Randomness for Free[☆,☆☆]

Krishnendu Chatterjee[a], Laurent Doyen[*,b], Hugo Gimbert[c], Thomas A. Henzinger[a]

[a]*IST Austria (Institute of Science and Technology Austria)*
[b]*LSV, ENS Cachan & CNRS, France*
[c]*LaBRI & CNRS, Bordeaux, France*

Abstract

We consider two-player zero-sum games on finite-state graphs. These games can be classified on the basis of the information of the players and on the mode of interaction between them. On the basis of information the classification is as follows: (a) partial-observation (both players have partial view of the game); (b) one-sided complete-observation (one player has complete observation); and (c) complete-observation (both players have complete view of the game). On the basis of mode of interaction we have the following classification: (a) concurrent (players interact simultaneously); and (b) turn-based (players interact in turn). The two sources of randomness in these games are randomness in the transition function and randomness in the strategies. In general, randomized strategies are more powerful than deterministic strategies, and probabilistic transitions give more general classes of games. We present a complete characterization for the classes of games where randomness is not helpful in: (a) the transition function (probabilistic transitions can be simulated by deterministic transitions); and (b) strategies (pure strategies are as powerful as randomized strategies). As a consequence of our characterization we obtain new undecidability results for these games.

1. Introduction

Games on graphs. Games played on graphs provide the mathematical framework to analyze several important problems in computer science as well as mathematics. In

[☆]A preliminary version of this paper appeared in the *Proceedings of the 35th International Symposium on Mathematical Foundations of Computer Science* (MFCS), Lecture Notes in Computer Science 6281, Springer, 2010, pp. 246-257.

[☆☆]This research was partly supported by Austrian Science Fund (FWF) Grant No P 23499- N23, FWF NFN Grant No S11407-N23 and S11402-N23 (RiSE), ERC Start grant (279307: Graph Games), Microsoft faculty fellows award, the ERC Advanced Grant QUAREM (267989: Quantitative Reactive Modeling), European project Cassting (FP7-601148), European project COMBEST, and the European Network of Excellence ArtistDesign.

[*]Corresponding author: Laurent Doyen, LSV, CNRS UMR 8643 & ENS Cachan, 61 avenue du Président Wilson, 94235 Cachan Cedex, France.

Email addresses: `Krishnendu.Chatterjee@ist.ac.at` (Krishnendu Chatterjee), `doyen@lsv.ens-cachan.fr` (Laurent Doyen), `hugo.gimbert@labri.fr` (Hugo Gimbert), `tah@ist.ac.at` (Thomas A. Henzinger)

particular, when the vertices and edges of a graph represent the states and transitions of a reactive system, then the synthesis problem (Church's problem) asks for the construction of a winning strategy in a game played on a graph [5, 24, 23, 21]. Game-theoretic formulations have also proved useful for the verification [1], refinement [18], and compatibility checking [14] of reactive systems. Games played on graphs are dynamic games that proceed for an infinite number of rounds. In each round, the players choose moves; the moves, together with the current state, determine the successor state. An outcome of the game, called a *play*, consists of the infinite sequence of states that are visited.

Strategies and objectives. A strategy for a player is a recipe that describes how the player chooses a move to extend a play. Strategies can be classified as follows: (a) *pure* strategies, which always deterministically choose a move to extend the play, and (b) *randomized* strategies, which may choose at a state a probability distribution over the available moves. Objectives are generally Borel-measurable sets [19]: the objective for a player is a Borel set B in the Cantor topology on S^ω (where S is the set of states), and the player satisfies the objective if the outcome of the game is a member of B. In verification, objectives are usually ω*-regular languages*. The ω-regular languages generalize the classical regular languages to infinite strings; they occur in the low levels of the Borel hierarchy (they lie in $\Sigma_3 \cap \Pi_3$) and they form a robust and expressive language for determining payoffs for commonly used specifications.

Classification of games. Games played on graphs can be classified according to the knowledge of the players about the state of the game, and the way of choosing moves. Accordingly, there are (a) *partial-observation* games, where each player only has a partial or incomplete view about the state and the moves of the other player; (b) *one-sided complete-observation* games, where one player has partial knowledge and the other player has complete knowledge about the state and moves of the other player; and (c) *complete-observation* games, where each player has complete knowledge of the game. According to the way of choosing moves, the games on graphs can be classified into *turn-based* and *concurrent* games. In turn-based games, in any given round only one player can choose among multiple moves; effectively, the set of states can be partitioned into the states where it is player 1's turn to play, and the states where it is player 2's turn. In concurrent games, both players may have multiple moves available at each state, and the players choose their moves simultaneously and independently.

Sources of randomness. There are two sources of randomness in these games. First is the randomness in the transition function: given a current state and moves of the players, the transition function defines a probability distribution over the successor states. The second source of randomness is the randomness in strategies (when the players play randomized strategies). In this work we study when randomness can be obtained for *free*; i.e., we study in which classes of games the probabilistic transitions can be simulated by deterministic transitions and the classes of games where pure strategies are as powerful as randomized strategies.

Motivation. The motivation to study this problem is as follows: (a) if for a class of games it can be shown that randomness is for free in the transition function, then all future works related to analysis of computational complexity, strategy complexity, and algorithmic solutions can focus on the simpler class with deterministic transitions (the randomness in transition function may be essential for modeling appropriate stochastic reactive systems, but the analysis can focus on the deterministic subclass); (b) if for a class of games it can be shown that randomness is for free in strategies, then all future works related to correctness results can focus on the simpler class of pure strategies, and the results would follow for the more general class of randomized strategies; and (c) the characterization of randomness for free will allow hardness results obtained for the more general class of games (such as games with randomness in the transition function) to be carried over to simpler class of games (such as games with deterministic transitions).

Contribution. The contributions of this paper are as follows:

1. *Randomness for free in the transition function.* We show that randomness in the transition function can be obtained for free for complete-observation concurrent games (and any class that subsumes complete-observation concurrent games) and for one-sided complete-observation turn-based games (and any class that subsumes this class). The reduction is polynomial for complete-observation concurrent games, and exponential for one-sided complete-observation turn-based games. It is known that for complete-observation turn-based games, a probabilistic transition function cannot be simulated by a deterministic transition function (see discussion in Section 3.4 for details), and thus we present a complete characterization when randomness can be obtained for free in the transition function.

2. *Randomness for free in the strategies.* We show that randomness in strategies is free for complete-observation turn-based games, and for 1-player partial-observation games (POMDPs). For all other classes of games randomized strategies are more powerful than pure strategies. It follows from a result of Martin [20] that for 1-player complete-observation games with probabilistic transitions (MDPs) pure strategies are as powerful as randomized strategies. We present a generalization of this result to the case of POMDPs. Our proof is totally different from Martin's proof and based on a new derandomization technique of randomized strategies.

3. *Concurrency for free in games.* We show that concurrency is obtained for free with partial-observation, both for one-sided complete-observation games as well as for general partial-observation games (see Section 3.5). It follows that for partial-observation games, future research can focus on the simpler model of turn-based games, and concurrency does not add anything in the presence of partial observation.

4. *New undecidability results.* As a consequence of our characterization of randomness for free, we obtain new undecidability results. In particular, using our results and results of Baier et al. [2] we show for one-sided complete-observation deterministic games, the problems of almost-sure winning for coBüchi objectives and positive winning for Büchi objectives are undecidable. Thus we obtain the first undecidability result for qualitative analysis (almost-sure and positive winning) of one-sided complete-observation deterministic games with ω-regular objectives.

Applications of our results. While we already show that our results allow us to obtain new undecidability results, they have also been used to simplify proofs and analysis of POMDPs and partial-observation games [6, 7, 8, 9, 16] (e.g. [7, Lemma 21] and [9, Claim 2. Lemma 5.1]) as well as extended to other settings such as probabilistic automata [17].

2. Definitions

In this section we present the definition of concurrent games of partial information and their subclasses, and notions of strategies and objectives. Our model of game is equivalent to the model of stochastic games with signals [22, 3] (in stochastic games with signals, the players receive signals which represent information about the game, which in our model is represented as observations). A *probability distribution* on a finite set A is a function $\kappa : A \to [0,1]$ such that $\sum_{a \in A} \kappa(a) = 1$. We denote by $\mathcal{D}(A)$ the set of probability distributions on A.

Concurrent games of partial observation. A *concurrent game of partial observation* (or simply a *game*) is a tuple $G = \langle S, A_1, A_2, \delta, \mathcal{O}_1, \mathcal{O}_2 \rangle$ with the following components:

1. *(State space).* S is a finite set of states;
2. *(Actions).* A_i ($i = 1,2$) is a finite set of actions for player i;
3. *(Probabilistic transition function).* $\delta : S \times A_1 \times A_2 \to \mathcal{D}(S)$ is a concurrent probabilistic transition function that given a current state s, actions a_1 and a_2 for both players gives the transition probability $\delta(s, a_1, a_2)(s')$ to the next state s'; for the sake of effectiveness, we assume that all probabilities in the transition function are rational;
4. *(Observations).* $\mathcal{O}_i \subseteq 2^S$ ($i = 1,2$) is a finite set of observations for player i that partition the state space S. These partitions uniquely define functions $\mathsf{obs}_i : S \to \mathcal{O}_i$ ($i = 1,2$) that map each state to its observation (for player i) such that $s \in \mathsf{obs}_i(s)$ for all $s \in S$.

We sometimes relax the assumption that games have a finite state space, and we allow the set S of states to be *countable*. This is useful in the context of game solving, where we get a countable state space after fixing an arbitrary strategy for one of the players in a game. In our results we explicitly mention when we consider countable state space and when we consider finite state space.

