{w} \centerdot \langle w_1, \varepsilon \rangle + \mathbf{w} \centerdot \langle w', \Delta'' \rangle) \\
= \; & \delta(\mathbf{w} \centerdot \langle w_1, \varepsilon \rangle + \mathbb{P}^{\delta,\mathsf{pp},\mathbf{w}}(\Delta_1)) \\
= \; & F^{\delta,\mathsf{pp},\mathbf{w}}(\mathbb{P}^{\delta,\mathsf{pp},\mathbf{w}})(s)
\end{aligned}
$$

$\square$

The following proposition is a discounted version of Theorem 2.19, where the static policy and payoff function are stated with respect to a discount factor that should be strictly less than 1.

Proposition B.10 Let $\delta \in [0, 1)$ be a discount and $\mathbf{w}$ a weight function. If pp is a max-seeking static policy with respect to δ and $\mathbf{w}$, then $\mathbb{P}^{\delta,\mathbf{w}}_{\max} = \mathbb{P}^{\delta,\mathsf{pp},\mathbf{w}}$.

Proof. By Lemma B.5, the function $F^{\delta,\mathsf{pp},\mathbf{w}}$ has a unique fixed point. By Lemmas B.6 and B.9, both $\mathbb{P}^{\delta,\mathbf{w}}_{\max}$ and $\mathbb{P}^{\delta,\mathsf{pp},\mathbf{w}}$ are fixed points of the same function $F^{\delta,\mathsf{pp},\mathbf{w}}$, which means that $\mathbb{P}^{\delta,\mathbf{w}}_{\max}$ and $\mathbb{P}^{\delta,\mathsf{pp},\mathbf{w}}$ coincide with each other. $\square$

In Proposition B.10 it is crucial to rule out the case $\delta = 1$, as the following example shows.

Example B.11 Consider the wMDP in Figure 8. Suppose we have a weight function $\mathbf{w}$ with $\mathbf{w}(s_0) = 0,\ \mathbf{w}(s_1) = 0,\ \mathbf{w}(s_2) = 1$. Recall that $\mathbf{w}(s_0)$ is the weight applied to the action benefit in transitions; however in the example all action benefits are 0 and therefore they will be more or less ignored. Note that $s_1 \Longrightarrow_{1,0} \overline{s_2}$ and $s_2 \Longrightarrow_{1,0} \overline{s_2}$ and therefore $\mathbb{P}^{1,\mathbf{w}}_{\max}(s_1) = \mathbb{P}^{1,\mathbf{w}}_{\max}(s_2) = 1$. Let us now look at which policies can attain this payoff, in particular for the state s_1.

According to Definition 2.16 there are three different static policies for the wMDP in Figure 8. All three are required to be undefined at state s_2 since it has no derivatives; so we concentrate on s_1. The first policy, pp_1, is also undefined at s_1. However pp_1 is not max-seeking for the discount $\delta = 1$ as it fails condition (1) in Definition B.3.

The second policy, pp_2 maps s_1 to the pair $\langle 0, \overline{s_1}\,\rangle$. Note that pp_2 is max-seeking for the discount $\delta = 1$ as it satisfies both parts of clause (2) in Definition B.3. However the payoff following this policy at state s_1 is 0; intuitively the policy follows the divergent trace continually through state s_1, accumulating the payoff 0. Formally, applying Definition B.8, $\mathbb{P}^{1,\mathsf{pp}_2,\mathbf{w}}(s_1) = 0$. Thus Proposition B.10 is in general false; pp_2 is max-seeking but $\mathbb{P}^{1,\mathbf{w}}_{\max}(s_1) \neq \mathbb{P}^{1,\mathsf{pp}_2,\mathbf{w}}(s_1)$.

Incidently the third possible static policy, pp_3 which maps s_1 to pair $\langle 0, \overline{s_1}\ {}_{1/2}\!\oplus\ \overline{s_2}\,\rangle$ is also max-seeking and it does attain that the maximum payoff. What is more interesting is to examine what happens when the discount δ is strictly less than 1.

If $\delta \in [0, 1)$, then pp_2 is no longer max-seeking. First note that from state s_1 we have the discounted hyper-derivation $\overline{s_1} \overset{\tau}{\Longrightarrow}_{\delta,0} \frac{\delta}{2-\delta} \cdot \overline{s_2}$ because

$$\begin{array}{rcl} \overline{s_1} & = & \overline{s_1} + \varepsilon \\ \overline{s_1} & \xrightarrow{\tau}_0 & \frac{1}{2} \cdot \overline{s_1} + \frac{1}{2} \cdot \overline{s_2} \\ \frac{1}{2} \cdot \overline{s_1} & \xrightarrow{\tau}_0 & \frac{1}{4} \cdot \overline{s_1} + \frac{1}{4} \cdot \overline{s_2} \\ & \vdots & \\ \frac{1}{2^k} \cdot \overline{s_1} & \xrightarrow{\tau}_0 & \frac{1}{2^{k+1}} \cdot \overline{s_1} + \frac{1}{2^{k+1}} \cdot \overline{s_2} \\ & \vdots & \end{array} \tag{29}$$

and $\sum_{k=1}^{\infty} \delta^k \cdot (\frac{1}{2^k} \cdot \overline{s_2}) = \frac{\delta}{2-\delta} \cdot \overline{s_2}$. From state s_2 we have the discounted hyper-derivation $\overline{s_2} \overset{\tau}{\Longrightarrow}_{\delta,0} \overline{s_2}$. Because of these hyper-derivations one can check that the discounted payoff function is given by $\mathbb{P}^{\delta,\mathbf{w}}_{\max}(s_1) = \frac{\delta}{2-\delta}$ and $\mathbb{P}^{\delta,\mathbf{w}}_{\max}(s_2) = 1$. Thus pp_2 is not max-seeking because it fails condition (2)(b) in Definition B.3. Its immediate payoff is $\frac{\delta}{2-\delta}$ which is strictly less than the immediate payoff obtained by following the other possible transition, $s_1 \xrightarrow{\tau}_0 \overline{s_1}\ {}_{1/2}\!\oplus\ \overline{s_2}$; $\mathbb{P}^{\delta,\mathbf{w}}_{\max}(\overline{s_1}\ {}_{1/2}\!\oplus\ \overline{s_2}) = \frac{1}{2} \cdot \frac{\delta}{2-\delta} + \frac{1}{2} \cdot 1 = \frac{1}{2-\delta}$, and $\frac{1}{2-\delta} > \frac{\delta}{2-\delta}$ for $\delta \in [0, 1)$.

Lemma B.4 assures us that some max-seeking policy always exists. In this case, with $\delta \in [0, 1)$, it happens to be unique, namely pp_3. Moreover one can check that by following it the transitions listed in (29) are realised, which yields the discounted hyper-SP-derivation $\overline{s_1} \overset{\tau}{\Longrightarrow}_{\delta,\mathsf{pp}_3,0} \frac{\delta}{2-\delta} \cdot \overline{s_2}$. Therefore, the maximum payoff $\frac{\delta}{2-\delta}$ from state s_1 can be attained; that is $\mathbb{P}^{\delta,\mathsf{pp}_2,\mathbf{w}}(s_1) = \frac{\delta}{2-\delta}$. $\square$

One of the key lemmas in proving the finite generalability theorem is the following, whose proof involves the mathematical concept of bounded continuity of real-valued functions. For convenience of presentation, we delegate the discussion on bounded continuity, culminating in Proposition D.2, to Section D.

Lemma B.12 Suppose $\overline{s} \stackrel{\tau}{\Longrightarrow}_w \Delta'$ with $\langle w, \Delta' \rangle = \sum_{i=0}^{\infty} \langle w_i, \Delta_i^{\times} \rangle$ for some properly related $\Delta_i^{\times}$ and some w_i with $w_0 = 0$. Let $\{\delta_j\}_{j=0}^{\infty}$ be a nondecreasing sequence of discount factors converging to 1. Then for any weight function $\mathbf{w}$ it holds that

$$\mathbf{w} \centerdot \langle w, \Delta' \rangle \ = \ \lim_{j\to\infty} \sum_{i=0}^{\infty} (\delta_j)^i (\mathbf{w} \centerdot \langle w_i, \Delta_i^{\times} \rangle).$$

Proof. We have three cases. If $w = \infty$ and $\mathbf{w}(s_0) > 0$, then it is easy to see that both sides of the equation are equal to ∞. Similarly, if $w = \infty$ and $\mathbf{w}(s_0) < 0$, both sides are equal to $-\infty$. Otherwise, $|\mathbf{w} \cdot \langle w, \Delta' \rangle| < \infty$ and we proceed as follows.

Let $f : \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ be the function defined by $f(i,j) = (\delta_j)^i (\mathbf{w} \centerdot \langle w_i, \Delta_i^{\times} \rangle)$. We check that f satisfies the four conditions in Proposition D.2.

