Game Equivalences and Metrics

Luca de Alfaro
Computer Engineering
University of California, Santa Cruz, CA
luca@soe.ucsc.edu

Rupak Majumdar
Department of Computer Science
University of California, Los Angeles, USA
rupak@cs.ucla.edu

Vishwanath Raman
Computer Science
University of California, Santa Cruz, USA
vishwa@soe.ucsc.edu

Mariëlle Stoelinga
FMT Group
University of Twente, the Netherlands
marielle@cs.utwente.nl

Abstract

1 Introduction

2 Games

2.1 Preliminary Definitions

For a finite set A, let $\text{Dist}(A)$ denote the set of probability distributions over A. We say that $p \in \text{Dist}(A)$ is *deterministic* if there is $a \in A$ such that $p(a) = 1$.

Given a set of states S, a *valuation over* S is a function $f : S \mapsto [0, 1]$ associating to every state $s \in S$ a value $0 \le f(s) \le 1$; we let $\mathcal{F}$ be the set of all valuations. For $c \in [0, 1]$, we denote by $\mathbf{c}$ the constant evaluation such that $\mathbf{c}(s) = c$ at all $s \in S$. We say that a valuation f is *boolean* if $f(s) \in \{0, 1\}$ for all $s \in S$. We order valuations pointwise: for $f, g \in \mathcal{F}$, we write $f \le g$ iff $f(s) \le g(s)$ at all $s \in S$; we remark that $\mathcal{F}$, under $\le$, forms a lattice.

Given $a, b \in \mathbb{R}$, we write $a \sqcup b = \max\{a, b\}$, and $a \sqcap b = \min\{a, b\}$; we also let $a \oplus b = \min\{1, \max\{0, a + b\}\}$. We extend $\sqcap, \sqcup, +, -, \oplus$ to valuations by interpreting them in pointwise fashion.

In this paper we will develop metrics for game structures. A *directed metric* is a function $d : S^2 \mapsto \mathbb{R}_{\ge 0}$ which satisfies $d(s, s) = 0$ and $d(s, t) \le d(s, u) + d(u, t)$ for all $s, t, u \in S$. We denote by $\mathsf{Me} = S^2 \mapsto \mathbb{R}$ the space of all metrics; this space, ordered pointwise, forms a lattice which we indicate with $(\mathsf{Me}, \le)$.

2.2 Game Structures

We assume a fixed, finite set $\mathcal{V}$ of *observation variables*. A (two-player, concurrent) *game structure* $G = \langle S, [\cdot], \mathcal{M}, \Gamma_1, \Gamma_2, \delta \rangle$ consists of the following components [AHK02, dAHK98]:

- A finite set S of states.
- A variable interpretation $[\cdot] : \mathcal{V} \times S \mapsto [0, 1]$, which associates with each variable $v \in \mathcal{V}$ a valuation $[x]$.
- A finite set $\mathcal{M}$ of moves.
- Two move assignments Γ_1, Γ_2: $S \mapsto 2^{\mathcal{M}} \setminus \emptyset$. For $i \in \{1, 2\}$, the assignment Γ_i associates with each state $s \in S$ the nonempty set $\Gamma_i(s) \subseteq \mathcal{M}$ of moves available to player-i at state s.
- A probabilistic transition function δ: $S \times \mathcal{M}^2 \mapsto \text{Dist}(S)$, that gives the probability $\delta(s, a_1, a_2)(t)$ of a transition from s to t when player-1 plays move a_1 and player-2 plays move a_2.

At every state $s \in S$, player 1 chooses a move $a_1 \in \Gamma_1(s)$, and simultaneously and independently player 2 chooses a move $a_2 \in \Gamma_2(s)$. The game then proceeds to the successor state $t \in S$ with probability $\delta(s, a_1, a_2)(t)$. We denote by $\text{Dest}(s, a_1, a_2) = \{t \in S \mid \delta(s, a_1, a_2)(t) > 0\}$ the set of *destination states* when actions a_1, a_2 are chosen at s. In the following, unless otherwise noted, the definitions refer to a game structure with components $G = \langle S, [\cdot], \mathcal{M}, \Gamma_1, \Gamma_2, \delta \rangle$. For player $i \in \{1, 2\}$, we write $\sim i = 3 - i$ for the opponent.

We consider various subclasses of game structures.

Turn-based game structures. A game structure G is *turn-based* if we can write S as the disjoint union of two sets: the set S_1 of *player-1 states,* and the set S_2 of *player-2 states,* such that $s \in S_1$ implies $|\Gamma_2(s)| = 1$, and $s \in S_2$ implies $|\Gamma_1(s)| = 1$. We say that G is *visibly turn-based* if G is turn-based, and if there is a special variable *turn* $\in \mathcal{Q}$, such that $[\mathit{turn}]s = 1$ iff $s \in S_1$, and $[\mathit{turn}]s = 0$ iff $s \in S_2$: thus, the variable *turn* indicates whose turn it is to play at a state.

Deterministic game structures. A game structure G is *deterministic* if, for all $s \in S$, $a_1 \in \mathcal{M}$, and $a_2 \in \mathcal{M}$, there exists a $t \in S$ such that $\delta(s, a_1, a_2)(t) = 1$; we denote such t by $\tau(s, a_1, a_2)$.

One-player game structures. We say that a game structure G is *one-player* if only one of the two players has a choice of moves. One-player game structures are also known as *Markov decision processes* [Der70, Ber95]. For $i \in \{1, 2\}$, we say that a structure is a player-i structure if $\forall s \in S$, $|\Gamma_{\sim i}(s)| = 1$. For one-player game structures, we omit the (single) move of the player without a choice of moves, and write $\delta(s, a)$ for the transition function.

Boolean game structures. ♣ do we need this? ♣ We say that a game structure G is *boolean* if, for all $v \in vars$ and $s \in S$, we have $[v]s \in \{0, 1\}$.

2.3 Valuation transformers

A *mixed move* is a probability distribution over the moves available to a player at a state. We denote by $\mathcal{D}_i(s) \subseteq \text{Dist}(\mathcal{M})$ the set of mixed moves available to player $i \in \{1, 2\}$ at $s \in S$, where:

$$\mathcal{D}_i(s) = \{\mathcal{D} \in \text{Dist}(\mathcal{M}) \mid \mathcal{D}(a) > 0 \textit{ implies } a \in \Gamma_i(s)\}.$$

The moves in $\mathcal{M}$ are sometimes called *pure moves,* in contrast to mixed moves. A *memoryless strategy* for player $i \in \{1, 2\}$ associates with each state $s \in S$ a mixed move in $\mathcal{D}_i(s)$.

We extend the transition function to mixed moves. For $s \in S$ and $x_1 \in \mathcal{D}_1(s)$, $x_2 \in \mathcal{D}_2(s)$, we write $\delta(s, x_1, x_2)$ for the next-state probability distribution induced by the mixed moves x_1 and x_2, defined for all $t \in S$ by

$$\delta(s, x_1, x_2)(t) = \sum_{a_1 \in \Gamma_1(s)} \sum_{a_2 \in \Gamma_2(s)} \delta(s, a_1, a_2)(t)\, x_1(a_1)\, x_2(a_2).$$

In the following, we sometimes restrict the moves of the players to pure moves. We identify a pure move $a \in \Gamma_i(s)$ available to player $i \in \{1, 2\}$ at a state s with a deterministic distribution that plays a with probability 1. Accordingly, we write $\mathcal{D}_i$ when we allow player i to play any mixed move, and write Γ_i when we restrict player i to only play pure moves.