Special cases. We consider the following special cases of partial-observation concurrent games, obtained either by restrictions in the observations, the mode of selection of moves, the type of transition function, or the number of players:

- *(Observation restriction).* The games with *one-sided complete-observation* are the special case of games where $\mathcal{O}_1 = \{\{s\} \mid s \in S\}$ (i.e., player 1 has complete observation) or $\mathcal{O}_2 = \{\{s\} \mid s \in S\}$ (player 2 has complete observation). The *games of complete-observation* are the special case of games where $\mathcal{O}_1 = \mathcal{O}_2 =$

$\{\{s\} \mid s \in S\}$, i.e., every state is visible to each player and hence both players have complete observation. If a player has complete observation we omit the corresponding observation sets from the description of the game.

- *(Mode of interaction restriction).* A *turn-based state* is a state s such that either (i) $\delta(s,a,b) = \delta(s,a,b')$ for all $a \in A_1$ and all $b, b' \in A_2$ (i.e, the action of player 1 determines the transition function and hence it can be interpreted as player 1's turn to play), we refer to s as a player-1 state, and we use the notation $\delta(s,a,-)$; or (ii) $\delta(s,a,b) = \delta(s,a',b)$ for all $a, a' \in A_1$ and all $b \in A_2$, we refer to s as a player-2 state, and we use the notation $\delta(s,-,b)$. A state s which is both a player-1 state and a player-2 state is called a *probabilistic state* (i.e., the transition function is independent of the actions of the players). We write $\delta(s,-,-)$ to denote the transition function in s. The *turn-based games* are the special case of games where all states are turn-based.

- *(Transition function restriction).* The *deterministic games* are the special case of games where for all states $s \in S$ and actions $a \in A_1$ and $b \in A_2$, there exists a state $s' \in S$ such that $\delta(s,a,b)(s') = 1$. We refer to such states s as deterministic states. For deterministic games, it is often convenient to assume that $\delta : S \times A_1 \times A_2 \to S$.

- *(Player restriction).* The *1½-player games*, also called *partially observable Markov decision processes* (or POMDPs), are the special case of games where the action set A_1 or A_2 is a singleton. Note that 1½-player games are turn-based. Games without player restriction are sometimes called 2½-player games.

The 1½-player games of complete-observation are Markov decision processes (or MDPs), and MDPs with all states deterministic can be viewed as graphs (and are often called 1-player games).

Classes of game graphs. We use the following abbreviations (Table 1a): we write Pa for partial-observation, Os for one-sided complete-observation, Co for complete-observation, C for concurrent, and T for turn-based. For example, CoC will denote complete-observation concurrent games, and OsT will denote one-sided complete-observation turn-based games. For $\mathcal{C} \in \{\mathsf{Pa}, \mathsf{Os}, \mathsf{Co}\} \times \{\mathsf{C}, \mathsf{T}\}$, we denote by $\mathcal{G}_{\mathcal{C}}$ the set of all $\mathcal{C}$ games. Note the following strict inclusions (see also Figure 2): partial observation (Pa) is more general than one-sided complete-observation (Os) and Os is more general than complete-observation (Co), and concurrent (C) is more general than turn-based (T). We will denote by $\mathcal{G}_D$ the set of all games with deterministic transition function. The results we establish in this article are summarized in Figure 3.

Plays. In concurrent games of partial observation, in each turn, player 1 chooses an action $a \in A_1$, player 2 chooses an action $b \in A_2$, and the successor of the current state s is chosen according to the probabilistic transition function $\delta(s,a,b)$. A *play* in a game G is an infinite sequence $\rho = s_0\, a_0 b_0\, s_1\, a_1 b_1\, s_2 \dots$ such that $\delta(s_i, a_i, b_i, s_{i+1}) > 0$ for all $i \geq 0$. The *prefix up to* s_n of the play ρ is denoted by $\rho(n)$. The set of plays in G is denoted $\mathsf{Plays}(G)$, and the set of corresponding finite prefixes (or histories) is denoted $\mathsf{Prefs}(G)$. The *observation sequence* of ρ for

Pa	partial observation		Σ_G	all player-1 strategies
Os	one-sided complete observation		Σ_G^O	observation-based pl.-1 strategies
Co	complete observation		Σ_G^P	pure player-1 strategies
C	concurrent		Π_G	all player-2 strategies
T	turn-based		Π_G^O	observation-based pl.-2 strategies
D	deterministic transition function		Π_G^P	pure player-2 strategies

(a) Classes of games (b) Classes of strategies in game G

Table 1: Abbreviations.

player i ($i = 1, 2$) is the unique infinite sequence $\mathsf{obs}_i(\rho) = o_0\, c_0\, o_1\, c_1\, o_2 \ldots$ such that $s_j \in o_j \in O_i$, and $c_j = a_j$ if $i = 1$, and $c_j = b_j$ if $i = 2$ for all $j \geq 0$.

Strategies. A *pure strategy* in a game G for player 1 is a function $\sigma : \mathsf{Prefs}(G) \to A_1$. A *randomized strategy* in G for player 1 is a function $\sigma : \mathsf{Prefs}(G) \to \mathcal{D}(A_1)$. A (pure or randomized) strategy σ for player 1 is *observation-based* if for all prefixes $\rho, \rho' \in \mathsf{Prefs}(G)$, if $\mathsf{obs}_1(\rho) = \mathsf{obs}_1(\rho')$, then $\sigma(\rho) = \sigma(\rho')$. We omit analogous definitions of strategies for player 2. We denote by Σ_G, Σ_G^O, Σ_G^P, Π_G, Π_G^O and Π_G^P the set of all player-1 strategies in G, the set of all observation-based player-1 strategies, the set of all pure player-1 strategies, the set of all player-2 strategies in G, the set of all observation-based player-2 strategies, and the set of all pure player-2 strategies, respectively (Table 1b). Note that if player 1 has complete observation, then $\Sigma_G^O = \Sigma_G$.

Objectives. An *objective* for player 1 in G is a set $\varphi \subseteq S^\omega$ of infinite sequences of states. A play $\rho = s_0\, a_0 b_0\, s_1\, a_1 b_1\, s_2 \ldots \in \mathsf{Plays}(G)$ *satisfies* the objective φ, denoted $\rho \models \varphi$, if $s_0 s_1 s_2 \ldots \in \varphi$. A Borel objective is a Borel-measurable set in the Cantor topology on S^ω [19]. We specifically consider ω-regular objectives specified as parity objectives (a canonical form to express all ω-regular objectives [26]). For a sequence $\bar{s} = s_0 s_1 s_2 \ldots$ we denote by $\mathrm{Inf}(\bar{s})$ the set of states that occur infinitely often in $\bar{s}$, that is, $\mathrm{Inf}(\bar{s}) = \{s \in S \mid s_j = s \text{ for infinitely many } j\text{'s}\}$. For $d \in \mathbb{N}$, let $p : S \to \{0, 1, \ldots, d\}$ be a *priority function*, which maps each state to a nonnegative integer priority. The *parity* objective $\mathsf{Parity}(p)$ requires that the minimum priority that occurs infinitely often be even. Formally, $\mathsf{Parity}(p) = \{\bar{s} \in S^\omega \mid \min\{p(s) \mid s \in \mathrm{Inf}(\bar{s})\} \text{ is even}\}$. The Büchi and coBüchi objectives are the special cases of parity objectives with two priorities, for $p : S \to \{0, 1\}$ and $p : S \to \{1, 2\}$ respectively. We say that an objective φ is *visible* for player i if for all $\rho, \rho' \in \mathsf{Plays}(G)$, if $\rho \models \varphi$ and $\mathsf{obs}_i(\rho) = \mathsf{obs}_i(\rho')$, then $\rho' \models \varphi$. For example if the priority function maps observations to priorities (i.e., $p : \mathcal{O}_i \to \{0, 1, \ldots, d\}$), then the parity objective is visible for player i.

Almost-sure winning, positive winning, and value function. An *event* is a measurable subset of S^ω, and given strategies σ and π for the two players, the probabilities of events are uniquely defined [27]. For a Borel objective φ, we denote by $\mathrm{Pr}_s^{\sigma,\pi}(\varphi)$ the probability that φ is satisfied by the play obtained from the starting state s when the strategies σ and π are used. Given a game structure G and a state s, an observation-based strategy σ for player 1 is *almost-sure winning* (resp., *positive winning*) for the

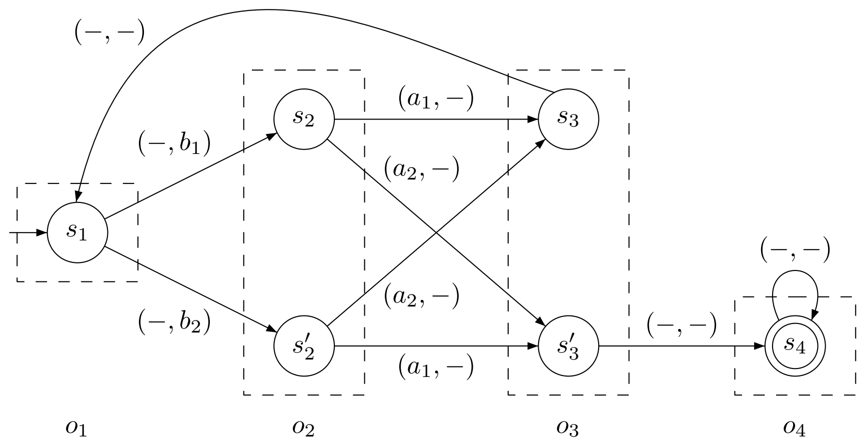

Figure 1: A game with one-sided complete observation (Example 1).

objective φ from s if for all observation-based randomized strategies π for player 2, we have $\Pr_s^{\sigma,\pi}(\varphi) = 1$ (resp., $\Pr_s^{\sigma,\pi}(\varphi) > 0$).