1. f satisfies condition **C1**. For all $i, j_1, j_2 \in \mathbb{N}$, if $j_1 \le j_2$ then $(\delta_{j_1})^i \le (\delta_{j_2})^i$. It follows that

$$|f(i,j_1)| \ = \ |(\delta_{j_1})^i (\mathbf{w} \centerdot \langle w_i, \Delta_i^{\times} \rangle)| \ \le \ |(\delta_{j_2})^i (\mathbf{w} \centerdot \langle w_i, \Delta_i^{\times} \rangle)| \ = \ |f(i,j_2)|.$$

2. f satisfies condition **C2**. For any $i \in \mathbb{N}$, we have

$$\lim_{j\to\infty} |f(i,j)| \ = \ \lim_{j\to\infty} |(\delta_j)^i (\mathbf{w} \centerdot \langle w_i, \Delta_i^{\times} \rangle)| \ = \ |\mathbf{w} \centerdot \langle w_i, \Delta_i^{\times} \rangle|. \tag{30}$$

3. f satisfies condition **C3**. For any $n \in \mathbb{N}$, the partial sum $S_n = \sum_{i=0}^{n} \lim_{j\to\infty} |f(i,j)|$ is bounded because

$$\begin{array}{rl} & \sum_{i=0}^{n} \lim_{j\to\infty} |f(i,j)| \\ = & \sum_{i=0}^{n} |\mathbf{w} \centerdot \langle w_i, \Delta_i^{\times} \rangle| \\ \le & \sum_{i=0}^{\infty} |\mathbf{w} \centerdot \langle w_i, \Delta_i^{\times} \rangle| \\ \le & \sum_{i=0}^{\infty} (w_i + |\Delta_i^{\times}|) \\ = & w + |\Delta'| \end{array}$$

   where the first equality is justified by (30).

4. f satisfies condition **C4**. For any $i, j_1, j_2 \in \mathbb{N}$, if $j_1 \le j_2$ then

$$\begin{array}{rl} & f(i,j_1) + |f(i,j_1)| \\ = & (\delta_{j_1})^i (\mathbf{w} \centerdot \langle w_i, \Delta_i^{\times} \rangle) + |(\delta_{j_1})^i (\mathbf{w} \centerdot \langle w_i, \Delta_i^{\times} \rangle)| \\ = & (\delta_{j_1})^i (\mathbf{w} \centerdot \langle w_i, \Delta_i^{\times} \rangle + |\mathbf{w} \centerdot \langle w_i, \Delta_i^{\times} \rangle|) \\ \le & (\delta_{j_2})^i (\mathbf{w} \centerdot \langle w_i, \Delta_i^{\times} \rangle + |\mathbf{w} \centerdot \langle w_i, \Delta_i^{\times} \rangle|) \\ = & f(i,j_2) + |f(i,j_2)|. \end{array}$$

Therefore, we can use Proposition D.2 to do the following inference.

$$\begin{array}{rl} & \lim_{j\to\infty} \sum_{i=0}^{\infty} (\delta_j)^i (\mathbf{w} \centerdot \langle w_i, \Delta_i^{\times} \rangle) \\ = & \sum_{i=0}^{\infty} \lim_{j\to\infty} (\delta_j)^i (\mathbf{w} \centerdot \langle w_i, \Delta_i^{\times} \rangle) \\ = & \sum_{i=0}^{\infty} \mathbf{w} \centerdot \langle w_i, \Delta_i^{\times} \rangle \\ = & \mathbf{w} \centerdot \sum_{i=0}^{\infty} \langle w_i, \Delta_i^{\times} \rangle \\ = & \mathbf{w} \centerdot \langle w, \Delta' \rangle \end{array}$$

□

Corollary B.13 Let $\{\delta_j\}_{j=0}^{\infty}$ be a nondecreasing sequence of discount factors converging to 1. For any static policy pp and weight function $\mathbf{w}$, it holds that $\mathbb{P}^{1,\mathsf{pp},\mathbf{w}} = \lim_{j\to\infty} \mathbb{P}^{\delta_j,\mathsf{pp},\mathbf{w}}$.

Proof. We need to show that $\mathbb{P}^{1,\mathsf{pp},\mathbf{w}}(s) = \lim_{j\to\infty} \mathbb{P}^{\delta_j,\mathsf{pp},\mathbf{w}}(s)$, for any state s. Note that for any discount δ_j, each state s enables a unique discounted hyper-SP-derivation $\overline{s} \stackrel{\tau}{\Longrightarrow}_{\delta_j,\mathsf{pp},w^j} \Delta^j$ such that $\langle\, w^j, \Delta^j \,\rangle = \sum_{i=0}^{\infty} (\delta_j)^i \langle\, w_i, \Delta_i^{\times} \,\rangle$ for some properly related $\Delta_i^{\times}$ and some w_i with $w_0 = 0$. Let $w = \sum_{i=0}^{\infty} w_i$ and $\Delta' = \sum_{i=0}^{\infty} \Delta_i^{\times}$. We have $\overline{s} \stackrel{\tau}{\Longrightarrow}_{1,\mathsf{pp},w} \Delta'$. Then we can infer that

$$\begin{array}{rl} & \lim_{j\to\infty} \mathbb{P}^{\delta_j,\mathsf{pp},\mathbf{w}}(s) \\ = & \lim_{j\to\infty} \mathbf{w} \centerdot \langle\, w^j, \Delta^j \,\rangle \\ = & \lim_{j\to\infty} \mathbf{w} \centerdot \sum_{i=0}^{\infty} (\delta_j)^i \langle\, w_i, \Delta_i^{\times} \,\rangle \\ = & \lim_{j\to\infty} \sum_{i=0}^{\infty} (\delta_j)^i (\mathbf{w} \centerdot \langle\, w_i, \Delta_i^{\times} \,\rangle) \\ = & \mathbf{w} \centerdot \langle\, w, \Delta' \,\rangle \qquad \text{by Lemma B.12} \\ = & \mathbb{P}^{1,\mathsf{pp},\mathbf{w}}(s) \end{array}$$

□

Theorem B.14 (Theorem 2.19) In a finitary wMDP, for any weight function $\mathbf{w}$ there exists a static policy pp such that $\mathbb{P}^{1,\mathbf{w}}_{\max} = \mathbb{P}^{1,\mathsf{pp},\mathbf{w}}$.

Proof. Let $\mathbf{w}$ be a weight function. By Lemma B.4 and Proposition B.10, for every discount factor $\delta < 1$ there exists a max-seeking static policy with respect to δ and $\mathbf{w}$ such that

$$\mathbb{P}^{\delta,\mathbf{w}}_{\max} = \mathbb{P}^{\delta,\mathsf{pp},\mathbf{w}}. \tag{31}$$

Since the wMDP is finitary, there are finitely many different static policies. There must exist a static policy pp such that (31) holds for infinitely many discount factors. In other words, for every nondecreasing sequence $\{\delta_n\}_{n=0}^{\infty}$ converging to 1, with $\delta_n < 1$ for all $n \geq 0$, there exists a sub-sequence $\{\delta_{n_j}\}_{j=0}^{\infty}$ converging to 1 and a static policy $\mathsf{pp}^\star$ such that

$$\mathbb{P}^{\delta_{n_j},\mathbf{w}}_{\max} = \mathbb{P}^{\delta_{n_j},\mathsf{pp}^\star,\mathbf{w}} \qquad \text{for all } j \geq 0. \tag{32}$$

For any state s, we infer as follows.

$$\begin{array}{rl} & \mathbb{P}^{1,\mathbf{w}}_{\max}(s) \\ = & \sup\{\mathbf{w} \centerdot \langle\, w, \Delta' \,\rangle \mid \overline{s} \stackrel{\tau}{\Longrightarrow}_w \Delta'\} \\ = & \sup\{\lim_{j\to\infty} \sum_{i=0}^{\infty} (\delta_{n_j})^i (\mathbf{w} \centerdot \langle\, w_i, \Delta_i^{\times} \,\rangle) \mid \overline{s} \stackrel{\tau}{\Longrightarrow}_w \Delta' \text{ with } \langle\, w, \Delta' \,\rangle = \sum_{i=0}^{\infty} \langle\, w_i, \Delta_i^{\times} \,\rangle\} \\ & \qquad \text{[by Lemma B.12]} \\ \leq & \lim_{j\to\infty} \sup\{\sum_{i=0}^{\infty} (\delta_{n_j})^i (\mathbf{w} \centerdot \langle\, w_i, \Delta_i^{\times} \,\rangle) \mid \overline{s} \stackrel{\tau}{\Longrightarrow}_w \Delta' \text{ with } \langle\, w, \Delta' \,\rangle = \sum_{i=0}^{\infty} \langle\, w_i, \Delta_i^{\times} \,\rangle\} \\ = & \lim_{j\to\infty} \sup\{\mathbf{w} \centerdot \sum_{i=0}^{\infty} (\delta_{n_j})^i (\langle\, w_i, \Delta_i^{\times} \,\rangle) \mid \overline{s} \stackrel{\tau}{\Longrightarrow}_w \Delta' \text{ with } \langle\, w, \Delta' \,\rangle = \sum_{i=0}^{\infty} \langle\, w_i, \Delta_i^{\times} \,\rangle\} \\ = & \lim_{j\to\infty} \sup\{\mathbf{w} \centerdot \langle\, w', \Delta'' \,\rangle \mid \overline{s} \stackrel{\tau}{\Longrightarrow}_{\delta_{n_j},w'} \Delta''\} \\ = & \lim_{j\to\infty} \mathbb{P}^{\delta_{n_j},\mathbf{w}}_{\max}(s) \\ = & \lim_{j\to\infty} \mathbb{P}^{\delta_{n_j},\mathsf{pp}^\star,\mathbf{w}}(s) \qquad \text{[by (32)]} \\ = & \mathbb{P}^{1,\mathsf{pp}^\star,\mathbf{w}}(s) \qquad \text{[by Corollary B.13]} \end{array}$$

The other direction, $\mathbb{P}^{1,\mathbf{w}}_{\max}(s) \geq \mathbb{P}^{1,\mathsf{pp}^\star,\mathbf{w}}(s)$, is trivial in view of Definitions B.2 and B.8. □

C Compactness arguments

In this appendix we give the detailed proofs of the two results from Section 3.2, Proposition 3.10 and Proposition 3.12 which rely on compactness arguments.

