Theoretical Computer Science 493 (2013) 80–97

Contents lists available at SciVerse ScienceDirect

Theoretical Computer Science

journal homepage: www.elsevier.com/locate/tcs

Determinacy and optimal strategies in infinite-state stochastic reachability games

Václav Brožek

Institute of Theoretical Computer Science, Faculty of Informatics, Masaryk University, Botanická 68a, 602 00, Brno, Czech Republic[1]

ARTICLE INFO

Keywords:
Stochastic games
Reachability
Determinacy
Optimal strategies

ABSTRACT

We consider perfect-information reachability stochastic games for 2 players on countable graphs. Such a game is strongly determined if, whenever we fix an inequality $\sim \in \{>, \geq\}$ and a threshold p, either Player Max has a strategy which forces the value of the game to satisfy $\sim p$ against any strategy of Player Min, or Min has a strategy which forces the opposite against any strategy of Max.

One of our results shows that whenever one of the players has an optimal strategy in every state of a game, then this game is strongly determined. This significantly generalises, e.g., recent results on finitely-branching reachability games. For strong determinacy, our methods are substantially different, based on which player has the optimal strategy, because the roles of the players are not symmetric. We also do not restrict the branching of the games, and where we provide an extension of results for finitely-branching games, we had to overcome significant complications and employ new methods as well.

The other result is finding a subclass of stochastic games where Player Max has an optimal strategy in each state. The subclass is defined by the property that if v is an accumulation point of the set of all values of a game then $v = 0$. These results complement recent results classifying the existence of an optimal strategy for Player Min, and our general strong-determinacy theorem applies here as well. We also apply our results for Max in the context of recently studied One-Counter stochastic games.

This work extends a workshop version of this paper which appeared in GandALF 2011, in particular, we prove a conjecture raised in that paper for the class of all reachability games.

1. Introduction, basic definitions and results

1.1. Background

Two-player turn-based zero-sum stochastic games, simply called "games" in this text, evolve randomly in discrete *transitions* from one of countably many *states* to another. The winning condition is some property of such infinite evolutions. Each state is either owned by Player Max, Player Min, or it is stochastic, and has a fixed set, possibly infinite, of available outgoing transitions. A graphical representation of such a game can be seen, e.g., in Fig. 1: transitions are edges, states are nodes shaped as circles, boxes, or diamonds, depending on whether they are stochastic, owned by Max, or by Min, respectively. The states and transitions define a *game graph*, an infinite path in this graph is called a *run*. The set of runs comes with a product topology over the discrete state space, i.e., open sets are generated by sets of runs sharing a common finite prefix. In stochastic states, the successor is sampled according to a fixed distribution, whereas players choose successors in

E-mail address: vaclav.brozek@gmail.com.

[1] Former affiliation.

0304-3975/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.tcs.2012.10.038

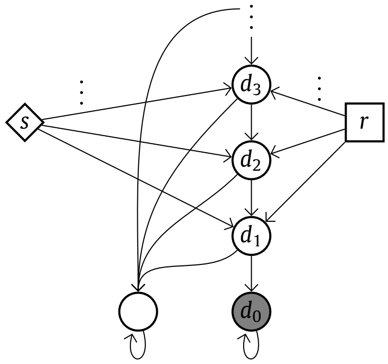

Fig. 1. A strongly determined game, where Min has no optimal strategy. All transition probabilities from stochastic states are uniformly distributed.

states they own, using *strategies*, which are functions from the possible histories of the play so far to the set of transitions from the current state. This induces a probabilistic measure for Borel-measurable sets of runs in a natural way.

A winning condition is a set W of runs. A run from W is won by Player Max, the other runs are won by Player Min (the games are zero-sum). For Borel measurable sets W, a fixed pair (σ, π) of strategies for Player Max and Min, respectively, and an initial state, s, the probability that Max wins is denoted by $\mathbb{P}_s^{\sigma,\pi}[W]$. The *value* of the game in s, denoted by $Val(s)$, is defined as

$$Val(s) := \sup_\sigma \inf_\pi \mathbb{P}_s^{\sigma,\pi}[W] = \inf_\pi \sup_\sigma \mathbb{P}_s^{\sigma,\pi}[W] \,. \tag{1}$$

The above equality, a consequence of a more general, Blackwell-determinacy result of Martin [1], implies that for every $\varepsilon > 0$ both of the players have so called ε-*optimal* strategies, σ_ε and π_ε, such that $\inf_\pi \mathbb{P}_s^{\sigma_\varepsilon,\pi}[W] \geq Val(s) - \varepsilon$, and $\sup_\sigma \mathbb{P}_s^{\sigma,\pi_\varepsilon}[W] \leq Val(s) + \varepsilon$. This may not be true for the case when $\varepsilon = 0$, where the so-called *optimal* (i.e., 0-optimal) strategies may not exist for either of the players.