The *value function* $\langle\langle 1\rangle\rangle_{val}^{G}(\varphi) : S \to \mathbb{R}$ for player 1 and objective φ assigns to every state of G the maximal probability with which player 1 can guarantee the satisfaction of φ with an observation-based strategy, against all observation-based strategies for player 2. Formally we define

$$\langle\langle 1\rangle\rangle_{val}^{G}(\varphi)(s) = \sup_{\sigma \in \Sigma_G^O} \inf_{\pi \in \Pi_G^O} \Pr_s^{\sigma,\pi}(\varphi).$$

The value of an observation-based strategy σ for player 1 and objective φ in state s is $val_1^{\sigma}(\varphi)(s) = \inf_{\pi \in \Pi_G^O} \Pr_s^{\sigma,\pi}(\varphi)$. Analogously for player 2, define $\langle\langle 2\rangle\rangle_{val}^{G}(\varphi)(s) = \inf_{\pi \in \Pi_G^O} \sup_{\sigma \in \Sigma_G^O} \Pr_s^{\sigma,\pi}(\varphi)$ and $val_2^{\pi}(\varphi)(s) = \sup_{\sigma \in \Pi_G^O} \Pr_s^{\sigma,\pi}(\varphi)$. For $\varepsilon \geq 0$, an observation-based strategy σ is *ε-optimal* for φ from s if $val_1^{\sigma}(\varphi)(s) \geq \langle\langle 1\rangle\rangle_{val}^{G}(\varphi)(s) - \varepsilon$. An *optimal* strategy is a 0-optimal strategy.

Example 1 ([10]). *Consider the game with one-sided complete observation (player* 2 *has complete information) shown in Figure 1. Consider the Büchi objective defined by the state* s_4 *(i.e., state* s_4 *has priority* 0 *and other states have priority* 1*). Because player* 1 *has partial observation (given by the partition* $\mathcal{O}_1 = \{\{s_1\}, \{s_2, s'_2\}, \{s_3, s'_3\}, \{s_4\}\}$*), she cannot distinguish between* s_2 *and* s'_2 *and therefore has to play the same actions with same probabilities in* s_2 *and* s'_2 *(while it would be easy to win by playing* a_2 *in* s_2 *and* a_1 *in* s'_2*, this is not possible). In fact, player* 1 *cannot win using a pure observation-based strategy. However, playing* a_1 *and* a_2 *uniformly at random in all states is almost-sure winning. Every time the game visits observation* o_2*, for any strategy of player* 2*, the game visits* s_3 *and* s'_3 *with probability* $\frac{1}{2}$*, and hence*

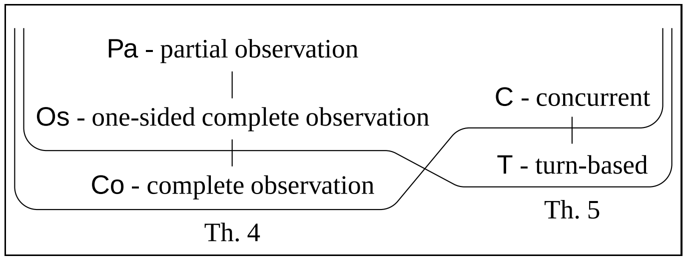

Figure 2: Hierarchy of the various classes of game graphs. According to Theorem 4 randomness is for free in the transition function for concurrent games even with complete observation, and according to Theorem 5 randomness is for free in the transition function for one-sided complete observation games even if they are turn-based. For $2^1\!/\!_2$-player games, randomness in the transition function is not for free only in complete-observation turn-based games.

also reaches s_4 with probability $\frac{1}{2}$. It follows that against all player-2 strategies the play eventually reaches s_4 with probability 1, and then stays there.

Theorem 1 ([20]). *Let G be a* CoT *stochastic game (with countable state space S) with initial state s and an objective $\varphi \subseteq S^\omega$. Then the following equalities hold: $\langle\!\langle 1 \rangle\!\rangle^G_{val}(\varphi)(s) = \langle\!\langle 2 \rangle\!\rangle^G_{val}(\varphi)(s) = \sup_{\sigma \in \Sigma^O_G \cap \Sigma^P_G} \inf_{\pi \in \Pi^O_G} \Pr^{\sigma,\pi}_s(\varphi)$.*

Discussion of Theorem 1. Theorem 1 can be derived as a consequence of Martin's proof of determinacy of Blackwell games [20]: the result states that for CoT stochastic games pure strategies can achieve the same value as randomized strategies, and as a special case, the result also holds for MDPs (for a detailed discussion how to obtain the result from [20] see [13, Lemma 10]). Note that Martin's determinacy result of $\langle\!\langle 1 \rangle\!\rangle^G_{val}(\varphi)(s) = \langle\!\langle 2 \rangle\!\rangle^G_{val}(\varphi)(s)$ also holds for CoC stochastic games (complete-observation concurrent stochastic games), but the equality with $\sup_{\sigma \in \Sigma^O_G \cap \Sigma^P_G} \inf_{\pi \in \Pi^O_G} \Pr^{\sigma,\pi}_s(\varphi)$ (which implies existence of pure ϵ-optimal strategies for $\epsilon > 0$) only holds for CoT stochastic games.

3. Randomness for Free in Transition Function

In this section we present a precise characterization of the classes of games where randomness in the transition function can be obtained for *free*: in other words, we present the precise characterization of classes of games with probabilistic transition function that can be reduced to the corresponding class with deterministic transition function. We present our results as three reductions: (a) the first reduction allows us to separate probability from the mode of interaction; (b) the second reduction shows how to simulate probability in transition function with CoC (complete-observation concurrent) deterministic transition function; and (c) the final reduction shows how to simulate probability in transition with OsT (one-sided complete-observation turn-based) deterministic transition function. We then show that for CoT (complete-observation

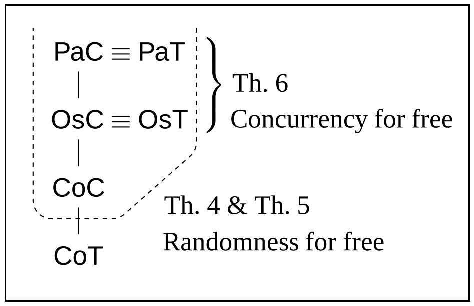

Figure 3: Summary of the results of Section 3.

turn-based) games, randomness in the transition function cannot be obtained for free, and conclude with the *concurrency for free* result that OsT and PaT games can simulate OsC and PaC games respectively.

A *reduction* from a class $\mathcal{G}$ of games to a class $\mathcal{G}'$ is a mapping that, from a game $G \in \mathcal{G}$ and an objective φ in G, returns a game $G' \in \mathcal{G}'$ and an objective φ' in G', and such that the state space S of G is (injectively) mapped to the state space S' of G'. In all our reductions we have $S \subseteq S'$, and thus the state-space mapping is the identity (on S). The mapping of objectives in our reductions is such that φ is the projection of φ' on S^ω. It follows that when φ is a parity objective defined with at most d priorities, then so is φ' (and in the sequel, we omit the definition of the priority function for φ'), and when φ is an objective in the k-th level of the Borel hierarchy, then so is φ'.

All our reductions are *local*: they consist of a gadget construction and replacement locally at every state. Additional properties of interest for reductions are as follows:

- A reduction is *almost-sure-preserving* (resp., *positive-preserving*), if for all states $s \in S$ in G: player 1 is almost-sure winning (resp., positive winning) in G from s if and only if player 1 is almost-sure winning (resp., positive winning) in G' from s.

- A reduction is *value-preserving* if $\langle\langle 1 \rangle\rangle_{val}^{G}(\varphi)(s) = \langle\langle 1 \rangle\rangle_{val}^{G'}(\varphi')(s)$ for all $s \in S$, and *threshold-preserving* if for all $\eta \in \mathbb{R}$, all states $s \in S$, and all $\bowtie \in \{>, \geq\}$: there exists an observation-based strategy $\sigma \in \Sigma_G^O$ for player 1 in G such that $\forall \pi \in \Pi_G^O : \mathrm{Pr}_s^{\sigma,\pi}(\varphi) \bowtie \eta$ if and only if there exists an observation-based strategy $\sigma' \in \Sigma_{G'}^O$ for player 1 in G' such that $\forall \pi' \in \Pi_{G'}^O : \mathrm{Pr}_s^{\sigma',\pi'}(\varphi') \bowtie \eta$.

All reductions presented in this paper are threshold-preserving. Note that threshold-preserving implies value-preserving, almost-sure-preserving ($\bowtie = \geq$, $\eta = 1$), and positive-preserving ($\bowtie = >$, $\eta = 0$).

A reduction *restriction-preserving* if when G is one-sided complete-observation, then so is G', when G is complete-observation, then so is G', and when G is turn-based, then so is G'. We say that a reduction is computable in *polynomial time* (resp., in *exponential time*) if the game G' can be constructed in polynomial time (resp., in exponential time) from G (assuming a reasonable encoding of games, such as explicit lists of binary-encoded states, observations, actions, and transitions, and rational probabilities encoded in binary).

An overview of the class of games for which randomness is for free in the transition function (which we establish in this section) is given in Figure 3.

3.1. *Separation of probability and interaction*

A concurrent game of partial observation G satisfies the *interaction separation* condition if the following restrictions are satisfied (see also Figure 4): the state space S can be partitioned into (S_A, S_P) such that (1) $\delta : S_A \times A_1 \times A_2 \to S_P$, and (2) $\delta : S_P \times A_1 \times A_2 \to \mathcal{D}(S_A)$ such that for all $s \in S_P$ and all $s' \in S_A$, and for all a_1, a_2, a'_1, a'_2 we have $\delta(s, a_1, a_2)(s') = \delta(s, a'_1, a'_2)(s') = \delta(s, -, -)(s')$. In other words, the choice of actions (or the interaction) of the players takes place at states in S_A and actions determine a unique successor state in S_P, and the transition function at S_P is probabilistic and independent of the choice of the players. In this section, we present a reduction of each class of games to the corresponding class satisfying interaction separation, and we present a reduction to games with uniform transition probabilities.