Given a valuation $f \in \mathcal{F}$, a state $s \in S$, and two mixed moves $x_1 \in \mathcal{D}_1(s)$ and $x_2 \in \mathcal{D}_2(s)$, we define the the expectation of f from s under x_1, x_2:

$$\mathbb{E}_s^{x_1, x_2}(f) = \sum_{t \in S} \delta(s, x_1, x_2)(t)\, f(t)$$

For a game structure G, for $i \in \{1, 2\}$ we define the *valuation transformer* $\text{PPre}_i : \mathcal{F} \mapsto \mathcal{F}$ by, for all $f \in \mathcal{F}$ and $s \in S$,

$$\text{PPre}_i(f)(s) = \sup_{x_i \in \mathcal{D}_i(s)} \inf_{x_{\sim i} \in \mathcal{D}_{\sim i}(s)} \mathbb{E}_s^{x_1, x_2}(f).$$

Intuitively, $\text{PPre}_i(f)(s)$ is the maximal expectation player i can achieve of f after one step from s: this is the classical "one-day" or "next-stage" operator of the theory of repeated games [FV97]. We define also a *deterministic* version of this operator, in which the players are forced to play pure moves:

$$\text{Pre}_i(f)(s) = \max_{x_i \in \mathcal{D}_{d\,i}(s)} \min_{x_{\sim i} \in \mathcal{D}_{d\,\sim i}(s)} \mathbb{E}_s^{x_1, x_2}(f)\ .$$

3 Quantitative μ-calculus

We consider the set of properties expressed by the *quantitative μ-calculus* ($q\mu$). As discussed in [dAM04], a wide set of properties can be encoded in $q\mu$, spanning from basic properties such as maximal reachability and safety probability, to the maximal probability of satisfying a general ω-regular specification.

Syntax. The syntax of quantitative μ-calculus is defined with respect to the set of observation variables $\mathcal{V}$ as well as a set *MVars* of *calculus variables,* which are distinct from the observation variables in $\mathcal{V}$. The syntax is given as follows:

$$\begin{aligned} \phi \ ::= c \mid v \mid Z \mid \neg\phi \mid \phi \vee \phi \mid \phi \wedge \phi \mid \phi \oplus c \mid \\ \text{PPre}_1(\phi) \mid \text{PPre}_2(\phi) \mid \mu Z.\, \phi \mid \nu Z.\, \phi \end{aligned}$$

for constants $c \in [0, 1]$, observation variables $v \in \mathcal{V}$, and calculus variables $Z \in \mathit{MVars}$. In the formulas $\mu Z.\, \phi$ and $\nu Z.\, \phi$, we furthermore require that all occurrences of the bound variable Z in ϕ occur in the scope of an even number of occurrences of the complement operator $\neg$. A formula ϕ is *closed* if every calculus variable Z in ϕ occurs in the scope of a quantifier μZ or νZ. From now on, with abuse of notation, we denote by $q\mu$ the set of closed formulas of $q\mu$. A formula is a *player-i formula,* for $i \in \{1, 2\}$, if ϕ does not contain the $\text{PPre}_{\sim i}$ operator; we denote with $q\mu_i$ the syntactic subset of $q\mu$ consisting only of closed player-i formulas. A formula is in *positive form* if the negation appears only in front of game variables, i.e., in the context $\neg v$; we denote with $q\mu^+$ and $q\mu_i^+$ the subsets of $q\mu$ and $q\mu_i$ consisting only of positive formulas.

Semantics. Given $a, b \in \mathbb{R}$, we write $a \sqcup b = \max\{a, b\}$, and $a \sqcap b = \min\{a, b\}$; we also let $a \oplus b = \min\{1, \max\{0, a + b\}\}$. We extend $\sqcap, \sqcup, +, -, \oplus$ to valuations by interpreting them in pointwise fashion.

A variable evaluation ξ: *MVars* $\mapsto \mathcal{F}$ is a function that maps every variable $Z \in$ *MVars* to a valuation in $\mathcal{F}$. We write $\xi[Z \mapsto f]$ for the evaluation that agrees with ξ on all variables, except that Z is mapped to f. Given a game structure G and a variable evaluation ξ, every formula ϕ of the quantitative μ-calculus defines a valuation $[\![\phi]\!]^G_\xi \in \mathcal{F}$ (the superscript G is omitted if the game structure is clear from the context):

$$\begin{aligned}
&[\![c]\!]_\xi = \mathbf{c} \\
&[\![\phi_1 \left\{ {\vee \atop \wedge} \right\} \phi_2]\!]_\xi = [\![\phi_1]\!]_\xi \left\{ {\sqcup \atop \sqcap} \right\} [\![\phi_2]\!]_\xi \\
&[\![v]\!]_\xi = [v] \\
&[\![\phi \oplus c]\!]_\xi = [\![\phi]\!]_\xi \oplus \mathbf{c} \\
&[\![Z]\!]_\xi = \xi(Z) \\
&[\![\mathrm{PPre}_i(\phi)]\!]_\xi = \mathrm{PPre}_i([\![\phi]\!]_\xi) \\
&[\![\neg\phi]\!]_\xi = \mathbf{1} - [\![\phi]\!]_\xi \\
&[\![\left\{ {\mu \atop \nu} \right\} Z.\, \phi]\!]_\xi = \left\{ {\inf \atop \sup} \right\} \{f \in \mathcal{F} \mid f = [\![\phi]\!]_{\xi[Z \mapsto f]}\}
\end{aligned}$$

where $i \in \{1, 2\}$. The existence of the required fixpoints is guaranteed by the monotonicity and continuity of all operators. The fixpoints can be computed by Picard iteration: $[\![\mu Z.\, \phi]\!]_\xi = \lim_{k\to\infty} f_k$ where $f_0 = \mathbf{0}$, and $f_{k+1} = [\![\phi]\!]_{\xi[Z \mapsto f_k]}$ for all $k \geq 0$; and $[\![\nu Z.\, \phi]\!]_\xi = \lim_{k\to\infty} f_k$ where $f_0 = \mathbf{1}$, and f_{k+1} is defined as in the μ case. If ϕ is closed, $[\![\phi]\!]_\xi$ is independent of ξ, and we write simply $[\![\phi]\!]$.

We recall that, as a consequence of von Neumann's minimax theorem, we have that

$$[\![\neg \mathrm{PPre}_1(\phi)]\!]_\xi = [\![\mathrm{PPre}_2(\neg\phi)]\!]_\xi$$

for all ϕ and ξ. Furthermore, we note that if the game structure is boolean, turn-based, and deterministic, the valuation $[\![\phi]\!]_\xi$ is boolean.

We also define a *deterministic* semantics $[\![\cdot]\!]^d$ for $q\mu$, in which the players can select only pure moves in the operators PPre_1, PPre_2. $[\![\cdot]\!]^d$ is defined as $[\![\cdot]\!]$, except for the clause

$$[\![\mathrm{PPre}_i(\phi)]\!]^d_\xi = \mathrm{Pre}_i([\![\phi]\!]^d_\xi)\ .$$

Example 1 Given a set $T \subseteq S$, the *characteristic valuation* $\mathbf{T}$ of T is defined by $\mathbf{T}(s) = 1$ if $s \in T$, and $\mathbf{T}(s) = 0$ otherwise. With this notation, the maximal probability with which player $i \in \{1, 2\}$ can ensure reaching $T \subseteq S$ is given by $[\![\mu Z.(\mathbf{T} \vee \mathrm{PPre}_i(Z))]\!]$, and the maximal probability with which player i can guarantee staying in T is given by $[\![\nu Z.(\mathbf{T} \wedge \mathrm{PPre}_i(Z))]\!]$ [dAM04]. The first property is called a *reachability* property, the second a *safety* property.