We consider a stronger notion of determinacy than (1), and call a game *strongly determined* if for every state s, every p, $0 \leq p \leq 1$, and $\rhd \in \{>, \geq\}$ either Player Max has a strategy $\bar{\sigma}$ such that $\forall \pi : \mathbb{P}_s^{\bar{\sigma},\pi}[W] \rhd p$, or Player Min has a strategy $\bar{\pi}$ such that $\forall \sigma : \mathbb{P}_s^{\sigma,\bar{\pi}}[W] \not\rhd p$. This notion is a natural requirement, along similar lines as (1): (1) asserts that for every threshold p each player either can, or cannot ensure the probability of winning to be arbitrarily close to p, independently of the strategy chosen by the other player. Strong determinacy, on the other hand, asserts that for every threshold p each player either can, or cannot ensure the probability of winning to be at least p, independently of the strategy chosen by the other player.

In games with finite game graphs and reachability winning conditions these two notions coincide, and so each question about "winning" for finite-graph games can be freely interpreted in terms of the strong determinacy, or the determinacy as in (1). When extending the algorithmic results from finite-graph games to the general case (see, e.g., [2–5]), we may have to face two possible ways of how to ask the questions, because, as shown below, these notions of determinacy may differ.

Calling the second type of determinacy "strong" is deliberate, to indicate that strong determinacy implies the determinacy of (1). Indeed, writing $L := \sup_\sigma \inf_\pi \mathbb{P}_s^{\sigma,\pi}[W]$ and $R := \inf_\pi \sup_\sigma \mathbb{P}_s^{\sigma,\pi}[W]$, if Max has a strategy $\bar{\sigma}$ such that $\forall \pi : \mathbb{P}_s^{\bar{\sigma},\pi}[W] \geq p$ then $p \leq L$. Similarly, if Player Min has a strategy $\bar{\pi}$ such that $\forall \sigma : \mathbb{P}_s^{\sigma,\bar{\pi}}[W] \leq p$ then $p \geq R$. $L \leq R$ follows from their definitions, and by strong determinacy, $\forall p : \neg(R > p > L)$, thus $R \leq L$, and strong determinacy implies determinacy.

On the other hand, it is easy to see that the existence of ε-optimal strategies for both players implies strong determinacy for cases where $|p - Val(s)| \geq 2\varepsilon$, the players simply use their ε-optimal strategies to win. This works even for $\varepsilon = 0$, thus whenever both players have optimal strategies, the game is strongly determined (for all p).

Both players have optimal strategies.	$\subsetneq$	strongly determined	$\subsetneq$	Borel games

The picture sums up the relations between the key three notions. Both example in Fig. 1, and Example 1 show that the left inclusion is proper, whereas the example from [4, Fig. 1] shows that the right one is proper. More precisely, Fig. 1 shows a game graph where Max owns the state r, Min owns s, and the rest are stochastic states. The winning condition is to reach $T = \{d_0\}$ (shaded). Each stochastic state d_i, $i \geq 1$ has two outgoing transitions with equal probabilities (0.5): to d_{i-1} and to an absorbing state outside T. Thus, the value of d_i, $i \geq 0$ is 2^{-i}. Consequently the value of r, which has transitions to all d_i, $i \geq 1$ is $\sup_{i\geq 0} 2^{-i} = 1/2$, and the value of s which has the same set of successors is $\inf_{i\geq 0} 2^{-i} = 0$. However, unlike Max who simply can choose $r \to d_1$ to ensure winning with probability at least 1/2 from r, Min has no strategy to visit T with probability 0 from s. However, the game is strongly determined, because starting from every state, at least one of the players has effectively only one strategy to play, e.g., starting from r, no states of Min are ever visited, etc.