Reduction to interaction separation. Let $G = \langle S, A_1, A_2, \delta, \mathcal{O}_1, \mathcal{O}_2 \rangle$ be a concurrent game of partial observation with an objective φ. We obtain a concurrent game of partial observation $G' = \langle S_A \cup S_P, A_1, A_2, \delta', \mathcal{O}'_1, \mathcal{O}'_2 \rangle$ where $S_A = S$, $S_P = S \times A_1 \times A_2$, and:

- *Observations.* For $i \in \{1, 2\}$, if $\mathcal{O}_i = \{\{s\} \mid s \in S\}$, then $\mathcal{O}'_i = \{\{s'\} \mid s' \in S_A \cup S_P\}$; otherwise $\mathcal{O}'_i = \{o \cup o \times A_1 \times A_2 \mid o \in \mathcal{O}_i\}$.
- *Transition function.* The transition function is as follows:
    1. We have the following three cases: (a) if s is a player 1 turn-based state, then pick an action a_2^* and for all a_2 let $\delta'(s, a_1, a_2) = (s, a_1, a_2^*)$; (b) if s is a player 2 turn-based state, then pick an action a_1^* and for all a_1 let $\delta'(s, a_1, a_2) = (s, a_1^*, a_2)$; and (c) otherwise, $\delta'(s, a_1, a_2) = (s, a_1, a_2)$;
    2. for all $(s, a_1, a_2) \in S_P$ we have $\delta'((s, a_1, a_2), -, -)(s') = \delta(s, a_1, a_2)(s')$.
- *Objective mapping.* Given the objective φ in G we obtain the objective $\varphi' = \{s_0 s'_0 s_1 s'_1 \ldots \mid s_0 s_1 \ldots \in \varphi\}$ in G'.

It is easy to map observation-based strategies of the game G to observation-based strategies in G' and vice-versa to preserve satisfaction of φ and φ' in G and G', respectively. Then we have the following theorem.

Theorem 2. *There exists a reduction from the class of partial-observation concurrent games (*PaC *games) to the class of* PaC *games with interaction separation such that this reduction is*

1. *threshold-preserving,*
2. *restriction-preserving, and*
3. *computable in polynomial time.*

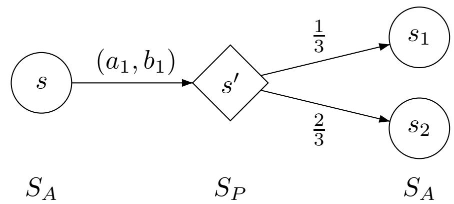

Figure 4: Example of interaction separation for $\delta(s, a_1, b_1)(s_1) = \frac{1}{3}$ and $\delta(s, a_1, b_1)(s_2) = \frac{2}{3}$.

Since the reduction is restriction-preserving, we have a reduction that separates the interaction and probabilistic transition maintaining the restriction of observation and mode of interaction.

Uniform-n-ary concurrent games. The class of *uniform-n-ary games* is the special class of games satisfying interaction separation and such that for every state $s \in S_P$ the probability $\delta(s, -, -)(s')$ to a successor state s' is a multiple of $\frac{1}{n}$. It follows from the results of [28] that every CoC game with rational transition probabilities can be reduced in polynomial time to an equivalent polynomial-size uniform-binary (i.e., $n = 2$) CoC game for all parity objectives. The reduction is achieved by adding dummy states to simulate the probability, and the reduction extends to all objectives (in the reduced game we need to consider the objective whose projection in the original game gives the original objective).

In the case of partial information, the reduction to uniform-binary games of [28] does not work. To see this, consider Figure 5 where two probabilistic states s_1, s_2 have the same observation (i.e., $\mathsf{obs}_1(s_1) = \mathsf{obs}_1(s_2)$) and the outgoing probabilities are $\langle \frac{1}{4}, \frac{3}{4} \rangle$ from s_1 and $\langle \frac{1}{3}, \frac{2}{3} \rangle$ from s_2. The corresponding uniform-binary game (given in Figure 5) is not equivalent to the original game because the number of steps needed to simulate the probabilities is not always the same from s_1 and from s_2. From s_1 two steps are always sufficient, while from s_2 more than two steps may be necessary (with probability $\frac{1}{4}$). Therefore with probability $\frac{1}{4}$, player 1 observing more than 2 steps would infer that the game was for sure in s_2, thus artificially improving his knowledge and increasing his value function.

Therefore in the case of a partial-observation game G satisfying interaction separation, we present a reduction to a uniform-n-ary game G' where $n = 1/r$ where r is the greatest common divisor of all probabilities in the original game G (a rational r is a divisor of a rational p if $p = q \cdot r$ for some integer q). Note that the number $n = 1/r$ is an integer. We denote by $[n]$ the set $\{0, 1, \ldots, n-1\}$. For a probabilistic state $s \in S_P$, we define the n-tuple $\mathsf{Succ}(s) = \langle s'_0, \ldots, s'_{n-1} \rangle$ in which each state $s' \in S$ occurs $n \cdot \delta(s, -, -)(s')$ times. Then, we can view the transition relation $\delta(s, -, -)$ as a function assigning the same probability $r = 1/n$ to each element of $\mathsf{Succ}(s)$ (and then adding up the probabilities of identical elements). Hence it is straightforward to obtain a uniform-n-ary game G'.

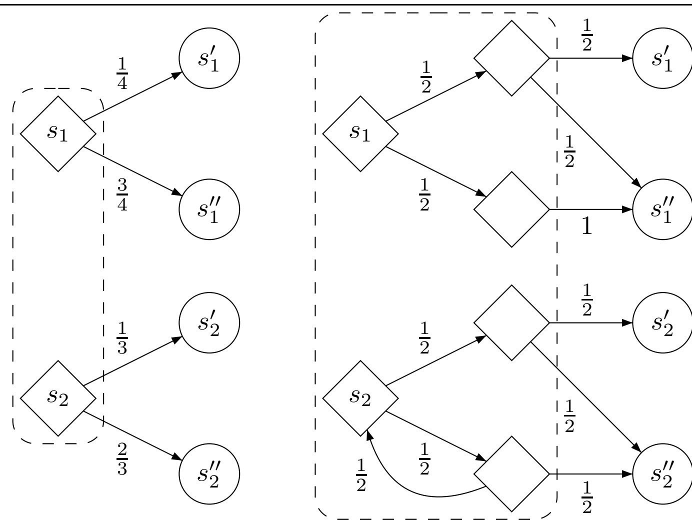

Figure 5: An example showing why the uniform-binary reduction cannot be used with partial observation.

Theorem 3. *There exists a reduction from the class of* PaC *games to the class of uniform-n-ary* PaC *games (where $1/n$ is the greatest common divisor of all probabilities in the original game) such that this reduction is*

1. *threshold-preserving,*
2. *restriction-preserving, and*
3. *computable in exponential time (and in polynomial time for* CoC *games [28]).*

Note that the above reduction is worst-case exponential (because so can be the inverse of the greatest common divisor of the transition probabilities). This is necessary to have the property that all probabilistic states in the game have the same number of successors. This property is crucial because it determines the number of actions available to player 1 in the reductions presented in Section 3.2 and 3.3, and the number of available actions should not differ in states that have the same observation.

3.2. Simulating probability by complete-observation concurrent determinism

In this section, we show that probabilistic states can be simulated by CoC deterministic gadgets (and hence also by OsC and PaC deterministic gadgets). By Theorem 2 and Theorem 3, we focus on uniform-n-ary games. A probabilistic state with uniform probability over the successors is simulated by a complete-observation concurrent deterministic state where the optimal strategy for both players is to play uniformly over the set of available actions.

Theorem 4. *Let $a \in \{\mathsf{Pa}, \mathsf{Os}, \mathsf{Co}\}$ and $b \in \{\mathsf{C}, \mathsf{T}\}$, and let $\mathcal{C} = ab$ and $\mathcal{C}' = a\mathsf{C}$. There exists a reduction from the class of games $\mathcal{G}_{\mathcal{C}}$ to the class of games $\mathcal{G}_{\mathcal{C}'} \cap \mathcal{G}_D$ (thus with deterministic transition function) such that this reduction is*

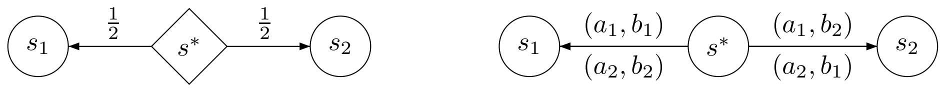

Figure 6: The reduction of uniform-binary CoC games.

1. *threshold-preserving, and*
2. *computable in polynomial time if* $a = \mathsf{Co}$*, and in exponential time if* $a = \mathsf{Pa}$ *or* $a = \mathsf{Os}$*.*

Proof. To prove the result we show that a uniform-n-ary probabilistic state can be simulated by a CoC deterministic gadget. For simplicity we present the details for the case when $n = 2$, and the gadget for the general case is presented later. Our reduction is as follows: we consider a uniform-binary CoC game such that there is only one probabilistic state, and reduce it to a CoC deterministic game. For uniform-binary CoC games with multiple probabilistic states the reduction can be applied to each state one at a time and we would obtain the desired reduction from uniform-binary CoC games to CoC deterministic games. It is easy to see that the reduction can be computed in polynomial time from uniform-n-ary games. The complexity result (item (2) of the theorem) then follows from Theorem 2 and Theorem 3.