Corollary C.1 Let Δ be a subdistribution in a bounded wMDP. The set $\{\langle w, \Delta' \rangle \mid \Delta \stackrel{\tau}{\Longrightarrow}_w \Delta'\}$ is compact and convex.

Proof. Let $\mathsf{pp}_1, ..., \mathsf{pp}_n$ ($n \geq 1$) be all the static policies in the bounded wMDP. Each policy determines a hyper-SP-derivation $\Delta \stackrel{\tau}{\Longrightarrow}_{\mathsf{pp}_i, w_i} \Delta'_i$. By Theorem 2.27, the weight w_i is finite. Let C be the convex closure of $\{\langle w_i, \Delta_i \rangle \mid 1 \leq i \leq n\}$. Let D be the set $\{\langle w, \Delta' \rangle \mid \Delta \stackrel{\tau}{\Longrightarrow}_w \Delta'\}$. By Theorem 2.20 we have $D \subseteq C$. On the other hand, it is easy to see from Lemma 2.11(1) that D is convex and thus $C \subseteq D$. Consequently, D coincides with C, the convex closure of a finite set. Therefore, it is Cauchy closed and bounded, thus being compact. □

In order to extend the above result to the relation $\stackrel{\alpha}{\Longrightarrow}$, for any $\alpha \in \mathsf{Act}$, we need some preliminary concepts.

Definition C.2 A subset $D \subseteq \mathbb{R} \times \mathcal{D}_{sub}(S)$ is said to be *finitely generable* whenever there is some finite set $F \subseteq \mathbb{R} \times \mathcal{D}_{sub}(S)$ such that $D = \updownarrow F$. Then a relation $\mathcal{R} \subseteq X \times \mathbb{R} \times \mathcal{D}_{sub}(S)$ is said to be *finitely generable* if for every x in X the set $x \cdot \mathcal{R}$ is finitely generable. □

Lemma C.3 If a set is finitely generable, then it is compact and convex.

Proof. A direct consequence of the definition of finite generability. □

Definition C.4 Let $\mathcal{R}_1, \mathcal{R}_2 \in \mathcal{D}_{sub}(S) \times (\mathbb{R} \times \mathcal{D}_{sub}(S))$ be two relations. We define their composition $\mathcal{R}_1;\mathcal{R}_2$ by letting $\Delta\ \mathcal{R}_1;\mathcal{R}_2\ \langle w, \Theta \rangle$ if there are some w_1, w_2, Θ' such that $\Delta\ \mathcal{R}_1\ \langle w_1, \Theta' \rangle$ and $\Theta'\ \mathcal{R}_2\ \langle w_2, \Theta \rangle$ with $w_1 + w_2 = w$. □

Lemma C.5 Let $\mathcal{R}_1, \mathcal{R}_2 \subseteq \mathcal{D}_{sub}(S) \times (\mathbb{R} \times \mathcal{D}_{sub}(S))$ be finitely generable. Moreover, $\mathcal{R}_2$ is both linear and decomposable. Then the relation $\mathcal{R}_1;\mathcal{R}_2$ is finitely generable.

Proof. Let $\mathcal{B}^i_\Phi$ be a finite set of pairs of reals and subdistributions such that $\Phi \cdot \mathcal{R}_i = \updownarrow \mathcal{B}^i_\Phi$ for $i = 1, 2$. By exploiting the linearity and decomposability of $\mathcal{R}_2$, we can check that

$$\Delta \cdot \mathcal{R}_1;\mathcal{R}_2 \ = \ \updownarrow \cup \{ \langle w, \varepsilon \rangle + \mathcal{B}^2_\Theta \mid \langle w, \Theta \rangle \in \mathcal{B}^1_\Delta \}.$$

where $\langle w, \varepsilon \rangle + \mathcal{B}^2_\Theta$ stands for the set $\{\langle w, \varepsilon \rangle + \langle v, \Gamma \rangle \mid \langle v, \Gamma \rangle \in \mathcal{B}^2_\Theta\}$. □

We are now ready to establish Proposition 3.10; it follows from this slightly more general result:

Lemma C.6 Let Δ be a subdistribution in a bounded wMDP. The set $\{\langle w, \Delta' \rangle \mid \Delta \stackrel{\alpha}{\Longrightarrow}_w \Delta'\}$ is compact and convex.

Proof. The relation $\stackrel{\alpha}{\Longrightarrow}$ is a composition of three stages: $\stackrel{\tau}{\Longrightarrow};\stackrel{\alpha}{\longrightarrow};\stackrel{\tau}{\Longrightarrow}$. In the proof of Corollary C.1 we have shown that $\stackrel{\tau}{\Longrightarrow}$ is finitely generable. Since a bounded wMDP is finitary, the relation $\stackrel{\alpha}{\longrightarrow}$ is also finitely generable. We observe that $\stackrel{\alpha}{\longrightarrow}$ is both linear and decomposable, so is $\stackrel{\tau}{\Longrightarrow}$ by Lemma 2.11. It follows from Proposition C.5 that $\stackrel{\alpha}{\Longrightarrow}$ is finitely generable. By Lemma C.3 we have that $\stackrel{\alpha}{\Longrightarrow}$ is compact and convex. □

Corollary C.7 In a bounded wMDP, the relation $\stackrel{\alpha}{\Longrightarrow}$ is the lifting of the compact and convex relation $\stackrel{\alpha}{\Longrightarrow}_S$, where $s \stackrel{\alpha}{\Longrightarrow}_S \Delta$ means $\overline{s} \stackrel{\alpha}{\Longrightarrow} \Delta$.

Proof. The relation $\stackrel{\alpha}{\Longrightarrow}_S$ is $\stackrel{\alpha}{\Longrightarrow}$ restricted to point distributions. We have shown that $\stackrel{\alpha}{\Longrightarrow}$ is compact and convex in Lemma C.6. Therefore, $\stackrel{\alpha}{\Longrightarrow}_S$ is compact and convex. Its lifting coincides with $\stackrel{\alpha}{\Longrightarrow}$, which follows from Proposition 2.11. □

Our next step is to show that each of the relations $\lhd^k$ is closed. This requires some results to be first established.

Lemma C.8 If $\mathcal{R} \subseteq S \times (\mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S))$ is compact, then so is its set of choice functions $\mathbf{Ch}(\mathcal{R})$.

Proof. Suppose that $\mathcal{R}$ is compact, that is closed and bounded. It is straightforward to show that $\mathbf{Ch}(\mathcal{R})$, under the metric defined on page 22, is therefore also closed and bounded. It follows that $\mathbf{Ch}(\mathcal{R})$ forms a complete metric space. Moreover, since $\mathcal{R}$ is bounded, $\mathbf{Ch}(\mathcal{R})$ is also totally bounded. Therefore, $\mathbf{Ch}(\mathcal{R})$ is compact, for a metric space is compact if and only if it is complete and totally bounded. □

Let $\beta(x)$ be a predicate with variable x ranging over some set X. We use the notation $\beta(\bullet)$ to represent the set $\{x \in X \mid \beta(x)\}$.

Lemma C.9 Suppose there is a continuous function $g : \mathbb{R}^2_{\geq 0} \to \mathbb{R}$ and two convex relations $\mathcal{R}_1, \mathcal{R}_2 \subseteq S \times (\mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S))$ such that $\mathcal{R}_1$ is compact and $\mathcal{R}_2$ is closed. Then the set

$$Z = \{ \langle r, \Theta \rangle \mid r \in \mathbb{R}_{\geq 0} \text{ and } \exists w \in \mathbb{R}_{\geq 0} : (\Theta \ \overline{\mathcal{R}_1} \ \langle w, \bullet \rangle) \cap (\Delta \ \overline{\mathcal{R}_2} \ \langle g(r,w), \bullet \rangle) \neq \emptyset \}$$

is closed.

Proof. We will use the continuous function $\mathcal{E}$, defined in the proof of Theorem 3.14; recall that it also maps closed sets to closed sets.