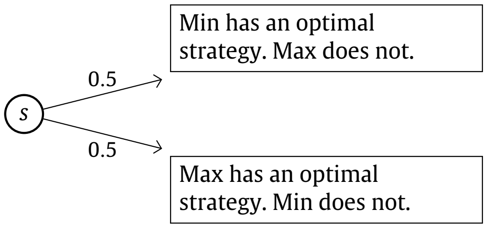

Fig. 2. A schematic representation of the game from [4, Fig. 1].

In the game from Example 1, which we show later, Player Min has only one (trivial) strategy, thus the game is strongly determined. However, there is a state r_0, such that for every fixed strategy of Max the probability of winning is strictly below $Val(r_0)$, in particular, Max does not always have an optimal strategy in that game.

The game from [4, Fig. 1], schematically represented here in Fig. 2, is composed of two halves, one of which is essentially equivalent to the game in Example 1, and the other is a similar game, where the roles of Max and Min are switched (infinite branching needed). As a consequence, in the first half, Player Max has no optimal strategy, whereas Player Min has a unique strategy there, which is therefore optimal; in the second half the situation is reversed. If we fix a strategy for Max, it plays optimally in the second half, but only achieves the optimal value minus some $\varepsilon > 0$ in the first half. Then Min can choose an $\varepsilon/2$-optimal strategy for the second half. Starting in state s, the balance of Min is at least $\varepsilon - \varepsilon/2 = \varepsilon/2 > 0$. But if we switch the order of fixing strategies, Max is the one with a positive balance. As a consequence, no player can ensure winning with probability at least $1/2$. In particular, this game is not strongly determined.

1.2. Our results

We are especially interested in the situation where W is an open set, and call such games *open* as well. This includes all *reachability* conditions, where W is the set of all runs visiting a state from a distinguished set of target states, T.

As noted before, the existence of optimal strategies for *both* players implies strong determinacy. But even for finitely-branching reachability games strong determinacy still holds, although Player Max may not have optimal strategies, and only Player Min always does [4]. Interestingly, we show here that this is a rule – one player having optimal strategies is enough to guarantee strong determinacy.

Theorem 1. *Let $\mathcal{G}$ be an open or closed stochastic game where one of the players has a strategy optimal in each state. Then $\mathcal{G}$ is strongly determined.*

For reachability, results of [4,6] imply (see Corollary 2) that Player Min has always optimal strategies if every state, s, owned by Min has at least one successor, t, such that $Val(s) = Val(t)$. This is always the case in finitely-branching games, where all states have only finite number of successors. On the other hand, even in very simple reachability games where every state has at most 2 successors, Player Max may not have an optimal strategy (cf. Example 1).

To complement the above results, and give more applications to Theorem 1, we give here a condition sufficient for the existence of optimal strategies for Player Max.

Theorem 2. *Let $\mathcal{G}$ be an open stochastic game. Player Max has an optimal strategy in all states, if*

$$\text{the set } V_\varepsilon := \{Val(s) \mid s \text{ a state of } \mathcal{G} \wedge Val(s) \geq \varepsilon\} \text{ is finite for all } \varepsilon > 0. \qquad (*)$$

In particular, $\mathcal{G}$ is not assumed to be finitely-branching. Condition $(*)$ is just saying that the set $V := \{Val(s) \mid s \text{ is a state}\}$ has no accumulation points, or the only such point is 0. To illustrate this on an example, for the game from Fig. 1: $V = \{0\} \cup \{2^{-i} \mid i \geq 0\}$. In this set, only 0 is an accumulation point, so Max has an optimal strategy in that game. The example is a toy one, and we could easily identify the optimal strategy here without Theorem 2, but it illustrates the concept.

It is a trivial task to construct a game where none of the players owns a single state, i.e., a Markov chain, and where the set V contains other accumulation points than 0. In Markov chains, however, each player has only one, trivial, strategy, which must thus be the optimal one. This shows that $(*)$ is not necessary. However, there are at least two reasons for which $(*)$ is interesting: first, we identify a class of recently studied infinite-state stochastic games which satisfy the assumption of Theorem 2, and for which the existence of optimal strategies for Max was not known before. This class, properly described later, consists of games generated by One-Counter automata [5,3,2], which satisfy a certain additional property, that can be tested algorithmically. As a special case, this class involves a maximising variant of Solvency Games [7].