4 Alternating and Concurrent Equivalences

For deterministic concurrent games, and for deterministic mixed moves, *alternating simulation* and *alternating bisimulation* relations [AHKV98] capture the notion of equivalence for deterministic μ-calculus formulas. Our first attempt is to generalize these relations by lifting equivalence of states to equivalence of distributions. Unless othewise specified, we develop our arguments for a set of observation variables $\mathcal{V}$ and a game structure denoted by the symbols $G = \langle S, [\cdot], \mathcal{M}, \Gamma_1, \Gamma_2, \delta \rangle$.

Game equivalences are based on the notion of local equivalence between states. Two states $s, t \in S$ are *locally equivalent,* written $s \equiv t$, if they assign the same values to all variables: for all $v \in \mathcal{V}$ we have $[v]s = [v]t$. For $s \in S$, we let $[s]_\equiv = \{t \in S | s \equiv t\}$ be the equivalence class of s for $\equiv$. We generalize $\equiv$ from states to distributions in the following way. Two distributions $p, q \in \mathrm{Dist}(S)$ are *congruent modulo* $\equiv$, written $p \equiv_R q$, if for all $s \in S$, we have $\sum_{t\in[s]_\equiv} p(t) = \sum_{t\in[s]_\equiv} q(t)$.

With the notion of local equivalence, we now define alternating simulation and bisimulation relations. The relations are parameterized by a player $i \in \{1, 2\}$, as well as whether mixed moves or only pure moves are allowed. For a relation $R \subseteq S \times S$, for $M \in \{\Gamma, \mathcal{D}\}$, for all $s, t \in S$ and $i \in \{1, 2\}$ consider the following conditions:

- **(loc)** $s\ R\ t$ implies $s \equiv t$;
- **(M-altsim)** $s\ R\ t$ implies
  $\forall x_i \in M_i(s)\,.\,\exists y_i \in M_i(t)\,.\,\forall y_{\sim i} \in M_{\sim i}(t)\,.\,\exists x_{\sim i} \in M_{\sim i}(s)\,.\,\delta(s, x_i, x_{\sim i}) \equiv_R \delta(t, y_i, y_{\sim i})$;
- **(M-sim)** $s\ R\ t$ implies
  $\forall x_i \in M_i(s)\,.\,\forall x_{\sim i} \in M_{\sim i}(s)\,.\,\exists y_i \in M_i(t)\,.\,\exists y_{\sim i} \in M_{\sim i}(t)\,.\,\delta(s, x_i, x_{\sim i}) \equiv_R \delta(t, y_i, y_{\sim i})$.

We can now define game simulation and bisimulation, as well as ordinary simulation and bisimulation, as follows:

- For $i \in \{1, 2\}$ and $M \in \{\Gamma, \mathcal{D}\}$, *player-i M-alternating simulation* $\preceq^M_i$ is the largest relation that satisfies (loc) and (M-altsim).
- For $i \in \{1, 2\}$ and $M \in \{\Gamma, \mathcal{D}\}$, *player-i M-alternating bimisulation* $\simeq^M_i$ is the largest symmetrical relation that satisfies (loc) and (M-altsim).
- For $M \in \{\Gamma, \mathcal{D}\}$, *$M$-Simulation* $\preceq^M_{12}$ (resp. *M-bisimulation* $\simeq^M_{12}$) is the largest relation (resp. largest symmetrical relation) that satisfies (loc) and (M-sim).

We denote with $\succeq^M_1$ the symmetrical relation to $\preceq^M_1$ (i.e., $s \succeq^M_1 t$ iff $t \preceq^M_1 s$, for all $s, t \in S$), and similarly for the other relations. In general, we omit the superscript $\mathcal{D}$ from the relations when players are allowed mixed moves, writing simply $\preceq_i$ (resp. $\simeq_i$) for $\preceq_i^{\mathcal{D}}$ (resp. $\simeq_i^{\mathcal{D}}$). We call the alternating simulation and bisimulation relations *deterministic* if players are restricted to use pure moves (i.e., $M = \Gamma$).

We shall evaluate alternating relations on games (as well as their metric generalizations in the next section) with respect to two criteria: logical characterizations and reciprocity. Logical characterization expresses the requirement that two equivalent states satisfy the same formulas of the logic $q\mu$. Reciprocity expresses the requirement that the relations are canonical.

Logical characterization As proved in [AHKV98], deterministic alternating bisimulation $\simeq_i^{\Gamma}$ for characterizes *alternating temporal logic* (ATL) [AHK02] on deterministic games. In fact, for all turn-based games and deterministic games with pure moves, deterministic alternating bisimulation characterizes $q\mu$ with the deterministic semantics. When extended to (not necessarily deterministic) game structures and mixed moves, alternating bisimulation similarly characterizes $q\mu$. We defer the proof of this result at the moment. It will follow from a more general result in the next section.

Theorem 1

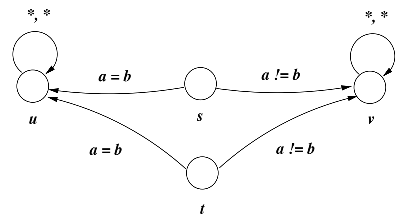
1. ***[AHKV98]*** *Let G be a deterministic game structure and s, t states of G. If $s \simeq_i^{\Gamma} t$ then for every $q\mu_i$ formula ϕ, we have $[\![\phi]\!]^d(s) = [\![\phi]\!]^d(t)$.*

2. *For any game structure G and states s, t of G, if $s \simeq_i t$ then for every $q\mu_i$ formula ϕ, we have $[\![\phi]\!](s) = [\![\phi]\!](t)$.*

Reciprocity. We say reciprocity holds for the alternating simulation and bisimulation relations if for all game structures, $\preceq_1 = \succeq_2$, and $\simeq_1 = \simeq_2$. The reciprocity property expresses an important canonicity of the equivalence relation: if it holds then there is *only one* game bisimulation on the game structure $\simeq_1 = \simeq_2$. Reciprocity in a game relation is a natural requirement for games for which a determinacy theorem holds, that is, for games in which the value obtained by player 1 is one minus the value obtained by player 2.

For concurrent games with pure moves, $\preceq_1$ and $\succeq_2$ (and also $\simeq_1$ and $\simeq_2$) may not coincide. This follows from the result that concurrent games, when played with pure moves, are not determined: while for all $s \in S$ and $f \in \mathcal{F}$ we have $\sup_{x_1 \in \mathcal{D}_1(s)} \inf_{x_2 \in \mathcal{D}_2} \mathbb{E}_s^{x_1,x_2}(f) = 1 - \sup_{x_2 \in \mathcal{D}_2} \inf_{x_1 \in \mathcal{D}_1(s)} \mathbb{E}_s^{x_1,x_2}(1-f)$, it is not necessarily the case that $\sup_{x_1 \in \mathcal{D}_{d1}(s)} \inf_{x_2 \in \mathcal{D}_{d2}} \mathbb{E}_s^{x_1,x_2}(f) = 1 - \sup_{x_2 \in \mathcal{D}_{d2}} \inf_{x_1 \in \mathcal{D}_{d1}(s)} \mathbb{E}_s^{x_1,x_2}(1-f)$.

For turn-based games, determinacy does hold [**?**, **?**] and reciprocity holds for the alternating simulation and bisimulation relations when restricted to turn-based games. In the case of concurrent games with mixed strategies, there is a similar determinacy theorem [**?**, dAM04], and again we shall show that reciprocity holds. We defer the proof of this theorem to a more general theorem in Section **??**.