Let $r, w \in \mathbb{R}_{\geq 0}$, $\Theta \in \mathcal{D}_{sub}(S)$, and $f \in S \to \mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S)$. Then define the following four functions

$$\begin{aligned} H_1 : &\ \langle \langle r, \Theta \rangle, f \rangle \mapsto \langle r, \langle \Theta, f \rangle \rangle \\ H_2 : &\ \langle r, \langle w, \Theta \rangle \rangle \mapsto \langle \langle r, w \rangle, \Theta \rangle \\ F_{\mathcal{E}} : &\ \langle r, \langle \Theta, f \rangle \rangle \mapsto \langle r, \mathcal{E}(\Theta, f) \rangle \\ G_g : &\ \langle \langle r, w \rangle, \Theta \rangle \mapsto \langle g(r,w), \Theta \rangle \end{aligned}$$

which are continuous. Finally let

$$Z' = \pi_1(H_1^{-1} \circ F_{\mathcal{E}}^{-1} \circ H_2^{-1} \circ G_g^{-1} \circ \mathcal{E}(\{\Delta\} \times \mathbf{Ch}(\mathcal{R}_2)) \cap (\mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S)) \times \mathbf{Ch}(\mathcal{R}_1))$$

where $\pi_1 : (\mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S)) \times \mathbf{Ch}(\mathcal{R}_1) \to \mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S)$ is the projection onto the first component of a pair. Since $\mathcal{R}_2$ is closed, it easily follows that $\mathbf{Ch}(\mathcal{R}_2)$ is also closed. Then the product $\{\Delta\} \times \mathbf{Ch}(\mathcal{R}_2)$ is closed. Its image under the closed function $\mathcal{E}$ is also closed. Since the four functions $G_g, H_2, F_{\mathcal{E}}, H_1$ are continuous and the inverse image of a closed set is closed, we know that $H_1^{-1} \circ F_{\mathcal{E}}^{-1} \circ H_2^{-1} \circ G_g^{-1} \circ \mathcal{E}(\{\Delta\} \times \mathbf{Ch}(\mathcal{R}_2))$ is closed. On the other hand, since $\mathcal{R}_1$ is compact, by Lemma C.8 the set of choice functions $\mathbf{Ch}(\mathcal{R}_1)$ is compact. It is then easy to see that $(\mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S)) \times \mathbf{Ch}(\mathcal{R}_1)$ is closed. It follows that the intersection of two closed sets

$$H_1^{-1} \circ F_{\mathcal{E}}^{-1} \circ H_2^{-1} \circ G_g^{-1} \circ \mathcal{E}(\{\Delta\} \times \mathbf{Ch}(\mathcal{R}_2)) \cap (\mathbb{R}_{\geq 0} \times \mathcal{D}_{sub}(S)) \times \mathbf{Ch}(\mathcal{R}_1)$$

is closed. By the tube lemma in topology theory, the projection π_1 is closed[1]. Therefore, we have that Z' is closed.

We now show that $Z = Z'$.

$$
\begin{array}{ll}
 & \langle r, \Theta \rangle \in Z' \\
\text{iff} & \langle \langle r, \Theta \rangle, f_1 \rangle \in H_1^{-1} \circ F_{\mathcal{E}}^{-1} \circ H_2^{-1} \circ G_g^{-1} \circ \mathcal{E}(\{\Delta\} \times \mathbf{Ch}(\mathcal{R}_2)) \text{ for some } f_1 \in \mathbf{Ch}(\mathcal{R}_1) \\
\text{iff} & \langle r, \langle \Theta, f_1 \rangle \rangle \in F_{\mathcal{E}}^{-1} \circ H_2^{-1} \circ G_g^{-1} \circ \mathcal{E}(\{\Delta\} \times \mathbf{Ch}(\mathcal{R}_2)) \text{ for some } f_1 \in \mathbf{Ch}(\mathcal{R}_1) \\
\text{iff} & \langle r, \mathcal{E}(\Theta, f_1) \rangle \in H_2^{-1} \circ G_g^{-1} \circ \mathcal{E}(\{\Delta\} \times \mathbf{Ch}(\mathcal{R}_2)) \text{ for some } f_1 \in \mathbf{Ch}(\mathcal{R}_1) \\
\text{iff} & \langle r, \mathrm{Exp}_{\Theta}(f_1) \rangle \in H_2^{-1} \circ G_g^{-1} \circ \mathcal{E}(\{\Delta\} \times \mathbf{Ch}(\mathcal{R}_2)) \text{ for some } f_1 \in \mathbf{Ch}(\mathcal{R}_1) \\
\text{iff} & \Theta \, \overline{R_1} \, \langle w, \Theta' \rangle \text{ and } \langle r, \langle w, \Theta' \rangle \rangle \in H_2^{-1} \circ G_g^{-1} \circ \mathcal{E}(\{\Delta\} \times \mathbf{Ch}(\mathcal{R}_2)) \text{ for some } \langle w, \Theta' \rangle \\
\text{iff} & \Theta \, \overline{R_1} \, \langle w, \Theta' \rangle \text{ and } \langle \langle r, w \rangle, \Theta' \rangle \in G_g^{-1} \circ \mathcal{E}(\{\Delta\} \times \mathbf{Ch}(\mathcal{R}_2) \text{ for some } \langle w, \Theta' \rangle \\
\text{iff} & \Theta \, \overline{R_1} \, \langle w, \Theta' \rangle \text{ and } \langle g(r, w), \Theta' \rangle \in \mathcal{E}(\{\Delta\} \times \mathbf{Ch}(\mathcal{R}_2) \text{ for some } \langle w, \Theta' \rangle \\
\text{iff} & \Theta \, \overline{R_1} \, \langle w, \Theta' \rangle \text{ and } \Delta \, \overline{\mathcal{R}_2} \, \langle g(r, w), \Theta' \rangle \text{ for some } \langle w, \Theta' \rangle \\
\text{iff} & (\Theta \, \overline{\mathcal{R}_1} \, \langle w, \bullet \rangle) \cap (\Delta \, \overline{\mathcal{R}_2} \, \langle g(r, w), \bullet \rangle) \neq \emptyset \text{ for some } w \\
\text{iff} & \langle r, \Theta \rangle \in Z.
\end{array}
$$

□

This lemma enables us to establish the second requirement of the appendix:

Proposition C.10 [Proposition 3.12] In a bounded wMDP, for every $k \in \mathbb{N}$, the relation $\lhd^k$ is closed and convex.

Proof. By induction on k. For $k = 0$ the result is obvious. So let us assume that $\lhd^k$ is closed and convex. We have to show that

$$
s \cdot \lhd^{(k+1)} \text{ is closed and convex, for every state } s \tag{33}
$$

For every $\alpha,\ v,\ \Delta$ let

$$
G_{\alpha,v,\Delta} = \{ \langle r, \Theta \rangle \mid r \in \mathbb{R}_{\geq 0} \text{ and } \exists w \in \mathbb{R}_{\geq 0} : (\Theta \cdot \overset{\alpha}{\Longrightarrow}_w) \cap (\Delta \cdot \overline{\lhd^k}_{r+w-v}) \neq \emptyset \}.
$$

By Corollary C.7, the relation $\overset{\alpha}{\Longrightarrow}$ is lifted from a compact and convex relation. By induction hypothesis we know that $\lhd^k$ is closed and convex. The function $g : \mathbb{R}^2_{\geq 0} \to \mathbb{R}_{\geq 0}$ given by $g(r, w) = r + w - v$ is continuous. So we can appeal to Lemma C.9 and conclude that each $G_{\alpha,v,\Delta}$ is closed. By Definition 2.2 it is also easy to see that $G_{\alpha,v,\Delta}$ is convex. But it follows that $s \cdot \lhd^{(k+1)}$ is also closed and convex as it can be written as

$$
\cap \{ G_{\alpha,v,\Delta} \mid s \xrightarrow{\alpha}_v \Delta \}.
$$

□

[1] In general, the projection $\pi_1 : X \times Y \to X$ is not closed. For example, if $X = Y = \mathbb{R}$, then π_1 maps the closed set $\{\langle x, y \rangle \in \mathbb{R}^2 \mid xy = 1\}$ into $\mathbb{R} \backslash \{0\}$ which is not closed. However, the tube lemma tells us that if X is any topological space and Y a compact space, then the projection map π_1 is closed.

D Bounded continuity

In this section we study the property of bounded continuity of real-valued binary functions, which plays a crucial role in the proof of Lemma B.12. We first consider nonnegative functions.

Proposition D.1 [Bounded continuity - nonnegative function] Given a function $f : \mathbb{N} \times \mathbb{N} \to \mathbb{R}_{\geq 0}$ which satisfies the following conditions

C1. f is monotonic on the second parameter, i.e. $j_1 \leq j_2$ implies $f(i, j_1) \leq f(i, j_2)$ for all $i, j_1, j_2 \in \mathbb{N}$.

C2. For any $i \in \mathbb{N}$, the limit $\lim_{j\to\infty} f(i, j)$ exists.