The reduction is illustrated in Figure 6 and is defined as follows. Consider a uniform-binary CoC game G with a single probabilistic state s^* with two successors s_1 and s_2. Construct the CoC deterministic game G' obtained from G by transforming the state s^* to a concurrent deterministic state as follows: the actions available for player 1 at s^* are a_1 and a_2, and the actions available for player 2 at s^* are b_1 and b_2; the transition function is as follows: $\delta(s^*, a_1, b_1) = \delta(s^*, a_2, b_2) = s_1$ and $\delta(s^*, a_1, b_2) = \delta(s^*, a_2, b_1) = s_2$. Note that the state space of G' is the same as in G, thus $\varphi' = \varphi$. Then for all objectives φ, we show that the reduction is threshold-preserving as follows.

1. First assume that there exists an observation-based strategy σ for player 1 in G such that $\forall \pi \in \Pi_G^O : \mathrm{Pr}_s^{\sigma,\pi}(\varphi) \bowtie \eta$ for some arbitrary $\eta \in \mathbb{R}$, $s \in S$, and $\bowtie \in \{>, \geq\}$, and construct a strategy σ' for player 1 in G' as follows: the strategy σ' copies the strategy σ for all histories other than when the current state is s^*, and if the current state is s^*, then the strategy σ' plays the actions a_1 and a_2 uniformly with probability $\frac{1}{2}$. Given the strategy σ', if the current state is s^*, then for any probability distribution over player 2's actions b_1 and b_2, the successor states are s_1 and s_2 with probability $\frac{1}{2}$ (i.e., it plays exactly the role of state s^* in G). It follows that for all strategies π' of player 2 in G', there is a strategy π in G (that plays like π' for all histories in G) such that $\mathrm{Pr}_s^{\sigma,\pi}(\varphi) = \mathrm{Pr}_s^{\sigma',\pi'}(\varphi)$ and thus $\mathrm{Pr}_s^{\sigma',\pi'}(\varphi) \bowtie \eta$.
2. Second assume that there exists an observation-based strategy σ' for player 1 in G' such that $\forall \pi' \in \Pi_{G'}^O : \mathrm{Pr}_s^{\sigma',\pi'}(\varphi) \bowtie \eta$ for some arbitrary $\eta \in \mathbb{R}$, $s \in S$,

and $\bowtie \in \{>, \geq\}$, and consider the strategy σ for player 1 in G that plays like σ' for all histories in G. Assume towards contradiction that against σ there exists a strategy $\pi \in \Pi_G^O$ such that $\neg\mathrm{Pr}_s^{\sigma,\pi}(\varphi) \bowtie \eta$. Then consider the strategy π' in G' that copies the strategy π for all histories other than when the current state is s^*, and if the current state is s^*, then the strategy π' plays the actions b_1 and b_2 uniformly with probability $\frac{1}{2}$. Given the strategy π' in G', if the current state is s^*, then for any probability distribution over player 1's actions a_1 and a_2, the successor states are s_1 and s_2 with probability $\frac{1}{2}$ (i.e., it plays exactly the role of state s^* in G). It follows that $\mathrm{Pr}_s^{\sigma',\pi'}(\varphi) = \mathrm{Pr}_s^{\sigma,\pi}(\varphi)$ and thus $\neg\mathrm{Pr}_s^{\sigma',\pi'}(\varphi) \bowtie \eta$, in contradiction with the assumption on σ'. Therefore, such a strategy π cannot exist, and we have $\mathrm{Pr}_s^{\sigma,\pi}(\varphi) \bowtie \eta$ for all $\pi \in \Pi_G^O$, which concludes the proof that the reduction is threshold-preserving.

Gadget for uniform-n-ary probability reduction. We now show how to simulate a probabilistic state s^*, with n successors $s_0, s_1, \ldots, s_{n-1}$ such that the transition probability is $1/n$ to each of the successors, by a concurrent deterministic state. In the concurrent deterministic state s^* there are n actions $a_0, a_1, \ldots, a_{n-1}$ available for player 1 and n actions $b_0, b_1, \ldots, b_{n-1}$ available for player 2. The transition function is as follows: for $0 \leq i < n$ and $0 \leq j < n$ we have $\delta(s^*, a_i, b_j) = s_{(i+j) \mod n}$. Intuitively, the transition function matrix is obtained as follows: the first row is filled with states $s_0, s_1, \ldots, s_{n-1}$, and from a row i, the row $i+1$ is obtained by moving the state of the first column of row i to the last column in row $i+1$ and left-shifting by one position all the other states; the construction is illustrated on an example with $n = 4$ successors in (1). The construction ensures that in every row and every column each state $s_0, s_1, \ldots, s_{n-1}$ appears exactly once. It follows that if player 1 plays all actions uniformly at random, then against any probability distribution of player 2 the successor states are $s_0, s_1, \ldots, s_{n-1}$ with probability $1/n$ each; and a similar result holds if player 2 plays all actions uniformly at random. The correctness of the reduction for uniform-n-ary probabilistic state is then exactly as for the case of $n = 2$.

$$\begin{bmatrix} s_0 & s_1 & s_2 & s_3 \\ s_1 & s_2 & s_3 & s_0 \\ s_2 & s_3 & s_0 & s_1 \\ s_3 & s_0 & s_1 & s_2 \end{bmatrix} \tag{1}$$

The desired result follows. ∎

3.3. Simulating probability by one-sided complete-observation turn-based determinism

We show that probabilistic states can be simulated by OsT (one-sided complete-observation turn-based) states, and by Theorem 2 we consider games that satisfy interaction separation. The reduction is illustrated in Figure 7: each probabilistic state s is transformed into a player-2 state with n successor player-1 states (where n is chosen such that the probabilities from s are integer multiples of $1/n$, in the example $n = 3$). Because all successors of s have the same observation, player 1 has no advantage in playing after player 2, and because by playing all actions uniformly at random each player can unilaterally decide to simulate the probabilistic state, the value and properties of strategies of the game are preserved.

Theorem 5. *Let $a \in \{\mathsf{Pa}, \mathsf{Os}, \mathsf{Co}\}$ and $b \in \{\mathsf{C}, \mathsf{T}\}$, and let $a' = \mathsf{Os}$ if $a = \mathsf{Co}$, and $a' = a$ otherwise. Let $\mathcal{C} = ab$ and $\mathcal{C}' = a'b$. There exists a reduction from the class of games $\mathcal{G}_{\mathcal{C}}$ to the class of games $\mathcal{G}_{\mathcal{C}'} \cap \mathcal{G}_D$ (thus with deterministic transition function) such that this reduction is*
1. *threshold-preserving, and*
2. *computable in polynomial time if $a = \mathsf{Co}$, and in exponential time if $a = \mathsf{Pa}$ or $a = \mathsf{Os}$.*

Proof. First, we present the proof for $a \neq \mathsf{Pa}$, assuming that player 2 has complete observation. A similar construction where player-1 instead of player-2 has complete observation is obtained symmetrically. Let $G = \langle S_A \cup S_P, A_1, A_2, \delta, \mathcal{O}_1 \rangle$ and assume w.l.o.g. (according to Theorem 2 and Theorem 3) that G satisfies interaction separation (i.e., states in S_A are deterministic states, and S_P are probabilistic states) and G is uniform-n-ary, i.e. all probabilities are equal to $\frac{1}{n}$. For each probabilistic state $s \in S_P$, let $\mathsf{Succ}(s) = \langle s'_0, \dots, s'_{n-1} \rangle$ be the n-tuple of states such that $\delta(s, -, -)(s'_i) = \frac{1}{n}$ for each $1 \leq i \leq n$.

We present a reduction that replaces the probabilistic states in G by a gadget with player-1 and player-2 turn-based states. From G, we construct the one-sided complete-observation game G' where player-2 has complete observation. The game $G' = \langle S', A'_1, A'_2, \delta', \mathcal{O}'_1 \rangle$ is defined as follows: $S' = S \cup (S \times [n]) \cup \{\mathsf{sink}\}$, $A'_1 = A_1 \cup [n]$, $A'_2 = A_2 \cup [n]$, $\mathcal{O}'_1 = \{o \cup (o \times [n]) \mid o \in \mathcal{O}_1\}$, and δ' is obtained from δ by applying the following transformation for each state $s \in S$:

1. if s is a deterministic state in G, then $\delta'(s, a, b) = \delta(s, a, b)$ for all $a \in A_1, b \in A_2$, and $\delta'(s, i, j) = \mathsf{sink}$ for all $i, j \in [n]$;

2. if s is a probabilistic state in G, then s is a player-2 state in G' and for all $i, j \in [n]$ we define $\delta'(s, -, i) = (s, i)$ and $\delta'((s, i), j, -) = s'_k$ such that s'_k is the element in position k in $\mathsf{Succ}(s)$ with $k = i + j \mod n$ (and let $\delta'(s, -, b) = \delta'((s, i), a, -) = \delta'(\mathsf{sink}, -, -) = \mathsf{sink}$ for all $a \in A_1, b \in A_2$).

Note that turn-based states in G remain turn-based in G' and the states (s, i) are player-1 states with the same observation as s. As usual, the objective φ' is defined as the set of plays in G' whose projection on S^ω belongs to φ.

Intuitively, each player in G' has the possibility to ensure exact simulation of the probabilistic states of G by playing actions in $[n]$ uniformly at random. For instance, if player 1 does so, then irrespective of the (possibly randomized) choice of player 2 among the states $(s, 1), \dots, (s, n)$, the states in $\mathsf{Succ}(s)$ are reached with probability $1/n$, as in G. The same property holds if player 2 plays the actions in $[n]$ uniformly at random, no matter what player 1 does. Therefore, by arguments similar to the proof of Theorem 4, player 1 can ensure the objective φ' in G' is satisfied with the same probability as φ in G, against any strategy of player 2, and the reduction is threshold-preserving.