(a) Alternating bisimilarity

Figure 1. Alternating simulation without game simulation

Theorem 2 *(1) Reciprocity does not hold for deterministic alternating simulation $\preceq_i^{\Gamma}$ and bisimulation $\simeq_i^{\Gamma}$. (2) Reciprocity holds for alternating simulation $\preceq_i^{\mathcal{D}}$ and bisimulation $\simeq_i^{\mathcal{D}}$.*

Relationship Unfortunately, as we demonstrate below, the deterministic and randomized alternating simulation and bisimulation relations are incomparable.

Theorem 3

1. *There exists a deterministic game structure G and states s, t of G such that $s \preceq_1^{\Gamma} t$ but $s \not\preceq_1^{\mathcal{D}} t$.*

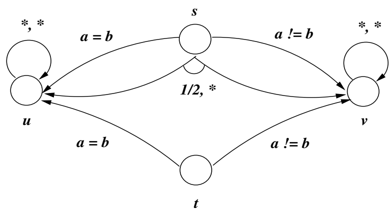
2. *There exists a game structure G and states s, t of G such that $s \preceq_1^{\mathcal{D}} t$ but $s \not\preceq_1^{\Gamma} t$.*

Proof For the first part, consider the deterministic game structure in Figure 1. It is a variant of the penny matching game. The set of moves is $\{0, 1, 2\}$, where move 2 is unavailable to either player from state s.

Assume that $u \not\equiv v$. For all pairs of moves a, b, where $a = b$, the successor state from s and t is u. For all pairs of moves a, b, where $a \neq b$, the successor state from s and t is v. u and v are absorbing.

Clearly, for player-1 alternating simulation, for every player-1 action a from s, there exists a player-1 action a' from t such that for every player-2 action b' from t, there exists a player-2 action b from s, that guarantees that the successor state from s and t is u. We have player-1 alternating simulation.

Now consider mixed strategies. From state s, if player-1 uses the strategy $x_1 = \{\frac{1}{2}, \frac{1}{2}, 0\}$ over moves $0, 1, 2$, then for any player-2 strategy we have, $\delta(s, x_1, \bullet)(u) = \frac{1}{2}$. But,



(a) Game Bisimilarity

Figure 2. Game simulation without alternating simulation

from state t, if player-2 chooses the strategy $y_2 = \{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\}$, then the probability of reaching u from state t will be $\delta(t, \bullet, y_2)(u) = \frac{1}{3}$, irrespective of the player-1 strategy from state t. We do not have player-1 game simulation.

For the second part, consider the game structure in Figure 2. The set of moves is $\{0, 1, \frac{1}{2}\}$, where the move $\frac{1}{2}$ is a player-1 move and not a player-2 move; it is like playing $\frac{1}{2}1 + \frac{1}{2}0$. The transitions for all moves are defined as follows: $\delta(s, a, b)(u) = 1, a = b$ and $\delta(s, a, b)(u) = 0, a \neq \frac{1}{2}, a \neq b$. If $a = \frac{1}{2}$, then $\delta(s, a, \bullet)(u) = (1 - x_1(0) - x_1(1))/2$, where x_1 is the player-1 strategy from state s. Let $\delta(t, a, b)(u) = 1, a = b$ and $\delta(t, a, b)(u) = 0, a \neq b$.

For any player-2 distribution over moves at states s and t, player-1 can restrict herself to the moves 0 and 1 at state s and mimic distributions at states s and t. Similarly, no matter what player-1 plays at the two states, player-2 can mimic transition probabilities from states s and t; for suppose player-1 plays move $\frac{1}{2}$ with a probability of α. The probability of transitioning to state u from s that player-2 can guarantee is $\frac{1}{2}(1 - \alpha) + \frac{1}{2}(\alpha)$, which is $\frac{1}{2}$. This is the same as the transition probability from state t to state u that player-2 can guarantee by playing the strategy $\{\frac{1}{2}, \frac{1}{2}\}$ over moves 0 and 1. Hence, we have game simulation.

For alternating simulation, since the move $\frac{1}{2}$ is not available to player-2 from state s, there exists a player-1 move, namely, $\frac{1}{2}$, such that there is no player-2 response from state s that guarantees a transition to state u. We do not have player-1 alternating simulation. ■

5 Metrics

5.1 A Metric Formulation of $\preceq$ and $\simeq$

For deterministic games and pure moves, the alternating simulation and bisimulation relations can be formulated as the kernel of certain metrics on the set of states, as follows. First, we define the *local distance* $[\equiv] \in \mathsf{Me}$ by:

$$[s \equiv t] = \max_{v \in \mathcal{V}} |[v](s) - [v](t)|$$

Next, we define the *metric transformer* $H_{\preceq_1} : \mathsf{Me} \mapsto \mathsf{Me}$ as follows, for $d \in \mathsf{Me}$, and $s, t \in S$:

$$\begin{aligned} & H_{\preceq_1}(d)(s,t) \\ & = [s \equiv t] \sqcup \sup_{x_1 \in \Gamma_1(s)} \inf_{y_1 \in \Gamma_1(t)} \sup_{y_2 \in \Gamma_1(t)} \inf_{x_2 \in \Gamma_1(s)} \\ & \qquad d(\tau(s, x_1, x_2), \tau(t, y_1, y_2)) \end{aligned} \tag{1}$$

It is easy to check that $H_{\preceq_1}$ is monotonic over the lattice $(\mathcal{F}, \leq)$; we can therefore define $[\preceq_1]$ as its least fixpoint in the lattice $(\mathsf{Me}, \leq)$. ♣ similarly altbis ♣ For similarity of notation with the case for relations, we will often write $[s \simeq_1 t]$ for $[\simeq_1](s, t)$, and similarly for other metrics.

Proposition 1 *For $i \in \{1, 2\}$, we have $[s \simeq_i t] = 0$ iff $s \simeq_i t$ and $[s \preceq_i t] = 0$ iff $s \preceq_i t$.*

5.2 Metrics for Markov Decision Processes

We consider the case for 1-MDPs; the case for 2-MDPs is symmetrical. We define two versions of the simulation metric.

A posteriori metrics. The first one is classical, from [vBW01, DGJP02]; we call it *a posteriori simulation metric.*

This metric is defined via a metric transformer $H^{1MDP}_{\sqsubseteq_1^r} : \mathsf{Me} \mapsto \mathsf{Me}$ as follows, for all $d \in \mathsf{Me}$ and $s, t \in S$:

$$\begin{aligned} & H^{1MDP}_{\sqsubseteq_1^r}(d)(s,t) \\ & = [s \equiv t] \sqcup \sup_{x_1 \in \mathcal{D}_{d1}(s)} \inf_{y_1 \in \mathcal{D}_{d1}(t)} D(\delta(s, x_1), \delta(t, y_1), d) \end{aligned}$$

where $D(\delta(s, x_1), \delta(t, y_1), d)$ is the distance between the two distributions $\delta(s, x_1)$ and $\delta(t, y_1)$, given the underlying metric d. This generalizes the deterministic case from Equation (6) to the probabilistic case and allows mixed moves.

For mixed moves, the distance $D(p, q)$ between two distributions p and q over S given an underlying metric d on states can be computed using a linear programming formulation [vBW01] as follows. Let the states in S be numbered from 0 to $N - 1$, with s_i representing the i^{th} state in S. We write p_i (respectively, q_i) to denote the probability $p(s_i)$ (respectively, $q(s_i)$). Let $N = |S| + 1$. For $i, j \in \{0, \ldots, N - 1\}$ let $c_{ij} = d(s_i, s_j)$, and let $c_{iN} = c_{Nj} = 1$ and $c_{NN} = 0$. Then, $D(\cdot)$ can be formulated as the following LP, known as the *transshipping formulation*:

$$\begin{aligned}
&\text{Minimize} && \sum_{i,j\in[0,N]} c_{ij}\lambda_{ij} \\
&\text{subject to} && \sum_{i\in[0,N]} \lambda_{ij} = p_j, \sum_{j\in[0,N]} \lambda_{ij} = q_j, \lambda_{ij} \geq 0
\end{aligned}$$

Intuitively, the p_j represent probability mass at the sources and q_j represent probability mass at the sinks, with c_{ij} representing the transhipment cost between source i and sink j; the task is to find a minimum cost allocation for moving probability mass from the sources to the sinks.