C3. For any $n \in \mathbb{N}$, the partial sum $S_n = \sum_{i=0}^{n} \lim_{j\to\infty} f(i, j)$ is bounded, i.e. there exists some $c \in \mathbb{R}_{\geq 0}$ such that $S_n \leq c$ for all $n \geq 0$.

then it holds that

$$\sum_{i=0}^{\infty} \lim_{j\to\infty} f(i,j) \;=\; \lim_{j\to\infty} \sum_{i=0}^{\infty} f(i,j).$$

Proof. Let ϵ be any positive real number. By **C3** the sequence $\{S_n\}_{n=0}^{\infty}$ is bounded and it is nondecreasing, so it converges to $\sum_{i=0}^{\infty} \lim_{j\to\infty} f(i, j)$. Then there exists some $n_\epsilon \in \mathbb{N}$ such that

$$0 \;\leq\; \sum_{i=0}^{\infty} \lim_{j\to\infty} f(i,j) - \sum_{i=0}^{n_\epsilon} \lim_{j\to\infty} f(i,j) \;\leq\; \frac{\epsilon}{2}. \tag{34}$$

By **C1** and **C2**, for any $i \in \mathbb{N}$, the sequence $\{f(i, j)\}_{j=0}^{\infty}$ is nondecreasing and converges to $\lim_{j\to\infty} f(i, j)$. Therefore, for each $i \in \mathbb{N}$, there exists some $m_{i,\epsilon,n_\epsilon} \in \mathbb{N}$ such that

$$\forall j \geq m_{i,\epsilon,n_\epsilon} : \qquad 0 \;\leq\; \lim_{j'\to\infty} f(i,j') - f(i,j) \;\leq\; \frac{\epsilon}{2(n_\epsilon + 1)}. \tag{35}$$

Let $m_\epsilon = \max\{m_{i,\epsilon,n_\epsilon} \mid 0 \leq i \leq n_\epsilon\}$. It follows from (35) that

$$\forall j \geq m_\epsilon : \qquad 0 \;\leq\; \sum_{i=0}^{n_\epsilon} \lim_{j'\to\infty} f(i,j') - \sum_{i=0}^{n_\epsilon} f(i,j) \;\leq\; \frac{\epsilon}{2}. \tag{36}$$

By summing up (34) and (36), we obtain

$$\forall j \geq m_\epsilon : \qquad 0 \;\leq\; \sum_{i=0}^{\infty} \lim_{j'\to\infty} f(i,j') - \sum_{i=0}^{n_\epsilon} f(i,j) \;\leq\; \epsilon. \tag{37}$$

By **C1** and **C2**, we have that $f(i, j) \leq \lim_{j'\to\infty} f(i, j')$ for any $i, j \in \mathbb{N}$. So for any $j, n \in \mathbb{N}$ the partial sum $\sum_{i=0}^{n} f(i, j)$ is bounded as

$$\sum_{i=0}^{n} f(i,j) \;\leq\; \sum_{i=0}^{n} \lim_{j'\to\infty} f(i,j') \;\leq\; c$$

according to **C3**. Thus it converges to $\sum_{i=0}^{\infty} f(i,j)$. Then for any $j \in \mathbb{N}$ there exists some $n_{j,\epsilon}$ such that

$$\forall n \geq n_{j,\epsilon}: \qquad 0 \;\leq\; \sum_{i=0}^{\infty} f(i,j) - \sum_{i=0}^{n} f(i,j) \;\leq\; \epsilon. \tag{38}$$

Now consider the particular case that $j = m_\epsilon$. Let $N_\epsilon = \max\{n_\epsilon, n_{m_\epsilon,\epsilon}\}$. We know from (37)

$$0 \;\leq\; \sum_{i=0}^{\infty} \lim_{j\to\infty} f(i,j) - \sum_{i=0}^{N_\epsilon} f(i,m_\epsilon) \;\leq\; \epsilon. \tag{39}$$

From (38) we infer that

$$-\epsilon \;\leq\; \sum_{i=0}^{N_\epsilon} f(i,m_\epsilon) - \sum_{i=0}^{\infty} f(i,m_\epsilon) \;\leq\; 0. \tag{40}$$

By summing up (39) and (40), we derive that

$$-\epsilon \;\leq\; \sum_{i=0}^{\infty} \lim_{j\to\infty} f(i,j) - \sum_{i=0}^{\infty} f(i,m_\epsilon) \;\leq\; \epsilon. \tag{41}$$

We conclude from (41) that

$$\lim_{j\to\infty} \sum_{i=0}^{\infty} f(i,j) \;=\; \sum_{i=0}^{\infty} \lim_{j\to\infty} f(i,j).$$

□

Proposition D.2 [Bounded continuity - general function] Given a function $f : \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ which satisfies the following conditions

C1. For all $i, j_1, j_2 \in \mathbb{N}$, we have $j_1 \leq j_2$ implies $|f(i,j_1)| \leq |f(i,j_2)|$.

C2. For any $i \in \mathbb{N}$, the limit $\lim_{j\to\infty} |f(i,j)|$ exists.

C3. For any $n \in \mathbb{N}$, the partial sum $S_n = \sum_{i=0}^{n} \lim_{j\to\infty} |f(i,j)|$ is bounded, i.e. there exists some $c \in \mathbb{R}_{\geq 0}$ such that $S_n \leq c$ for all $n \geq 0$.

C4. For all $i, j_1, j_2 \in \mathbb{N}$, we have $j_1 \leq j_2$ implies $f(i,j_1) + |f(i,j_1)| \leq f(i,j_2) + |f(i,j_2)|$.

then it holds that

$$\sum_{i=0}^{\infty} \lim_{j\to\infty} f(i,j) \;=\; \lim_{j\to\infty} \sum_{i=0}^{\infty} f(i,j).$$

Proof. For any $i, j \in \mathbb{N}$, we have $f(i,j) + |f(i,j)| \leq 2|f(i,j)| \leq 2\ \lim_{j\to\infty} |f(i,j)|$ by **C1** and **C2**. Therefore, for any $i \in \mathbb{N}$, the sequence $\{f(i,j) + |f(i,j)|\}_{j=0}^{\infty}$ has a limit. That is, we have the condition

C5. for any $i \in \mathbb{N}$, the limit $\lim_{j\to\infty}(f(i,j) + |f(i,j)|)$ exists.

Moreover, it holds that $\lim_{j\to\infty}(f(i,j)+|f(i,j)|) \;\leq\; 2\;\lim_{j\to\infty}|f(i,j)|$. It follows that

C6. for any $n \in \mathbb{N}$, the partial sum $\sum_{i=0}^{n}\lim_{j\to\infty}(f(i,j)+|f(i,j)|) \;\leq\; 2\;\sum_{i=0}^{n}\lim_{j\to\infty}|f(i,j)| \leq 2c$.

By Proposition D.1 and conditions **C1**, **C2** and **C3**, we infer that

$$\lim_{j\to\infty}\sum_{i=0}^{\infty}|f(i,j)| \;=\; \sum_{i=0}^{\infty}\lim_{j\to\infty}|f(i,j)|. \tag{42}$$

By Proposition D.1 and conditions **C4**, **C5** and **C6**, we infer that

$$\lim_{j\to\infty}\sum_{i=0}^{\infty}(f(i,j)+|f(i,j)|) \;=\; \sum_{i=0}^{\infty}\lim_{j\to\infty}(f(i,j)+|f(i,j)|). \tag{43}$$

Since $\sum_{i=0}^{\infty}f(i,j) \;=\; \sum_{i=0}^{\infty}(f(i,j)+|f(i,j)|)-\sum_{i=0}^{\infty}|f(i,j)|$, we then have

$$\begin{array}{rcl}
\lim_{j\to\infty}\sum_{i=0}^{\infty}f(i,j) & = & \lim_{j\to\infty}(\sum_{i=0}^{\infty}(f(i,j)+|f(i,j)|)-\sum_{i=0}^{\infty}|f(i,j)|) \\
 & & \text{[existence of the two limits by (42) and (43)]} \\
 & = & \lim_{j\to\infty}\sum_{i=0}^{\infty}(f(i,j)+|f(i,j)|)-\lim_{j\to\infty}\sum_{i=0}^{\infty}|f(i,j)| \\
 & & \text{[by (42) and (43)]} \\
 & = & \sum_{i=0}^{\infty}\lim_{j\to\infty}(f(i,j)+|f(i,j)|)-\sum_{i=0}^{\infty}\lim_{j\to\infty}|f(i,j)| \\
 & = & \sum_{i=0}^{\infty}(\lim_{j\to\infty}(f(i,j)+|f(i,j)|)-\lim_{j\to\infty}|f(i,j)|) \\
 & = & \sum_{i=0}^{\infty}\lim_{j\to\infty}(f(i,j)+|f(i,j)|-|f(i,j)|) \\
 & = & \sum_{i=0}^{\infty}\lim_{j\to\infty}f(i,j)
\end{array}$$

□

E Completeness for benefits testing

Here we outline the details for the proof of Theorem 4.12, which underlies the completeness of benefits testing for amortised weighted simulation. They are a variation on the proof of the corresponding result in [DvGHM09].