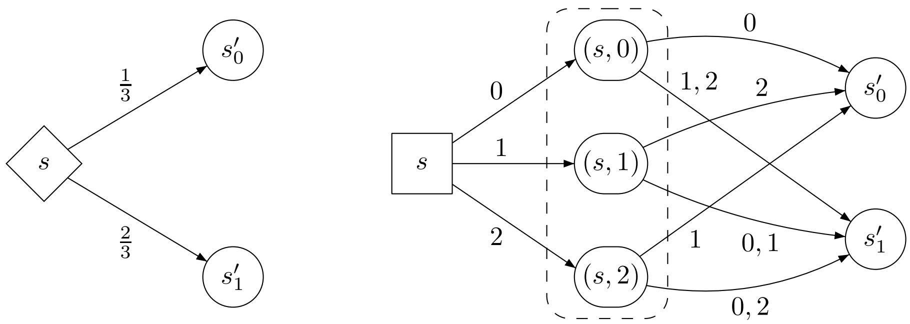

Figure 7: For the probabilistic state s (on the left), we have $\mathsf{Succ}(s) = \langle s'_0, s'_1, s'_1 \rangle$ and $n = 3$ is the gcd of the probabilities denominators. Therefore, we apply the reduction of Theorem 5 to obtain the turn-based game on the right, where s is a player-2 state.

The reduction can be easily adapted to the case $a = \mathsf{Pa}$ of games with partial information for both players. Since the construction of G' is polynomial, the complexity result (item (2) of the theorem) follows from Theorem 2 and Theorem 3. ∎

3.4. Impossibility Results

We have shown that for CoC games and OsT games, randomness is for free in the transition function. We complete the picture (Figure 2) by showing that for CoT (complete-observation turn-based) games, randomness in the transition function cannot be obtained for free.

Remark 1 (Role of probabilistic transition in CoT games and POMDPs). *It follows from the result of Martin [20] that for all* CoT *deterministic games and all objectives, the values are either 1 or 0; however, even* MDP*s with reachability objectives can have values in the interval* $[0, 1]$ *(not value 0 and 1 only). It follows that "randomness in the transition function" cannot be replaced by "randomness in the strategies" in* CoT *deterministic games. For* POMDP*s, we show in Theorem 7 that pure strategies are sufficient, and it follows that for* POMDP*s with deterministic transition function the values are 0 or 1, and since* MDP*s with reachability objectives can have values other than 0 and 1 it follows that randomness in the transition function cannot be obtained for free for* POMDP*s. The probabilistic transitions also play an important role in the complexity of solving games in case of* CoT *games: for example,* CoT *deterministic games with reachability objectives can be solved in linear time, but with probabilistic transition function the problem is in NP* $\cap$ *coNP and no polynomial-time algorithm is known. In contrast, for* CoC *games we present a polynomial-time reduction from probabilistic to deterministic transition function. Table 2 summarizes our results characterizing the classes of games where randomness in the transition function can be obtained for free.*

	2½-player			1½-player	
	complete	one-sided	partial	MDP	POMDP
turn-based	not (Rmk. 1)	free (Th. 5)	free (Th. 5)	not (Rmk. 1)	not (Rmk. 1)
concurrent	free (Th. 4)	free (Th. 4)	free (Th. 4)	(NA)	(NA)

Table 2: When randomness is for free in the transition function. In particular, probabilities can be eliminated in all classes of 2-player games except complete-observation turn-based games. In the table, Rmk. 1 refers to Remark 1, Th. 5 refers to Theorem 5, and Th. 4 refers to Theorem 4.

3.5. Concurrency for free

The idea of the reduction in Theorem 5 can be extended to prove that concurrency is for free in one-sided complete-observation games, i.e., we present a polynomial reduction of OsC games to OsT games, and from PaC games to PaT games.

Theorem 6. *There exists a reduction from* OsC *games to* OsT *games, and from* PaC *games to* PaT *games, such that these reductions are*

1. *threshold-preserving, and*
2. *computable in polynomial time.*

Proof. We present the reduction from OsC games to OsT games, for the case where player 1 has complete information. The reduction for one-sided games where player 2 has complete information is symmetric. Finally, the reduction from PaC games to PaT games is obtained analogously.

Let $G = \langle S, A_1, A_2, \delta, \mathcal{O}_2 \rangle$ be a OsC game where player 1 has complete information, and we construct a OsT game $G' = \langle S', A_1, A_2, \delta', \mathcal{O}'_1 \rangle$ as follows:

1. $S' = S \cup (S \times A_1)$,
2. $\mathcal{O}'_2 = \{o \cup (o \times A_1) \mid o \in \mathcal{O}_2\}$, and
3. δ' is defined as follows, for each state $s \in S$ and actions $a \in A_1$, $b \in A_2$: $\delta'(s, a, -) = (s, a)$ and $\delta'((s, a), -, b) = \delta(s, a, b)$.

Hence the transition function δ' lets player 1 play first an action a, then player 2 plays an action b, and the successor state of s is chosen according to the transition relation $\delta(s, a, b)$ from the original game. As usual, the objective $\varphi' = \{s_0(s_0, a_0)s_1(s_1, a_1)\cdots \mid s_0 s_1 \cdots \in \varphi \wedge \forall i \geq 0 : a_i \in A_1\}$ in G' requires that the projection of a play on S^ω satisfies φ. Since player 1 plays first in G', player 1 can achieve the objective φ' in G' with at most the same probability as for φ in G, and since for all $s \in S$ and actions $a \in A_1$, the states s and (s, a) are indistinguishable for player 2, player 2 does not know the last action chosen by player 1 and therefore does not gain any advantage in playing after player 1 rather than concurrently. Therefore the reduction is threshold-preserving and since it is computable in polynomial time, the result follows. ∎

Role of concurrency in complete-observation games. We have shown that concurrency can be obtained for free in partial-observation games (OsT and PaT games). In contrast, for complete-observation games, the value is irrational in general for concurrent games with deterministic transitions (CoC deterministic games) [11], while the value is always rational in turn-based stochastic games with rational probabilities (CoT stochastic games) [12]. This rules out any value-preserving reduction of CoC (deterministic) games to CoT (stochastic) games with rational probabilities.

4. Randomness for Free in Strategies

In this section we present our results for randomness for free in strategies. We start with a remark.

Remark 2 (Randomness in strategies). *It is known from the results of [15] that in* CoC *games randomized strategies are more powerful than pure strategies: values achieved by pure strategies are lower than values achieved by randomized strategies and randomized almost-sure winning strategies may exist whereas no pure almost-sure winning strategy exists. Similar results also hold in the case of* OsT *games (see [10] for an example, also see Example 1). By contrast we show that in POMDPs, restricting the set of strategies to pure strategies does not decrease the value nor affect the existence of almost-sure and positive winning strategies.*

We start with a lemma, and then present our results precisely in Theorem 7. The main argument in the proof of Lemma 1 relies on showing that the value $\Pr_s^\sigma(\varphi)$ of any randomized observation-based strategy σ is equal to the average of the values $\Pr_s^{\sigma_i}(\varphi)$ of (uncountably many) pure observation-based strategies σ_i. Therefore, one of the pure strategies σ_i has to achieve at least the value of the randomized strategy σ. The theory of integration and Fubini's theorem make this argument precise.

Lemma 1. *Let G be a* POMDP *(with countable state space S), let $s_* \in S$ be an initial state, and let $\varphi \subseteq S^\omega$ be an objective. For every randomized observation-based strategy $\sigma \in \Sigma_G^O$ there exists a* pure *observation-based strategy $\sigma_P \in \Sigma_G^P \cap \Sigma_G^O$ such that $\Pr_{s_*}^\sigma(\varphi) \leq \Pr_{s_*}^{\sigma_P}(\varphi)$.*

Proof. Let $G = \langle S, A_1, \delta, \mathcal{O}_1 \rangle$ be a POMDP (remember that A_2 is a singleton in POMDPs and therefore $\mathcal{O}_2$ is irrelevant), let $\sigma : \mathsf{Prefs}(G) \to \mathcal{D}(A_1)$ be a randomized observation-based strategy, and fix $s_* \in S$ an initial state.

To simplify notations, we suppose that $A_1 = \{0, 1\}$ contains only two actions, and that given a state $s \in S$ and an action $a \in \{0, 1\}$ there are only two possible successors $L(s, a) \in S$ and $R(s, a) \in S$ chosen with respective probabilities $\delta(s, a, L(s, a))$ and $\delta(s, a, R(s, a)) = 1 - \delta(s, a, L(s, a))$. The proof for an arbitrary finite set of actions and more than two successors is essentially the same, with more complicated notations.

There is a natural way to "derandomize" the randomized strategy σ. Fix an infinite sequence $x = (x_n)_{n\in\mathbb{N}} \in [0, 1]^\omega$ and define the pure strategy $\sigma_x : \mathsf{Prefs}(G) \to A_1$ as follows. For every play prefix $h = s_0\, a_1\, s_1\, a_2\, s_2 \ldots s_n$, let

$$\sigma_x(h) = \begin{cases} 0 & \text{if } x_n \leq \sigma(h)(0) \\ 1 & \text{otherwise.} \end{cases}$$

Intuitively, the sequence x fixes in advance the sequence of results of coin tosses used for playing with σ. Note that if σ is observation-based, then for every sequence x the strategy σ_x is both observation-based and pure.