By the duality theorem for linear programming, the dual of the above transhipping problem yields the same distance. Let $C(d)$ be a feasible vector space, where $k \in C(d)$, implies that $k_i \geq 0$, $k_i \leq 1$ and $k_i - k_j \leq d(s_i, s_j) = c_{ij}, \forall i, j \in [0, N]$. Then the dual formulation is:

$$\begin{aligned}
&\text{Maximize} && \sum_{i\in[0,N]} p_i k_i - \sum_{i\in[0,N]} q_i k_i \\
&\text{subject to} && k \in C(d)
\end{aligned} \tag{2}$$

The vector k can be interpreted as a predicate, that assigns a value in the interval $[0, 1]$ to every state depending on the state transition probabilities induced by the player strategy.

Using the LP formulation for D above (Equation (2)), we can re-write $H^{1MDP}_{\sqsubseteq^r_1}(d)(s, t)$ as:

$$\begin{aligned}
&H^{1MDP}_{\sqsubseteq^r_1}(d)(s,t) \\
&= [s \equiv t] \sqcup \sup_{x_1\in\mathcal{D}_{d1}(s)} \inf_{y_1\in\mathcal{D}_{d1}(t)} \sup_{k\in C(d)} \big(\mathbb{E}^{x_1}_s(k) - \mathbb{E}^{y_1}_t(k)\big)
\end{aligned} \tag{3}$$

The *a posteriori simulation* metric is the least fixpoint of the transformer $H^{1MDP}_{\sqsubseteq^r_1}$.

A priori metrics. From the formulation in (2), we derive an alternate definition for the equivalence metric. We call it the *a priori metric*. We interpret the predicate k as being determined a priori: the player plays to maximize the expectation of the predicate. We define $H^{1MDP}_{\preceq^r_1} : \mathsf{Me} \mapsto \mathsf{Me}$ as follows, for all $d \in \mathsf{Me}$ and $s, t \in S$:

$$\begin{aligned}
&H^{1MDP}_{\preceq^r_1}(d)(s,t) \\
&= [s \equiv t] \sqcup \sup_{k\in C(d)} \sup_{x_1\in\mathcal{D}_{d1}(s)} \inf_{y_1\in\mathcal{D}_{d1}(t)} \big(\mathbb{E}^{x_1}_s(k) - \mathbb{E}^{y_1}_t(k)\big)
\end{aligned}$$

The *a priori simulation* metric is the least fixpoint of the transformer $H^{1MDP}_{\preceq_1}$.

We will now show that for one-player game structures, a posteriori metrics coincide with a priori metrics. Let p^s_u and p^t_u be vectors such that, $\forall a_i \in \mathcal{M}$, we have $p^s_u(i) = \delta(s, a_i)(u)$ and $p^t_u(i) = \delta(t, a_i)(u)$, for $i \in \{1, \ldots, |\mathcal{M}|\}$. We define

$$f_0 = \mathbb{E}^{x_1}_s(k) - \mathbb{E}^{y_1}_t(k) = \sum_{u\in S}\big(x_1^T p^s_u - y_1^T p^t_u\big)k_u$$

Theorem 4 *For all one-player game structures, for all $s, t \in S$, and for any metric $d \in \mathsf{Me}$, we have $H^{1MDP}_{\sqsubseteq^r_1}(d)(s,t) = H^{1MDP}_{\preceq^r_1}(d)(s,t)$. Thus, for all one-player game structures, the a posteriori simulation metric is identical to the a priori simulation metric.*

Proof We need to show that $\sup_k \sup_{x_1} \inf_{y_1} f_0 = \sup_{x_1} \inf_{y_1} \sup_k f_0$. First, we interchange the two sups in the left hand side:

$$\begin{aligned}
&\sup_k \sup_{x_1} \inf_{y_1} f_0 \\
&= \sup_k \big(\sup_{x_1} \inf_{y_1} \big\{\sum_{u\in S}(x_1^T p^s_u - y_1^T p^t_u)k_u\big\}\big) \\
&= \sup_{x_1} \big(\sup_{k} \inf_{y_1} \big\{\sum_{u\in S}(x_1^T p^s_u - y_1^T p^s_u)k_u\big\}\big)
\end{aligned} \tag{4}$$

Consider the term $\sup_k \inf_{y_1}(\sum_{u\in S}(x_1^T p^s_u - y_1^T p^t_u)k_u)$. Each term is bilinear in the two variables y_1 and k_u, since we have a fixed x_1. Further, the domain of y_1 is the $|\mathcal{M}|$ dimensional probability simplex and the domain of k is a compact convex subset of a linear space. By the generalized minimax theorem [**?**],

$$\begin{aligned}
&\sup_k \inf_{y_1} \big\{\sum_{u\in S}(x_1^T p^s_u - y_1^T p^t_u)k_u\big\} \\
&= \inf_{y_1} \sup_k \big\{\sum_{u\in S}(x_1^T p^s_u - y_1^T p^t_u)k_u\big\}
\end{aligned} \tag{5}$$

Therefore, from (4) and (5), we get

$$\begin{aligned}
&\sup_k \big(\sup_{x_1} \inf_{y_1} \big\{\sum_{u\in S}(x_1^T p^s_u - y_1^T p^t_u)k_u\big\}\big) \\
&= \sup_{x_1} \inf_{y_1} \big(\sup_{k} \big\{\sum_{u\in S}(x_1^T p^s_u - y_1^T p^t_u)k_u\big\}\big)
\end{aligned}$$

$$\sup_k \sup_{x_1} \inf_{y_1} f_0 = \sup_{x_1} \inf_{y_1} \sup_k f_0$$

■

The metrics defined above logically characterize $q\mu$ and the distance between two states is exactly equal to the difference in property evaluations at the two states [**?**]. ♣ I am assuming that we don't have to give out the logical characterization here but can simply cite Prakash et al ♣.

5.3 Metrics for Concurrent Games

We now extend the simulation and bisimulation metrics from one-player game structures to general game structures.

As we shall see, unlike for one-player games, the a priori and the a posteriori metrics do not coincide over games. In particular, we show that the a priori formulation satisfies both a tight logical characterization as well as reciprocity, while, perhaps surprisingly, the more natural a posteriori version does not.