Lemma E.1 Let Δ be a distribution and T, T_i be tests.

1. $o \in \textbf{Outcomes}(\Delta \,||\, \omega)$ iff $o = \langle\, 0, \vec{\omega}\,\rangle$.
2. $o \in \textbf{Outcomes}(\Delta \,||\, a_0.T)$ and $o \neq \vec{0}$ iff $\Delta \stackrel{a}{\Longrightarrow}_w \Delta'$ and $o = o' + \langle\, w, \vec{0}\,\rangle$ for some $o' \in \textbf{Outcomes}(\Delta' \,||\, T)$.
3. $o \in \textbf{Outcomes}(\Delta \,||\, T_1 {}_p\!\oplus T_2)$ iff $o = p_i \cdot o_1 + (1-p)\cdot o_2$ for some $o_i \in \textbf{Outcomes}(\Delta \,||\, T_i)$.
4. $o \in \textbf{Outcomes}(\Delta \,||\, (\tau_0.T_1 + \tau_0.T_2))$ if there are $q \in [0,1]$, weight w and distributions Δ_1, Δ_2 such that $\Delta \stackrel{\tau}{\Longrightarrow}_w q\cdot\Delta_1 + (1-q)\cdot\Delta_2$ and $o = q\cdot o_1 + (1-q)\cdot o_2 + \langle\, w, \vec{0}\,\rangle$ for certain $o_i \in \textbf{Outcomes}(\Delta_i \,||\, T_i)$.

Proof.

1. The states in the support of $\Delta \mid\mid \omega$ has a unique outgoing transition labelled by ω. Therefore, $\Delta \mid\mid \omega$ is the unique extreme derivative of itself. As $\mathbf{Success}(\Delta \mid\mid \omega) = \vec{\omega}$, we have $\mathbf{Outcomes}(\Delta \mid\mid \omega) = \{\langle 0, \vec{\omega} \rangle\}$.

2. ($\Leftarrow$) Suppose $\Delta \stackrel{a}{\Longrightarrow}_w \Delta'$, $o' \in \mathbf{Outcomes}(\Delta' \mid\mid T)$ and $o = o' + \langle w, \vec{0} \rangle$. With loss of generality we may assume that $\Delta \stackrel{\tau}{\Longrightarrow}_{w_1} \Delta_1 \stackrel{a}{\longrightarrow}_{w_2} \Delta_2 \stackrel{\tau}{\Longrightarrow}_{w_3} \Delta'$ with $w = w_1 + w_2 + w_3$. Using Lemma 3.3, we have that $\Delta \mid\mid a_0.T \stackrel{\tau}{\Longrightarrow}_{w_1} \Delta_1 \mid\mid a_0.T \stackrel{a}{\longrightarrow}_{w_2} \Delta_2 \mid\mid T \stackrel{\tau}{\Longrightarrow}_{w_3} \Delta' \mid\mid T$. It follows that $o \in \mathbf{Outcomes}(\Delta \mid\mid a_0.T)$.

   ($\Rightarrow$) Suppose $o \in \mathbf{Outcomes}(\Delta \mid\mid a_0.T)$ and $o \neq \vec{0}$. Then there must be a Δ' such that $\Delta \stackrel{\tau}{\Longrightarrow}_{w_1} \stackrel{a}{\longrightarrow}_{w_2} \Delta'$ and some $o' \in \mathbf{Outcomes}(\Delta' \mid\mid T)$ exists with $o = o' + \langle w_1 + w_2, \vec{0} \rangle$.

3. ($\Leftarrow$) Suppose $o_i \in \mathbf{Outcomes}(\Delta \mid\mid T_i)$ for $i = 1, 2$. Then $\Delta \mid\mid T_i \stackrel{\tau}{\Longrightarrow}_{w_i} \Gamma_i$ for some stable Γ_i with $o_i = \langle w_i, \mathbf{Success}(\Gamma_i) \rangle$. By Proposition 2.11(4) we have $\Delta \mid\mid T_1 \,{}_p\!\oplus\, T_2 \stackrel{\tau}{\Longrightarrow}_w \Gamma$ with $w = pw_1 + (1-p)w_2$ and $\Gamma = p \cdot \Gamma_1 + (1-p) \cdot \Gamma_2$. Clearly, Γ is also stable and $\mathbf{Success}(\Gamma) = p \cdot \mathbf{Success}(\Gamma_1) + (1-p) \cdot \mathbf{Success}(\Gamma_2)$. Hence, $o \in \mathbf{Outcomes}(\Delta \mid\mid T_1 \,{}_p\!\oplus\, T_2)$.

   ($\Rightarrow$) Suppose $o \in \mathbf{Outcomes}(\Delta \mid\mid T_1 \,{}_p\!\oplus\, T_2)$. Then there is a stable Γ such that $\Delta \mid\mid T_1 \,{}_p\!\oplus\, T_2 \stackrel{\tau}{\Longrightarrow}_w \Gamma$ and $o = \langle w, \mathbf{Success}(\Gamma) \rangle$. By Proposition 2.11(3) there are Γ_i for $i = 1, 2$, such that $\Delta \mid\mid T_i \stackrel{\tau}{\Longrightarrow}_{w_i} \Gamma_i$ and $w = pw_1 + (1-p)w_2$ and $\Gamma = p \cdot \Gamma_1 + (1-p) \cdot \Gamma_2$. As Γ_1 and Γ_2 are stable, we have $\langle w_i, \mathbf{Success}(\Gamma_i) \rangle \in \mathbf{Outcomes}(\Delta \mid\mid T_i)$. Moreover, $o = p \cdot \langle w_1, \mathbf{Success}(\Gamma_1) \rangle + (1-p) \cdot \langle w_2, \mathbf{Success}(\Gamma_2) \rangle$.

4. Suppose $\Delta \stackrel{\tau}{\Longrightarrow}_w q \cdot \Delta_1 + (1-q) \cdot \Delta_2$ and $o_i \in \mathbf{Outcomes}(\Delta_i \mid\mid T_i)$. Then there are stable Γ_i with $\Delta_i \mid\mid T_i \stackrel{\tau}{\Longrightarrow}_{w_i} \Gamma_i$ and $o_i = \langle w_i, \mathbf{Success}(\Gamma_i) \rangle$. Using Lemma 3.3, we have that $\Delta \mid\mid (\tau_0.T_1 + \tau_0.T_2) \stackrel{\tau}{\Longrightarrow}_w q \cdot (\Delta_1 \mid\mid (\tau_0.T_1 + \tau_0.T_2)) + (1-q) \cdot (\Delta_2 \mid\mid (\tau_0.T_1 + \tau_0.T_2)) \stackrel{\tau}{\longrightarrow}_0 q \cdot \Delta_1 \mid\mid T_1 + (1-q) \cdot \Delta_2 \mid\mid T_2 \stackrel{\tau}{\Longrightarrow}_{w'} \Gamma$ with $w' = qw_1 + (1-q)w_2$ and $\Gamma = q \cdot \Gamma_1 + (1-q) \cdot \Gamma_2$. Clearly, Γ is stable and $\mathbf{Success}(\Gamma) = q \cdot \mathbf{Success}(\Gamma_1) + (1-q) \cdot \mathbf{Success}(\Gamma_2)$. Hence, $q \cdot o_1 + (1-q) \cdot o_2 + \langle w, \vec{0} \rangle \in \mathbf{Outcomes}(\Delta \mid\mid T_1 \,{}_p\!\oplus\, T_2)$.

□

The converse to part (4) of Lemma E.1 also holds, though its proof is much more complicated.

Lemma E.2 If $o \in \mathbf{Outcomes}(\Delta \mid\mid (\tau_0.T_1 + \tau_0.T_2))$ then there are $q \in [0, 1]$, weight w and distributions Δ_1, Δ_2 such that $\Delta \stackrel{\tau}{\Longrightarrow}_w q \cdot \Delta_1 + (1-q) \cdot \Delta_2$ and $o = q \cdot o_1 + (1-q) \cdot o_2 + \langle w, \vec{0} \rangle$ for certain $o_i \in \mathbf{Outcomes}(\Delta_i \mid\mid T_i)$.

Proof. By mimicking the corresponding proof in [DvGHM09]. □

Proposition E.3 In a bounded wMDP, for every formula $\phi \in \mathcal{L}$ there exists a pair (T_ϕ, v_ϕ) with T_ϕ a multi-success test and $v_\phi \in [0, 1]^\Omega$ such that, for any weight r and distribution Δ,

(1) If $\langle r, \Delta \rangle \models \phi$ then $\exists o \in \mathbf{Outcomes}(\Delta \mid\mid T_\phi) : v_\phi \leq o + \langle r, \vec{0} \rangle$.

(2) If $\exists o \in \mathbf{Outcomes}(\Delta \mid\mid T_\phi) : v_\phi \leq o + \langle r, \vec{0} \rangle$ then there exists some weight r' such that $r' \geq r$ and $\langle r', \Delta \rangle \models \phi$.