To prove the lemma, we show that $[0,1]^\omega$ can be equipped with a probability measure ν such that the mapping $x \mapsto \mathrm{Pr}_{s_*}^{\sigma_x}(\varphi)$ from $[0,1]^\omega$ to $[0,1]$ is measurable, and:

$$\mathrm{Pr}_{s_*}^{\sigma}(\varphi) = \int_{x\in[0,1]^\omega} \mathrm{Pr}_{s_*}^{\sigma_x}(\varphi)\, d\nu(x) \ . \tag{2}$$

Suppose that (2) holds. Then there exists $x \in [0,1]^\omega$ (actually many x's) such that $\mathrm{Pr}_{s_*}^{\sigma}(\varphi) \leq \mathrm{Pr}_{s_*}^{\sigma_x}(\varphi)$ and since strategy σ_x is deterministic, this proves the lemma.

To complete the proof, it is thus enough to construct a probability measure ν on $[0,1]^\omega$ such that (2) holds.

We start with the definition of the probability measure ν. The set $[0,1]^\omega$ is equipped with the sigma-field generated by *sequence-cylinders* which are defined as follows. For every finite sequence $x = x_0, x_1, \ldots, x_n \in [0,1]^*$ the sequence-cylinder $\mathcal{C}(x)$ is the subset $[0,x_0] \times [0,x_1] \times \ldots \times [0,x_n] \times [0,1]^\omega \subseteq [0,1]^\omega$. According to Tulcea's theorem [4], there is a unique product probability measure ν on $[0,1]^\omega$ such that $\nu(\mathcal{C}(\epsilon)) = 1$ and for every sequence $x_0, \ldots, x_n, x_{n+1}$ in $[0,1]$,

$$\nu(\mathcal{C}(x_0, \ldots, x_n, x_{n+1})) = x_{n+1} \cdot \nu(\mathcal{C}(x_0, \ldots, x_n)) \ .$$

Now that ν is defined, it remains to prove that the mapping $x \mapsto \mathrm{Pr}_{s_*}^{\sigma_x}(\varphi)$ from $[0,1]^\omega$ to $[0,1]$ is measurable and that (2) holds. For that, we introduce the following mapping:

$$f_{s_*,\sigma} : [0,1]^\omega \times [0,1]^\omega \to (SA_1)^\omega,$$

that associates with every pair of sequences $((x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}})$ the infinite history $h = s_0\, a_1\, s_1\, a_2 \ldots \in (SA_1)^\omega$ defined recursively as follows. First $s_0 = s_*$, and for every $n \in \mathbb{N}$,

$$a_{n+1} = \begin{cases} 0 & \text{if } x_n \leq \sigma(s_0\, a_1\, s_1 \cdots s_n)(0), \\ 1 & \text{otherwise.} \end{cases}$$

$$s_{n+1} = \begin{cases} L(s_n, a_{n+1}) & \text{if } y_n \leq \delta(s_n, a_{n+1}, L(s_n, a_{n+1})), \\ R(s_n, a_{n+1}) & \text{otherwise.} \end{cases}$$

Intuitively, $(x_n)_{n\in\mathbb{N}}$ fixes in advance the coin tosses used by the strategy, while $(y_n)_{n\in\mathbb{N}}$ takes care of the coin tosses used by the probabilistic transitions, and $f_{s_*,\sigma}$ produces the resulting description of the play. Thanks to the mapping $f_{s_*,\sigma}$, randomness related to the use of the randomized strategy σ is separated from randomness due to transitions of the game, which allows to represent the randomized strategy σ by mean of a probability measure over the set of pure strategies $\{\sigma_x \mid x \in [0,1]^\omega\}$.

We equip both sets $(SA_1)^\omega$ and $[0,1]^\omega \times [0,1]^\omega$ with sigma-fields that make $f_{s_*,\sigma}$ measurable. First, $(SA_1)^\omega$ is equipped with the sigma-field generated by cylinders, defined as follows. An *action-cylinder* is a subset $\mathcal{C}(h) \subseteq (SA_1)^\omega$ such that $\mathcal{C}(h) = h(SA_1)^\omega$ for some $h \in (SA_1)^*$. A *state-cylinder* is a subset $\mathcal{C}(h) \subseteq (SA_1)^\omega$ such that $\mathcal{C}(h) = h(A_1S)^\omega$ for some $h \in (SA_1)^*S$. The set of *cylinders* is the union of the sets of action-cylinders and state-cylinders. Second, $[0,1]^\omega \times [0,1]^\omega$ is equipped with the sigma-field generated by products of sequence-cylinders. Checking that $f_{s_*,\sigma}$ is measurable is an elementary exercise.

Now we define two probability measures μ and μ' on $(SA_1)^\omega$ and prove that they coincide. On one hand, the measurable mapping $f_{s_*,\sigma} : [0,1]^\omega \times [0,1]^\omega \to (SA_1)^\omega$ defines naturally a probability measure μ' on $(SA_1)^\omega$. Equip the set $[0,1]^\omega \times [0,1]^\omega$ with the product measure $\nu \times \nu$. Then for every measurable subset $B \subseteq (SA_1)^\omega$,

$$\mu'(B) = (\nu \times \nu)(f_{s_*,\sigma}^{-1}(B)) \ .$$

On the other hand, the strategy σ and the initial state s_* naturally define another probability measure μ on $(SA_1)^\omega$. According to Tulcea's theorem [4], there exists a unique product probability measure μ on $(SA_1)^\omega$ such that $\mu(\mathcal{C}(s_*)) = 1$, $\mu(\mathcal{C}(s)) = 0$ for $s \in S \setminus \{s_*\}$, and for $h = s_0\, a_1\, s_1\, a_2 \cdots s_n \in (SA_1)^* S$ and $(a,t) \in A_1 \times S$,

$$\begin{aligned}\mu(\mathcal{C}(ha)) &= \mu(\mathcal{C}(h)) \cdot \sigma(h)(a)\\ \mu(\mathcal{C}(hat)) &= \mu(\mathcal{C}(ha)) \cdot \delta(s_n,a,t).\end{aligned}$$

To prove that μ and μ' coincide, it is enough to prove that μ and μ' coincide on the set of cylinders, that is for every cylinder $\mathcal{C}(h) \subseteq (SA_1)^\omega$,

$$\mu(\mathcal{C}(h)) = (\nu \times \nu)(f_{s_*,\sigma}^{-1}(\mathcal{C}(h))) \ . \tag{3}$$

This is obvious for $h = s_*$ and $h = s \in S \setminus \{s_*\}$. The general case goes by induction. Let $h = s_0\, a_1\, s_1\, a_2 \cdots s_n \in (SA_1)^* S$ and $(a,t) \in A_1 \times S$. Let $I = [0,1]$. Let $I_a = [0, \sigma(h)(a)]$ if $a = 0$ and $I_a = [\sigma(h)(a), 1]$ if $a = 1$. Let $I_t = [0, \delta(s_n,a,t)]$ if $t = L(s_n,a)$ and $I_t = [\delta(s_n,a,t), 1]$ if $t = R(s_n,a)$. Then:

$$\begin{aligned}\mu(\mathcal{C}(ha) \mid \mathcal{C}(h)) &= \sigma(h)(a)\\ &= (\nu \times \nu)((I \times I)^n (I_a \times I)(I \times I)^\omega)\\ &= (\nu \times \nu)(f_{s_*,\sigma}^{-1}(\mathcal{C}(ha)) \mid f_{s_*,\sigma}^{-1}(\mathcal{C}(h)))\\ \mu(\mathcal{C}(hat) \mid \mathcal{C}(ha)) &= \delta(s_n,a,t)\\ &= (\nu \times \nu)((I \times I)^n (I \times I_t)(I \times I)^\omega)\\ &= (\nu \times \nu)(f_{s_*,\sigma}^{-1}(\mathcal{C}(hat)) \mid f_{s_*,\sigma}^{-1}(\mathcal{C}(ha))) \ ,\end{aligned}$$

which proves that (3) holds for every cylinder $\mathcal{C}(h)$.

Now all the tools needed to prove (2) have been introduced, and we can state the main relation between $f_{s_*,\sigma}$ and $\mathrm{Pr}_{s_*}^\sigma(\varphi)$. Let $\varphi' \subseteq (SA_1)^\omega$ be the set of histories $s_0\, a_1\, s_1\, a_2 \ldots$ such that $s_0 s_1 \cdots \in \varphi$, and let $\mathbf{1}_\varphi$ and $\mathbf{1}_{\varphi'}$ be the indicator functions of φ and φ'. Then:

$$\begin{aligned}\mathrm{Pr}_{s_*}^\sigma(\varphi) &= \int_{p \in S^\omega} \mathbf{1}_\varphi(p)\, d\mathrm{Pr}_{s_*}^\sigma(p) = \int_{p \in (SA_1)^\omega} \mathbf{1}_{\varphi'}(p)\, d\mu(p) = \int_{p \in (SA_1)^\omega} \mathbf{1}_{\varphi'}(p)\, d\mu'(p)\\ &= \int_{(x,y) \in [0,1]^\omega \times [0,1]^\omega} \mathbf{1}_{\varphi'}(f_{s_*,\sigma}(x,y))\, d(\nu \times \nu)(x,y)\\ &= \int_{x \in [0,1]^\omega} \left(\int_{y \in [0,1]^\omega} \mathbf{1}_{\varphi'}(f_{s_*,\sigma}(x,y))\, d\nu(y) \right) d\nu(x) \ ,\end{aligned} \tag{4}$$

	2½-player			1½-player	
	complete	one-sided	partial	MDP	POMDP
turn-based	$\epsilon > 0$ (Th. 1)	not (Rmk. 2)	not (Rmk. 2)	$\epsilon \geq 0$ (Th. 7)	$\epsilon \geq 0$ (Th. 7)
concurrent	not (Rmk. 2)	not (Rmk. 2)	not (Rmk. 2)	(NA)	(NA)

Table 3: When pure (ϵ-optimal) strategies are as powerful as randomized strategies. The case $\epsilon = 0$ in complete-observation turn-based games is open. In the table, Th. 1 refers to Theorem 1, Rmk. 2 refers to Remark 2, Th. 7 refers to Theorem 7.

where the first and second equalities are by definition of $\Pr_{s_*}^{\sigma}(\varphi)$, the third equality holds because $\mu = \mu'$, the fourth equality is a basic property of image measures, and the last equality holds by Fubini's theorem [4] that we can use since $\mathbf{1}_{\varphi'} \circ f_{s_*,\sigma}$ is positive.