A posteriori metrics for concurrent games. The first one is classical, from [**?**, **?**]; we call it *a posteriori simulation metric*. The metric is defined as the least fixpoint of a metric transformer. We define a metric transformer $H^{G}_{\sqsubseteq_1^r}$: $\mathsf{Me} \mapsto \mathsf{Me}$ as follows, for all $d \in \mathsf{Me}$ and $s, t \in S$:

$$\begin{aligned} & H_{\sqsubseteq_1^r}(d)(s,t) \\ & = [s \equiv t] \sqcup \sup_{x_1 \in \mathcal{D}_1(s)} \inf_{y_1 \in \mathcal{D}_1(t)} \sup_{y_2 \in \mathcal{D}_1(t)} \inf_{x_2 \in \mathcal{D}_1(s)} \\ & \qquad D(\delta(s,x_1,x_2), \delta(t,y_1,y_2), d) \end{aligned} \tag{6}$$

As before, $D(\cdot,\cdot)$ is the distance between two distributions given the underlying metric d. By expanding out the definition of $D(\cdot,\cdot)$, we get that (6) is equivalent to

$$\begin{aligned} & [s \equiv t] \sqcup \sup_{x_1 \in \mathcal{D}_1(s)} \inf_{y_1 \in \mathcal{D}_1(t)} \sup_{y_2 \in \mathcal{D}_1(t)} \inf_{x_2 \in \mathcal{D}_1(s)} \\ & \qquad \sup_k \big(\mathbb{E}_s^{x_1,x_2}(k) - \mathbb{E}_t^{y_1,y_2}(k)\big) \end{aligned} \tag{7}$$

where k ranges over $C(d)$, i.e., $0 \le k_i \le 1$ and $k_i - k_j \le d(i,j)$ for all $i, j \in \{0, \ldots, N\}$.

A priori metrics for concurrent games. As with one-player game structures, we derive an alternate metric transformer $H_{\preceq_1^r}$: $\mathsf{Me} \mapsto \mathsf{Me}$ as follows, for all $d \in \mathsf{Me}$ and $s, t \in S$:

$$\begin{aligned} & H_{\preceq_1^r}(d)(s,t) \\ & = [s \equiv t] \sqcup \sup_k \sup_{x_1 \in \mathcal{D}_1(s)} \inf_{y_1 \in \mathcal{D}_1(t)} \sup_{y_2 \in \mathcal{D}_1(t)} \inf_{x_2 \in \mathcal{D}_1(s)} \\ & \qquad \big(\mathbb{E}_s^{x_1,x_2}(k) - \mathbb{E}_t^{y_1,y_2}(k)\big) \end{aligned}$$

The least fixpoint of $H_{\preceq_1^r}$ is called the *a priori simulation metric*. Similar to the case of MDPs, k is a known predicate in this case. In player-1 alternating simulation, the player picks strategies at the two states s and t that maximizes the expectation of the predicate picked a priori, considering the worst player-2 response at these states.

Comparing *a priori* and *a posteriori* metrics We now show that a priori simulation metrics are always bounded from above by posteriori metrics. Further, equality may not hold.

Theorem 5

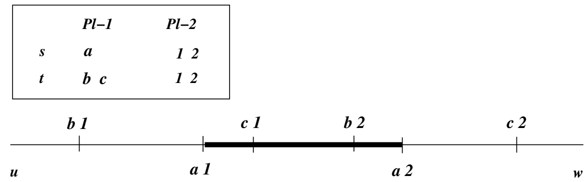


Figure 3. A Game that shows a priori but not posteriori refinement

1. *For every game structure G with states S, for every metric $d \in \mathsf{Me}$, and for every $s, t \in S$, we have $H_{\preceq_1^r}(d)(s,t) \le H_{\sqsubseteq_1^r}(d)(s,t)$. Hence, the a priori simulation metric is bounded above by the a posteriori simulation metric.*

2. *There is a game structure G for which the a priori simulation metric is strictly smaller than the a posteriori simulation metric.*

Proof We need to show that,

$$\begin{aligned} & \sup_k \big(\sup_{x_1} \inf_{y_1} \sup_{y_2} \inf_{x_2} \big\{\mathbb{E}_s^{x_1,x_2}(k) - \mathbb{E}_t^{y_1,y_2}(k)\big\}\big) \\ & \le \sup_{x_1} \inf_{y_1} \sup_{y_2} \inf_{x_2} \big(\sup_k \big\{\mathbb{E}_t^{y_1,y_2}(k) - \mathbb{E}_s^{x_1,x_2}(k)\big\}\big) \\ \text{i.e.,} \quad & \sup_{x_1} \sup_k \inf_{y_1} \sup_{y_2} \inf_{x_2} \big\{\mathbb{E}_s^{x_1,x_2}(k) - \mathbb{E}_t^{y_1,y_2}(k)\big\} \\ & \le \sup_{x_1} \inf_{y_1} \sup_k \sup_{y_2} \inf_{x_2} \big\{\mathbb{E}_t^{y_1,y_2}(k) - \mathbb{E}_s^{x_1,x_2}(k)\big\} \end{aligned} \tag{8}$$

$$\begin{aligned} \text{i.e.,} \quad & \sup_k \inf_{y_1} \sup_{y_2} \inf_{x_2} \big\{\mathbb{E}_s^{x_1,x_2}(k) - \mathbb{E}_t^{y_1,y_2}(k)\big\} \\ & \le \inf_{y_1} \sup_k \sup_{y_2} \inf_{x_2} \big\{\mathbb{E}_t^{y_1,y_2}(k) - \mathbb{E}_s^{x_1,x_2}(k)\big\} \end{aligned} \tag{9}$$

(12) is certainly true, since $\sup_k \inf_{y_1} g_0 \le \inf_{y_1} \sup_k g_0$, where $g_0 = \sup_{y_2} \inf_{x_2} \mathbb{E}_s^{x_1,x_2}(k) - \mathbb{E}_t^{y_1,y_2}(k)$.

(11) follows from the result for MDPs. When player-1 strategies from state s and t are known, the game reduces to an MDP. If the player-1 strategies are a_1 and b_1 at s and t, then $\sup_k \sup_{y_2} \inf_{x_2} (\mathbb{E}_t^{b_1,y_2}(k) - \mathbb{E}_s^{a_1,x_2}(k)) = \sup_{y_2} \inf_{x_2} \sup_k (\mathbb{E}_t^{b_1,y_2}(k) - \mathbb{E}_s^{a_1,x_2}(k))$.

For the second part, we now show an example where a priori distances are strictly less than a posteriori distances. Consider a game, where player-2 has two moves labelled $\{1, 2\}$ from states s and t, while player-1 has a single move $\{a\}$ at state s and moves $\{b, c\}$ at state t.

In Figure 3, the point $b1$ represents the transition probability, $P_t^{b,1}$, to states u and w, when player-1 plays move b and player-2 plays move 1 from state t. Similarly for the other points on the probability simplex. The line segment between the points $b1$ and $b2$ represents the transition probabilities when the player-2 strategy is a convex combination of the pure strategies that pick moves 1 and 2. Assume that the following ordered relation holds amongst the transition probabilities to state u for all combinations of player-1 and player-2 pure strategies,

$$P_t^{b,1}(u) > P_s^{a,1}(u) > P_t^{c,1}(u) > P_t^{b,2}(u) > P_s^{a,2}(u) > P_t^{c,2}(u) \quad (10)$$

Further,
$$P_s^{a,1}(u) < \frac{P_t^{b,1}(u) + P_t^{c,1}(u)}{2} \quad (11)$$

$$P_s^{a,2}(u) > \frac{P_t^{b,2}(u) + P_t^{c,2}(u)}{2} \quad (12)$$

The relation for state w is reversed. In figure 2, the solid line represents the space of all distributions over states, induced by the single player-1 move $\{a\}$ and the convex combination of player-2 moves $\{1, 2\}$ at state s.

For a player-1 distribution x_1 from s, we define the set $\tau_s(x_1) = \{P_s^{x_1,x_2} | x_2 \in \text{Dist}_2(s)\}$. A strategy y_1 at state t mimics a strategy x_1 at state s if, $\tau_t(y_1) \subseteq \tau_s(x_1)$.