T_ϕ is called a *characteristic test* of ϕ and v_ϕ its *target value*.

Proof. For any $\phi \in \mathcal{L}$ we define the pair T_ϕ and v_ϕ by structural induction.

- Let $\phi = \mathsf{tt}$. Take $T_\phi := \omega_0.\mathbf{0}$ for some $\omega \in \Omega$ and $v_\phi := \langle\, 0, \vec{\omega}\,\rangle$.
- Let $\phi = \langle\alpha\rangle_v \psi$. By induction, ψ has a characteristic test T_ψ with target value v_ψ. Take $T_\phi := a_0.T_\psi$ and $v_\phi := v_\psi + \langle\, v, \vec{0}\,\rangle$.
- Let $\phi = \phi_1 \wedge \phi_2$. Choose Ω-disjoint tests T_1, T_2 for ϕ_1 and ϕ_2, with target values v_1, v_2. Let $p \in (0,1)$ be chosen arbitrarily. We define $T_\phi := T_1 \,{}_p\!\oplus T_2$ and $v_\phi := p \cdot v_1 + (1-p) \cdot v_2$.
- Let $\phi = \phi_1 \,{}_p\!\oplus \phi_2$. Choose Ω-disjoint tests T_1, T_2 for ϕ_1 and ϕ_2 with target values v_1, v_2, and two fresh success actions ω_1, ω_2. Let $T_i' := T_i \,{}_{\frac{1}{2}}\!\oplus w_i$ and $v_i' := \frac{1}{2} v_i + \frac{1}{2}\langle\, 0, \vec{\omega_i}\,\rangle$. Note that for $i = 1,2$ we have that T_i' is also a characteristic test of ϕ_i with target value v_i. We define $T_\phi := \tau_0.T_1' + \tau_0.T_2'$ and $v_\phi := p \cdot v_1' + (1-p) \cdot v_2'$.

We now check by induction on ϕ that (1) and (2) above hold.

(1)

- Let $\phi = \mathsf{tt}$. For any configuration $\langle\, r, \Delta\,\rangle$, there exists some $o \in \textbf{Outcomes}(\Delta \mid\mid \omega_0.\mathbf{0})$ with $\langle\, 0, \vec{\omega}\,\rangle \le o \le o + \langle\, r, \vec{0}\,\rangle$, using Lemma E.1(1).
- Let $\phi = \langle\alpha\rangle_v \psi$. Suppose $\langle\, r, \Delta\,\rangle \models \phi$. Then there are w, Δ' with $\Delta \overset{\alpha}{\Longrightarrow}_w \Delta'$ and $\langle\, r + w - v, \Delta'\,\rangle \models \psi$. By induction, there exists $o_\psi \in \textbf{Outcomes}(\Delta' \mid\mid T_\psi)$ with $v_\psi \le o_\psi + \langle\, r + w - v, \vec{0}\,\rangle$. By Lemma E.1(2), there is some $o \in \textbf{Outcomes}(\Delta \mid\mid a_0.T_\psi)$ with $o = o_\psi + \langle\, w, \vec{0}\,\rangle$. It follows that $v_\phi = v_\psi + \langle\, v, \vec{0}\,\rangle \le o + \langle\, r, \vec{0}\,\rangle$ as required.
- Let $\phi = \phi_1 \wedge \phi_2$. Suppose $\langle\, r, \Delta\,\rangle \models \phi$. Then $\langle\, r, \Delta\,\rangle \models \phi_i$ for $i = 1,2$. By induction, there exists $o_i \in \textbf{Outcomes}(\Delta \mid\mid T_i)$ with $v_i \le o_i + \langle\, r, \vec{0}\,\rangle$. By Lemma E.1(3), we have $o := p \cdot v_1 + (1-p) \cdot v_2 \in \textbf{Outcomes}(\Delta \mid\mid T_\phi)$, and $v_\phi \le o + \langle\, r, \vec{0}\,\rangle$.
- Let $\phi = \phi_1 \,{}_p\!\oplus \phi_2$. Suppose $\langle\, r, \Delta\,\rangle \models \phi$. Then there are $r_1, r_2, \Delta_1, \Delta_2$ such that $\langle\, r, \Delta\,\rangle = p \cdot \langle\, r_1, \Delta_1\,\rangle + (1-p) \cdot \langle\, r_2, \Delta_2\,\rangle$ and $\langle\, r_i, \Delta_i\,\rangle \models \phi_i$ for $i = 1,2$. By induction,there exists some $o_i \in \textbf{Outcomes}(\Delta_i \mid\mid T_i)$ with $v_i \le o_i + \langle\, r_i, \vec{0}\,\rangle$. By Lemma E.1(1), we have $\langle\, 0, \vec{\omega_i}\,\rangle \in \textbf{Outcomes}(\Delta_i \mid\mid \omega_i)$. Since $T_i' = T_i \,{}_{\frac{1}{2}}\!\oplus \omega_i$, by Lemma E.1(3), there is some $o_i' := \frac{1}{2} \cdot o_i + \frac{1}{2} \cdot \langle\, 0, \vec{\omega_i}\,\rangle \in \textbf{Outcomes}(\Delta_i \mid\mid T_i')$. We note that

  $$v_i' \;:=\; \frac{1}{2} \cdot v_i + \frac{1}{2} \cdot \langle\, 0, \vec{\omega_i}\,\rangle \;\le\; \frac{1}{2} \cdot (o_i + \langle\, r_i, \vec{0}\,\rangle) + \frac{1}{2} \cdot \langle\, 0, \vec{\omega_i}\,\rangle \;=\; o_i' + \frac{1}{2} \cdot \langle\, r_i, \vec{0}\,\rangle.$$

  By Lemma E.1(4), there exists some $o := p \cdot o_1' + (1-p) \cdot o_2' \in \textbf{Outcomes}(\Delta \mid\mid (\tau_0.T_1' + \tau_0.T_2'))$. Therefore,

  $$v_\phi \;\le\; p \cdot (o_1' + \frac{1}{2} \cdot \langle\, r_1, \vec{0}\,\rangle) + (1-p) \cdot (o_2' + \frac{1}{2} \cdot \langle\, r_2, \vec{0}\,\rangle) \;=\; o + \frac{1}{2} \cdot \langle\, r, \vec{0}\,\rangle \;\le\; o + \langle\, r, \vec{0}\,\rangle.$$

(2)

- Let $\phi = \mathsf{tt}$. For any configuration $\langle\, r, \Delta\,\rangle$, we have $\langle\, r, \Delta\,\rangle \models \phi$.