To complete the proof, we show that for every $x \in [0,1]^{\omega}$,

$$\int_{y \in [0,1]^{\omega}} \mathbf{1}_{\varphi'}(f_{s_*,\sigma}(x,y))\, d\nu(y) = \Pr_s^{\sigma_x}(\varphi), \tag{5}$$

Equation (4) holds for every observation-based strategy σ, hence in particular for strategy σ_x. But strategy σ_x has the following property: for every $x' \in]0,1[^{\omega}$ and every $y \in [0,1]^{\omega}$, $f_{s_*,\sigma_x}(x',y) = f_{s_*,\sigma}(x,y)$. Together with (4), this gives (5). This completes the proof, since (4) and (5) immediately give (2). ∎

We obtain the following result as a consequence of Lemma 1.

Theorem 7. *Let G be a* POMDP *(with countable state space S), let $s_* \in S$ be an initial state, and let $\varphi \subseteq S^{\omega}$ be an objective. Then the following assertions hold:*

1. $\sup_{\sigma \in \Sigma_G^O} \Pr_{s_*}^{\sigma}(\varphi) = \sup_{\sigma \in \Sigma_G^O \cap \Sigma_G^P} \Pr_{s_*}^{\sigma}(\varphi)$.
2. *If there is a randomized optimal (resp., almost-sure winning, positive winning) strategy for φ from s_*, then there is a pure optimal (resp., almost-sure winning, positive winning) strategy for φ from s_*.*

Theorem 7 shows that the result of Theorem 1 can be generalized to POMDPs, and a stronger result (item (2) of Theorem 7) can be proved for POMDPs (and MDPs as a special case). It remains open whether a result similar to item (2) of Theorem 7 can be proved for CoT stochastic games. Note that it was already shown in [13, Example 1] that in CoT stochastic games with Borel objectives optimal strategies need not exist. The results summarizing when randomness can be obtained for free for strategies is shown in Table 3.

Undecidability result for POMDPs. The results of [2] show that the emptiness problem for finite-state probabilistic coBüchi (resp., Büchi) automata under the almost-sure (resp., positive) semantics [2] is undecidable. As a consequence it follows that for finite-state POMDPs the problem of deciding if there is a pure observation-based almost-sure (resp., positive) winning strategy for coBüchi (resp., Büchi) objectives is undecidable, and as a consequence of Theorem 7 we obtain an analogous undecidability result for randomized strategies. The undecidability result holds even if the coBüchi (resp., Büchi) objectives is visible.

Corollary 1. *Let G be a finite-state* POMDP *with initial state s_* and let $\mathcal{T} \subseteq S$ be a subset of states (or union of observations). Whether there exists a pure or randomized almost-sure winning strategy for player 1 from s_* in G for the objective* $\mathsf{coBuchi}(\mathcal{T})$ *is undecidable; and whether there exists a pure or randomized positive winning strategy for player 1 from s_* in G for the objective* $\mathsf{Buchi}(\mathcal{T})$ *is undecidable.*

Undecidability result for one-sided complete-observation turn-based games. The undecidability results of Corollary 1 also holds for finite-state OsT stochastic games (as they subsume finite-state POMDPs as a special case). It follows from Theorem 5 that finite-state OsT stochastic games can be reduced to finite-state OsT deterministic games. The reduction holds for randomized strategies and thus we obtain the first undecidability result for finite-state OsT deterministic games (Corollary 2), solving the open question of [10]. Note that for pure strategies, OsT deterministic games with a parity objective are EXPTIME-complete [25, 10].

Corollary 2. *Let G be a finite-state* OsT *deterministic game with initial state s_* and let $\mathcal{T} \subseteq S$ be a subset of states (or union of observations). Whether there exists a randomized almost-sure winning strategy for player 1 from s_* in G for the objective* $\mathsf{coBuchi}(\mathcal{T})$ *is undecidable; and whether there exists a randomized positive winning strategy for player 1 from s_* in G for the objective* $\mathsf{Buchi}(\mathcal{T})$ *is undecidable.*

5. Conclusion

In this work we have presented a precise characterization for classes of games where randomization can be obtained for free in transition functions and in strategies. As a consequence of our characterization we obtain new undecidability results. The other impact of our characterization is as follows: for the class of games where randomization is free in transition function, future algorithmic and complexity analysis can focus on the simpler class of deterministic games; and for the class of games where randomization is free in strategies, future analysis of such games can focus on the simpler class of pure strategies. Thus our results will be useful tools for simpler analysis techniques in the study of games, as already demonstrated in [6, 7, 8, 9, 16, 17].

Finally, note that it can be expected that randomness would not be for free in both the transition function and the strategies, and the results of this paper show that the classes of games in which randomness is for free in the transition function (Table 2) are those in which randomized strategies are more powerful than pure strategies (Table 3), i.e. randomness is not for free in strategies when randomness is for free in the transition function.

References

[1] R. Alur, T. A. Henzinger, and O. Kupferman. Alternating-time temporal logic. *Journal of the ACM*, 49:672–713, 2002.

[2] C. Baier, N. Bertrand, and M. Größer. On decision problems for probabilistic Büchi automata. In *FoSSaCS*, LNCS 4962, pages 287–301. Springer, 2008.

[3] N. Bertrand, B. Genest, and H. Gimbert. Qualitative determinacy and decidability of stochastic games with signals. In *Proc. of LICS*, pages 319–328. IEEE Computer Society, 2009.

[4] P. Billingsley. *Probability and Measure*. Wiley-Interscience, 1995.

[5] J. R. Büchi and L. H. Landweber. Solving sequential conditions by finite-state strategies. *Transactions of the AMS*, 138:295–311, 1969.

[6] P. Cerný, K. Chatterjee, T. A. Henzinger, A. Radhakrishna, and R. Singh. Quantitative synthesis for concurrent programs. In *CAV*, pages 243–259, 2011.

[7] K. Chatterjee and M. Chmelik. POMDPs under probabilistic semantics. *CoRR*, abs/1408.2058, 2014 (Conference version: UAI, 2013).

[8] K. Chatterjee, M. Chmelik, and M. Tracol. What is decidable about partially observable Markov decision processes with omega-regular objectives. In *CSL*, pages 165–180, 2013.

[9] K. Chatterjee and L. Doyen. Partial-observation stochastic games: How to win when belief fails. *ACM Trans. Comput. Log.*, 15(2):16, 2014.

[10] K. Chatterjee, L. Doyen, T. A. Henzinger, and J.-F. Raskin. Algorithms for omega-regular games of incomplete information. *Logical Methods in Computer Science*, 3(3:4), 2007.

[11] K. Chatterjee and T. A. Henzinger. Semiperfect-information games. In *FSTTCS'05*. LNCS 3821, Springer, 2005.

[12] K. Chatterjee, M. Jurdziński, and T. A. Henzinger. Quantitative stochastic parity games. In *SODA'04*, pages 121–130. SIAM, 2004.

[13] K. Chatterjee, R. Majumdar, and M. Jurdziński. On Nash equilibria in stochastic games. In *CSL'04*, pages 26–40. LNCS 3210, Springer, 2004.

[14] L. de Alfaro and T. A. Henzinger. Interface theories for component-based design. In *EMSOFT'01*, LNCS 2211, pages 148–165. Springer, 2001.

[15] H. Everett. Recursive games. In *Contributions to the Theory of Games III*, volume 39 of *Annals of Mathematical Studies*, pages 47–78, 1957.

[16] H. Gimbert and Y. Oualhadj. Deciding the value 1 problem for $\sharp$-acyclic partially observable Markov decision processes. In *SOFSEM*, pages 281–292, 2014.

[17] J. Goubault-Larrecq and R. Segala. Random measurable selections. In *Horizons of the Mind*, pages 343–362, 2014.

[18] T. A. Henzinger, O. Kupferman, and S. Rajamani. Fair simulation. *Information and Computation*, 173:64–81, 2002.

[19] A. Kechris. *Classical Descriptive Set Theory*. Springer, 1995.

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

[21] R. McNaughton. Infinite games played on finite graphs. *Annals of Pure and Applied Logic*, 65:149–184, 1993.

[22] J.-F. Mertens, S. Sorin, and S. Zamir. Repeated games. *Core Discussion Papers*, 9422, 1994.

[23] A. Pnueli and R. Rosner. On the synthesis of a reactive module. In *POPL'89*, pages 179–190. ACM Press, 1989.

[24] P. J. Ramadge and W. M. Wonham. Supervisory control of a class of discrete-event processes. *SIAM Journal of Control and Optimization*, 25(1):206–230, 1987.

[25] J. H. Reif. The complexity of two-player games of incomplete information. *Journal of Computer and System Sciences*, 29(2):274–301, 1984.

[26] W. Thomas. Languages, automata, and logic. In *Handbook of Formal Languages*, volume 3, Beyond Words, chapter 7, pages 389–455. Springer, 1997.

[27] M. Y. Vardi. Automatic verification of probabilistic concurrent finite-state systems. In *FOCS*, pages 327–338. IEEE Computer Society Press, 1985.

[28] U. Zwick and M. Paterson. The complexity of mean payoff games on graphs. *Theoretical Computer Science*, 158:343–359, 1996.