- $s \not\sqsubseteq_1 t$**: There is no mixed player-1 strategy over** $\{b, c\}$ **from state** t **that mimics** a **from state** s**.** From the figure it is the case that for any $\alpha \in [0, 1]$, player-2 has a spoiling strategy from state t, that will induce a distribution outside the solid line. Intuitively, for the transition probability to be within the solid line, player-1, in the absence of a known k, will have to choose an α such that $\alpha > \frac{1}{2}$ and $(1-\alpha) > \frac{1}{2}$, which clearly cannot be the case. $\forall \alpha \in [0, 1], \tau_t(\alpha c + (1-\alpha)b) \not\subseteq \tau_s(a)$ and $[s \sqsubseteq_1 t] > 0$.

- $s \preceq_1 t$**: For every** k**, we can find an** $\alpha \in [0, 1]$**, such that** $\alpha c + (1-\alpha)b$ **mimics** a **for that** k**.** Let $\tau_s(x_1) \downarrow k = \{d^T k | d \in \tau_s(x_1)\}$. For a known k, we need to prove that,

$$\tau_t(y_1) \downarrow k \subseteq \tau_s(x_1) \downarrow k \quad (13)$$

If the value of k is the same at states u and v, $k_u = k_w$. A choice of $\alpha = \frac{1}{2}$ will be optimal for both players at t, with $P_t^{c,1}(u) > P_t^{y_1,y_2}(u) > P_t^{b,2}$. Player-2 can mimic this distribution from state s.

Let the value of $k_u = 1$ and $k_w = 0$. Since player-1 plays to maximize the expectation of this predicate k, $\tau_t(y_1) = \tau_t(b) = \{z | P_t^{b,1}(u) \geq z(u) \geq P_t^{b,2}(u)\}$, which is the line segment between the points $b1$ and $b2$. Similarly, if $k_u = 0$ and $k_w = 1$, the induced transition probabilities $\tau_t(c)$ will be on the line segment between the points $c1$ and $c2$. Therefore, for all values of k in between these extremal values,

$$\tau_t(y_1) \downarrow k \subseteq \tau_t(b) \downarrow k \cap \tau_t(c) \downarrow k \subseteq \tau_s(x_1) \downarrow k$$

■

Logical Characterization

Theorem 6 *For states s and $t \in S$, for all $\phi \in q\mu$, $|\phi(s) - \phi(t)| \leq [s \preceq_1 t]$.*

Proof Define the function *Depth* by induction on formulas as $Depth(p) = 0$, $Depth(\neg p) = 0$; $Depth(\phi_1 \{^\vee_\wedge\} \phi_2) = max(Depth(\phi_1), Depth(\phi_2))$; $Depth(\phi_1 \oplus c) = Depth(\phi_1)$; and $Depth(\text{PPre}_i(\phi)) = Depth(\phi_i) + 1$.

Let ϕ be a formula that has the same valuation as a particular feasible k. We prove the result by induction on the depth of the formulas.

Assume $|\phi(s) - \phi(t)| \leq [s \preceq_1 t]$. For the inductive case, consider the operator PPre,

$$\begin{aligned} &\text{PPre}_1(\phi)(s) - \text{PPre}_1(\phi)(t) \\ &= \sup_{x_1} \inf_{x_2} \mathbb{E}_s^{x_1,x_2}(\phi) - \sup_{y_1} \inf_{y_2} \mathbb{E}_t^{y_1,y_2}(\phi) \\ &= \sup_{x_1} \inf_{y_1} \sup_{y_2} \inf_{x_2} \big\{\mathbb{E}_s^{x_1,x_2}(\phi) - \mathbb{E}_t^{y_1,y_2}(\phi)\big\} \\ &\leq [s \preceq_t t] \end{aligned} \quad (14)$$

(11) holds since $0 \leq \phi(s) \leq 1$, $0 \leq \phi(t) \leq 1$ and $|\phi(s) - \phi(t)| \leq [s \preceq_1 t]$, which implies that $\big\{\mathbb{E}_s^{x_1,x_2}(\phi) - \mathbb{E}_t^{y_1,y_2}(\phi)\big\} \leq [s \preceq_t t]$. Hence, for all $\phi \in q\mu$, $|\phi(s) - \phi(t)| \leq [s \preceq_1 t]$. ■

Theorem 7 *For states s and $t \in S$, if $s \not\preceq_1 t$, there exists a $\phi \in q\mu$ that witnesses it: if $[s \preceq_1 t] > 0$, then there exists $\phi \in q\mu$, with $|\phi(s) - \phi(t)| = [s \preceq_1 t]$.*

Proof ■

Theorem 8 *A priori metrics logically characterize $q\mu$. For all states s and $t \in S$, $\sup_\phi |\phi(s) - \phi(t)| = [s \preceq_1 t]$.*

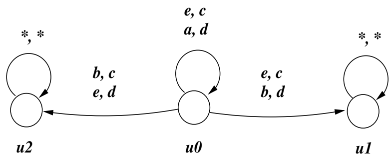
Proof From Theorem 6, we have $|\phi(s) - \phi(t)| \leq [s \preceq_1 t]$ or $\sup_\phi |\phi(s) - \phi(t)| \leq [s \preceq_1 t]$. We will show that for all $\epsilon > 0$, $\sup_\phi |\phi(s) - \phi(t)| \geq [s \preceq_1 t] - \epsilon$. ■

♣ Not sure if we want the following result. ♣

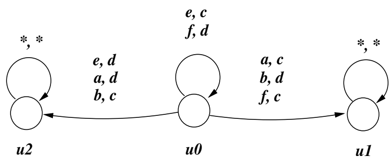
We conclude our results on logical characterization, by showing that boolean relations are weaker than their quantitative counterparts. We consider the case of alternating bisimulation and show that there is a game structure with two states that cannot be distinguished using a boolean logic but can using $q\mu$.

Theorem 9 $[\simeq_1^b] \leq [\simeq_1^q]$

Consider the games in Figure 2. u_0 is the initial state. In Game-1, player-1 actions are $\{a, b, e\}$ and player-2 actions are $\{c, d\}$. In Game-2, player-1 actions are $\{a, b, e, f\}$ and player-2 actions are $\{c, d\}$. Let $U = \{u_0, u_1, u_2\}$. Let $d(2, 1) = \frac{2}{3}$, $d(2, 0) = 1$ and $d(1, 0) = \frac{1}{3}$.



(a) Game 1



(b) Game 2

Figure 4. To show that $\simeq^b \neq \simeq^q$

- **The case of $\simeq_1^b$**: Under a boolean interpretation, every formula evaluates to either 0 or 1. A formula ϕ in this setting is interpreted as an indicator function, that has the value 1 in the subset of states in which it holds.

  In Game 1, the probability with which player-1 can guarantee a transition to u_2 is $\frac{1}{2}$: by playing actions b and e with probability $\frac{1}{2}$. The probability with which player-1 can guarantee a transition to u_1 or staying at u_0 is $\frac{1}{2}$. Similarly, the probability with which player-1 can satisfy the complementary goal of not transitioning to u_2 is 1, achieved by playing the action a with probability 1. This holds for the complementary goals $\neg \bigcirc \{u_0\}$ and $\neg \bigcirc \{u_1\}$, where $\bigcirc$ has the usual LTL interpretation of *next step*.

  In Game 2, one can easily verify that the maximum value for player-1 of the goals $\bigcirc u$ and $\neg \bigcirc u$, with $u \in U$, is the same as in Game 1. Therefore, for all values of ϕ, under a boolean interpretation, the simulation distance between u_0 in Game 1 and u_0 in Game 2 is 0.

- **The case of $\simeq_1^q$**: Consider the quantitative interpretation. Let a feasible evaluation of ϕ be $\phi(u_0) = 0$, $\phi(u_1) = \frac{1}{3}$ and $\phi(u_2) = 1$.