- Let $\phi = \langle\alpha\rangle_v \psi$. Suppose there exists some $o \in \mathbf{Outcomes}(\Delta \mid\mid T_\phi)$ with $v_\phi \leq o + \langle\, r, \vec{0}\,\rangle$. It is easy to see that $o \neq \vec{0}$ because $o(\omega) \geq v_\phi(\omega) \neq 0$ for some $\omega \in \Omega$. By Lemma E.1(2) we have $\Delta \stackrel{a}{\Longrightarrow}_w \Delta'$ and $o = o' + \langle\, w, \vec{0}\,\rangle$ for some $o' \in \mathbf{Outcomes}(\Delta' \mid\mid T_\psi)$. It follows that $v_\psi + \langle\, v, \vec{0}\,\rangle \leq o' + \langle\, w, \vec{0}\,\rangle + \langle\, r, \vec{0}\,\rangle$. In other words, $v_\psi \leq o' + \langle\, r + w - v, \vec{0}\,\rangle$. By induction, there is some weight $r' \geq r + w - v$ with $\langle\, r', \Delta'\,\rangle \models \psi$. Let $r'' := \max(0, r' - w + v)$. Clearly, we have $r'' \geq r' - w + v \geq r$. It holds that $\langle\, r'', \Delta\,\rangle \models \phi$. To see this, we consider two cases: (i) if $r'' = r' - w + v$ then $r' - w + v \geq 0$ and by the definition of $\models$ we get $\langle\, r' - w + v, \Delta\,\rangle \models \phi$; (ii) if $r'' = 0$ then $r' - w + v \leq 0$, i.e. $w - v \geq r'$, which implies $\langle\, w - v, \Delta'\,\rangle \models \psi$ by Lemma 3.15 and then $\langle\, 0, \Delta\,\rangle \models \phi$.
- Let $\phi = \phi_1 \wedge \phi_2$. Suppose there exists $o \in \mathbf{Outcomes}(\Delta \mid\mid T_\phi)$ with $v_\phi \leq o + \langle\, r, \vec{0}\,\rangle$. By Lemma E.1(3) we have $o = p \cdot o_1 + (1-p) \cdot o_2$ for certain $o_i \in \mathbf{Outcomes}(\Delta \mid\mid T_i)$. Recall that T_1, T_w are Ω-disjoint tests. There exists weight r_i that $v_i \leq o_i + \langle\, r_i, \vec{0}\,\rangle$ for both $i = 1, 2$. To see this, we observe that (i) $v_i(\omega) \leq o_i(\omega)$ for all $\omega \in \Omega$ for if $v_i(\omega) > o_i(\omega)$ for some $i = 1$ or 2 then ω must occur in T_i but not in T_{3-i}, thus $v_{3-i}(\omega) = 0$ and $v_\phi(\omega) > o(\omega)$, in contradiction with the assumption; (ii) if x_i and y_i are the weight components of v_i and o_i respectively, then we can simply choose $r_i := \max(0, x_i - y_i)$ to ensure that $x_i \leq y_i + r_i$. By induction, there exists some weight $r'_i \geq r_i$ such that $\langle\, r'_i, \Delta\,\rangle \models \phi_i$, for $i = 1$ and 2. Let $r'' = \max(r'_1, r'_2, r)$. By Lemma 3.15 we have $\langle\, r'', \Delta\,\rangle \models \phi_i$, hence $\langle\, r'', \Delta\,\rangle \models \phi$.
- Let $\phi = \phi_1\, {}_p\!\oplus\, \phi_2$. Suppose there is some $o \in \mathbf{Outcomes}(\Delta \mid\mid T_\phi)$ such that $v_\phi \leq o + \langle\, r, \vec{0}\,\rangle$. By Lemma E.2, there are q, w, Δ_1, Δ_2 such that $\Delta \stackrel{\tau}{\Longrightarrow}_w q \cdot \Delta_1 + (1-q) \cdot \Delta_2$ and $o = q \cdot o'_1 + (1-q) \cdot o'_2 + \langle\, w, \vec{0}\,\rangle$ for certain $o'_i \in \mathbf{Outcomes}(\Delta_i \mid\mid T'_i)$. Now $v'_i(\omega_i) = o'_i(\omega_i) = \frac{1}{2}$ for both $i = 1$ and 2, so using that T_1, T_2 are Ω-disjoint tests, $\frac{1}{2}p = p \cdot v'_1(\omega_1) = v_\phi(\omega_1) \leq o(\omega_1) = q \cdot o'_1(\omega_1) = \frac{1}{2}q$ and likewise $\frac{1}{2}(1-p) = (1-p) \cdot v'_2(\omega_2) = v_\phi(\omega_2) \leq o(\omega_2) = (1-q) \cdot o'_2(\omega_2) = \frac{1}{2}(1-q)$. Together, these inequalities say that $p = q$. Exactly as in the previous case one obtains $v'_i \leq o'_i + \langle\, r_i, \vec{0}\,\rangle$ for some weight r_i, where $i = 1, 2$. Given that $T'_i = T_i\, {}_{\frac{1}{2}}\!\oplus\, w_i$, using Lemma E.1(3), it must be that $o'_i = \frac{1}{2}o_i + \frac{1}{2}\vec{\omega_i}$ for some $o_i \in \mathbf{Outcomes}(\Delta_i \mid\mid T_i)$ with $v_i \leq o_i + 2r_i$. By induction, there exists some $r'_i \geq 2r_i$ such that $\langle\, r'_i, \Delta\,\rangle \models \phi_i$, for $i = 1$ and 2. Let $r'' = \max(r,\ pr'_1 + (1-p)r'_2)$. We have $\langle\, r'', \Delta\,\rangle \models \phi$, using Lemma 3.15.

□

Corollary E.4 [Theorem 4.12] In a bounded wMDP, if $\Delta \sqsubseteq^r_{\mathrm{mmay}} \Theta$ then there exists some r' such that $r' \geq r$ and $\mathcal{L}(0, \Delta) \subseteq \mathcal{L}(r', \Theta)$.

Proof. For any $\phi \in \mathcal{L}(0, \Delta)$, we have $\langle\, 0, \Delta\,\rangle \models \phi$. Let T_ϕ be a characteristic test of ϕ with target value v_ϕ. By Proposition E.3(1), there exists some $o \in \mathbf{Outcomes}(\Delta \mid\mid T_\phi)$ such that $v_\phi \leq o$. Since $\Delta \sqsubseteq^r_{\mathrm{mmay}} \Theta$, there is some $o' \in \mathbf{Outcomes}(\Theta \mid\mid T_\phi)$ such that $o \leq o' + \langle\, r, \vec{0}\,\rangle$. It follows that $v_\phi \leq o'_i + \langle\, r, \vec{0}\,\rangle$. By Proposition E.3(2), there exists some weight r' such that $r' \geq r$ and $\langle\, r', \Theta\,\rangle \models \phi$, i.e. $\phi \in \mathcal{L}(r', \Theta)$. □

References

[BC00] Marco Bernardo and Rance Cleaveland. A theory of testing for Markovian processes. In *Proceedings of the 11th International Conference on Concurrency Theory*, volume 1877 of *Lecture Notes in Computer Science*, pages 305–319. Springer, 2000.

[BK08] C. Baier and J.-P. Katoen. *Principles of Model Checking*. The MIT Press, 2008.

[CDH09] Krishnendu Chatterjee, Laurent Doyen, and Thomas A. Henzinger. Probabilistic weighted automata. In *Proceedings of the 20th International Conference on Concurrency Theory*, volume 5710 of *Lecture Notes in Computer Science*, pages 244–258, 2009.

[DH11] Yuxin Deng and Matthew Hennessy. On the semantics of markov automata. In *Proceedings of the 38th International Colloquium on Automata, Languages and Programming*, volume 6756 of *Lecture Notes in Computer Science*, pages 307–318. Springer, 2011.

[DKV09] Manfred Droste, Werner Kuich, and Heiko Vogler, editors. *Handbook of Weighted Automata*. Springer, 2009.

[DvGHM09] Yuxin Deng, Rob van Glabbeek, Matthew Hennessy, and Carroll Morgan. Testing finitary probabilistic processes. In *Proceedings of the 20th International Conference on concurrency theory*, volume 5710 of *Lecture Notes in Computer Science*, pages 274–288. Springer, 2009.

[DvGMZ07] Yuxin Deng, Rob van Glabbeek, Carroll Morgan, and Chenyi Zhang. Scalar outcomes suffice for finitary probabilistic testing. In *Proceedings of the 16th European Symposium on Programming*, volume 4421 of *Lecture Notes in Computer Science*, pages 363–378. Springer, 2007.

[EHZ10] Christian Eisentraut, Holger Hermanns, and Lijun Zhang. On probabilistic automata in continuous time. In *Proceedings of the 25th Annual IEEE Symposium on Logic in Computer Science*, pages 342–351. IEEE Computer Society, 2010.

[Her02] H. Hermanns. *Interactive Markov Chains: The Quest for Quantified Quality*, volume 2428 of *Lecture Notes in Computer Science*. Springer, 2002.

[Hil96] Jane Hillston. *A Compositional Approach to Performance Modelling*. Cambridge University Press, 1996.

[JLY01] Bengt Jonsson, Kim G. Larsen, and Wang Yi. Probabilistic extensions of process algebras. In *Handbook of Process Algebra*, pages 685–710. Elsevier, 2001.

[KAK05] Astrid Kiehn and S. Arun-Kumar. Amortised bisimulations. In *Proceedings of the 25th IFIP WG 6.1 International Conference on Formal Techniques for Networked and Distributed Systems*, volume 3731 of *Lecture Notes in Computer Science*, pages 320–334. Springer, 2005.

[Koz83] Dexter Kozen. Results on the propositional mu-calculus. *Theoretical Computer Science*, 27:333–354, 1983.

[Lip65] S. Lipschutz. *Schaum's outline of theory and problems of general topology.* McGraw-Hill, 1965.

[LSV07] Nancy Lynch, Roberto Segala, and Frits Vaandrager. Observing branching structure through probabilistic contexts. *SIAM Journal on Computing*, 37:977–1033, 2007.

[Mat02] Jiri Matousek. *Lectures on Discrete Geometry*, volume 212 of *Graduate Texts in Mathematics*. Springer, 2002.

[MO98] Markus Müller-Olm. Derivation of characteristic formulae. *Electronic Notes in Theoretical Computer Science*, 18:159–170, 1998.

[NH84] R. De Nicola and M. C. B. Hennessy. Testing equivalences for processes. *Theoretical Computer Science*, 34(1–2):83–133, November 1984.

[Put94] Martin L. Puterman. *Markov Decision Processes.* Wiley, 1994.

[RKNP04] J. Rutten, M. Kwiatkowska, G. Norman, and D. Parker. *Mathematical Techniques for Analyzing Concurrent and Probabilistic Systems,* P. Panangaden and F. van Breugel (eds.), volume 23 of *CRM Monograph Series.* American Mathematical Society, 2004.

[Seg95] Roberto Segala. Modeling and verification of randomized distributed real-time systems. Technical Report MIT/LCS/TR-676, PhD thesis, MIT, Dept. of EECS, 1995.

[Seg96] Roberto Segala. Testing probabilistic automata. In *Proceedings of the 7th International Conference on Concurrency Theory*, volume 1119 of *Lecture Notes in Computer Science*, pages 299–314. Springer, 1996.

[Tar55] Alfred Tarski. A lattice-theoretical fixpoint theorem and its application. *Pacific Journal of Mathematics*, 5:285–309, 1955.