  In Game 1, the player-1 optimal strategy is $\{0, \frac{1}{2}, \frac{1}{2}\}$ over the moves a, b, e. This yields a player-1 expectation of ϕ at state s of $\frac{1}{2} \cdot \phi(u_2) + \frac{1}{2} \cdot \phi(u_0) + 0 \cdot \phi(u_1)$.

  In Game 2, the player-1 optimal strategy is $\{0, 0, 0, 1\}$ over the moves a, b, e, f. This yields a player-1 expectation of ϕ at state t to be $0 \cdot \phi(u_2) + 0 \cdot \phi(u_0) + 1 \cdot \phi(u_1)$, which gives a simulation distance of $\frac{1}{6}$. ■

Reciprocity The following theorem is an important result that applies to a priori simulation metrics. We show that the directed simulation distance between two states s and t coincides for player-1 and player-2.

Theorem 10 *(1) Reciprocity holds for a priori simulation metrics. $H_{\preceq_1^r}(d)(s,t) = H_{\preceq_2^r}(d)(s,t)$. (2) Reciprocity fails for a posteriori simulation metrics. Player-1 and player-2 a posteriori game simulation distances do not coincide.*

Proof We need to show that,

$$\begin{aligned}&\sup_k\Big(\sup_{x_1}\inf_{y_1}\sup_{y_2}\inf_{x_2}\big\{\mathbb{E}_s^{x_1,x_2}(k) - \mathbb{E}_t^{y_1,y_2}(k)\big\}\Big)\\ = &\sup_k\Big(\sup_{y_2}\inf_{x_2}\sup_{x_1}\inf_{y_1}\big\{\mathbb{E}_t^{y_1,y_2}(k) - \mathbb{E}_s^{x_1,x_2}(k)\big\}\Big)\end{aligned}$$

$$\begin{aligned}&\sup_k\Big(\sup_{x_1}\inf_{y_1}\sup_{y_2}\inf_{x_2}\big\{\mathbb{E}_s^{x_1,x_2}(k) - \mathbb{E}_t^{y_1,y_2}(k)\big\}\Big)\\ = &\sup_k\Big(\sup_{x_1}\inf_{x_2}\mathbb{E}_s^{x_1,x_2}(k) - \sup_{y_1}\inf_{y_2}\mathbb{E}_t^{y_1,y_2}(k)\Big) \quad (15)\\ = &\sup_k\Big(\inf_{x_2}\sup_{x_1}\mathbb{E}_s^{x_1,x_2}(k) - \inf_{y_2}\sup_{y_1}\mathbb{E}_t^{y_1,y_2}(k)\Big) \quad (16)\\ = &\sup_k\Big(\inf_{x_2}\sup_{x_1}\mathbb{E}_s^{x_1,x_2}(1-k) - \inf_{y_2}\sup_{y_1}\mathbb{E}_t^{y_1,y_2}(1-k)\Big) \quad (17)\\ = &\sup_k\Big(\inf_{x_2}\sup_{x_1}\mathbb{E}_s^{x_1,x_2}(-k) - \inf_{y_2}\sup_{y_1}\mathbb{E}_t^{y_1,y_2}(-k)\Big)\\ = &\sup_k\Big(\sup_{y_2}\inf_{y_1}\mathbb{E}_t^{y_1,y_2}(k) + \inf_{x_2}\sup_{x_1}\mathbb{E}_s^{x_1,x_2}(-k)\Big)\\ = &\sup_k\Big(\sup_{y_2}\inf_{x_2}\sup_{x_1}\inf_{y_1}\big\{\mathbb{E}_t^{y_1,y_2}(k) - \mathbb{E}_s^{x_1,x_2}(k)\big\}\Big) \quad (18)\end{aligned}$$

Since the variables x_1, x_2, y_1, y_2 are probability distributions, by the minimax theorem, (5) follows from (4). We get (6) by a variable replacement of k by $1-k$.

To show reciprocity fails for a posteriori simulation metrics, consider the game from Figure 1. Let player-2 moves from state t be $\{1,2,3,4\}$, where moves 1 and 2 are the same as before while the move pair $(*, \{3,4\})$ is identical to $(a, \{1,2\})$ from state s. That is, $P_t^{*,3}(u) = P_s^{a,1}(u)$ and $P_t^{*,4}(u) = P_s^{a,2}(u)$. Since we have strengthened player-2's moves at state s, it is still the case that $[s \sqsubseteq_1 t] > 0$. We need to prove that, $[s \sqsubseteq_2 t] = 0$. Player-2 can restrict herself to moves 3 and 4 at state t and mimic state transition probabilities at both states s and t irrespective of player-1 strategies at state t. Therefore, $[s \sqsubseteq_2 t] = 0$. ■

In a turn-based game, only one player has a choice of moves at every state. Therefore, comparing any two player-i states is the same as comparing two states in an MDP. Therefore, a priori and a posteriori metrics carry over in definition and properties from the case for MDPs.

The following theorem states that for turn-based game structures, reciprocity holds for both a priori and a posteriori metrics.

Theorem 11 *For all turn-based games reciprocity holds.*

Proof The result follows from the equivalence of a priori player-1 and player-2 metrics and the fact that for MDPs, a priori metrics and a posteriori metrics coincide. ■

6 Special Cases

Fall-back We say fall-back holds in a game structure if $\preceq_1 = \preceq_2 = \preceq_{12}$, and $\simeq_1 = \simeq_2 = \simeq_{12}$.

If the fall-back property holds, then the alternating relations coincide with the classical, non-alternating relations of [Mil90]; we note that fall-back holds only for visibly turn-based game structures.

References

[AHK02] R. Alur, T.A. Henzinger, and O. Kupferman. Alternating time temporal logic. *J. ACM*, 49:672–713, 2002.

[AHKV98] R. Alur, T.A. Henzinger, O. Kupferman, and M.Y. Vardi. Alternating refinement relations. In *CONCUR 98: Concurrency Theory. 9th Int. Conf.*, volume 1466 of *Lect. Notes in Comp. Sci.*, pages 163–178. Springer-Verlag, 1998.

[Ber95] D.P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, 1995. Volumes I and II.

[dAHK98] L. de Alfaro, T.A. Henzinger, and O. Kupferman. Concurrent reachability games. In *Proc. 39th IEEE Symp. Found. of Comp. Sci.*, pages 564–575. IEEE Computer Society Press, 1998.

[dAM04] L. de Alfaro and R. Majumdar. Quantitative solution of omega-regular games. *Journal of Computer and System Sciences*, 68:374–397, 2004. A preliminary version appeared in STOC 01: 33rd Annual ACM Symposium on Theory of Computing, 2001.

[Der70] C. Derman. *Finite State Markovian Decision Processes*. Academic Press, 1970.

[DGJP02] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Approximating labelled markov processes. *Information and Computation*, 2002.

[FV97] J. Filar and K. Vrieze. *Competitive Markov Decision Processes*. Springer-Verlag, 1997.

[GH82] Y. Gurevich and L. Harrington. Trees, automata, and games. In *Proc. 14th ACM Symp. Theory of Comp.*, pages 60–65. ACM Press, 1982.

[Mar98] D.A. Martin. The determinacy of Blackwell games. *The Journal of Symbolic Logic*, 63(4):1565–1581, 1998.

[Mil90] R. Milner. Operational and algebraic semantics of concurrent processes. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume B, pages 1202–1242. Elsevier Science Publishers (North-Holland), Amsterdam, 1990.

[vBW01] F. van Breugel and J. Worrel. An algorithm for quantitative verification of probabilistic transition systems. In *CONCUR 01: Concurrency Theory. 12th Int. Conf.*, volume 2154 of *Lect. Notes in Comp. Sci.*, pages 336–350, 2001.