Second, in Examples 1 and 2, we show games where Player Max lacks optimal strategies. These games are rather simple, and violate $(*)$ only "very slightly", in particular, they (i) are finitely-branching, and in fact have both the out-degree and in-degree of the game graph bounded by 2, (ii) do not contain states of Player Min at all, (iii) all transition probabilities in stochastic states are uniformly distributed, and (iv) V has only one accumulation point. This point is 1 in Example 1, and 1/2 in Example 2. In the latter case, the accumulation point is approached only from above, and $V \cap [0, 1/2) = \{0\}$. Thus it is not possible to weaken the assumption $(*)$ in Theorem 2 by allowing other accumulation points than 0.

1.3. Related work and open questions

This is a significantly extended version of a previous workshop publication [8]. Theorem 1, one of the two main results here, is new, and has not been included in the workshop version. It superseded a weaker result on strong determinacy [8, Theorem 2]. The proof techniques used to prove Theorem 1 are new and do not extend any proofs presented in [8]. Although the structure of the proof of Proposition 1 is borrowed from the proof of [4, Theorem 3.3], the underlying methods are completely different, to cope with infinite branching, as the proof of [4, Theorem 3.3] relied on several places on the fact that only finite branching is present.

Blackwell games are more general than our stochastic games, players there choose their moves simultaneously, not knowing the concurrent choice of the opponent. A famous determinacy result in the sense of (1) for Blackwell games is given by Martin [1].

Finitely-branching reachability games have been studied as a theoretical background for some algorithmic results concerning so called "BPA" games (i.e., games with graphs generated by stateless pushdown automata) in [4,6].

In 1992, Condon published a seminal work on finite-state reachability stochastic games [9], studying the complexity of their analysis, and stating a famous, and still open, problem of whether we can find out in polynomial time whether the value of a finite-state stochastic reachability game is at least $1/2$. In regard to the existence of optimal strategies and strong determinacy, finite-state games are not interesting, however: optimal strategies always exist there.

Theorem 1 sheds light on which open games are determined, and proves Conjecture 1 from the workshop version of this paper [8] for open games. The conjecture in general, however, stays unanswered, as far as we know, so we do not currently know whether Theorem 1 could be generalised for higher levels of the Borel hierarchy of objectives. Other open questions include finding new interesting classes of games where one of the players is guaranteed to have optimal strategies.

1.4. Outline of the paper

We briefly formalise the necessary notions, and recall some important known facts in Section 2. Section 3 proves Theorem 1, divided into two parts according to the player who has optimal strategies. This section is new and has not been included in [8] in any form. In Section 4 we prove Theorem 2, first for the special case of games without Player Min, and then in full generality. Finally, in Section 5 we briefly explain what are One Counter games, and apply our results to them.

2. Preliminaries

Definition 1. A *game graph*, $G = (S, \to, \delta)$, has a countable set S of *states*, partitioned into sets S_0, S_1, S_2 of stochastic states, states of Player Max, and Player Min, respectively; a countable *transition relation* $\to\ \subseteq S \times S$ such that $\forall r \in S : \exists s \in S : r \to s$; and a probability weight function $\delta : S_0 \to (S \to [0, 1])$ such that for all $r \in S_0$ we have $\forall r \to s : \delta(r)(s) > 0$ and $\sum_{r\to s} \delta(r)(s) = 1$.

A *run* is an infinite path in a game graph. For a finite path w, also called a *history*, we denote the states it visits by $w(0), w(1), \ldots, w(k)$, and call $k = len(w)$ the *length* of w. $Run(w)$ is the set of all runs extending w. Unions of sets of the form $Run(w)$ are called open sets, they are open in the product topology over the discrete spaces S. Closing the set of open sets under complement and countable union defines the set of (Borel-)measurable sets.

Definition 2. A *game*,[2] $\mathcal{G}$, is given by a game graph, G, and a Borel-measurable set of runs, W, called the *winning condition*. If there is some $T \subseteq S$ so that $W = \bigcup\{Run(w) \mid w \text{ ends in } T\}$ then W is a *reachability* condition, $\mathcal{G}$ is called a *reachability game*, and T is the *target set* for W.

A *strategy* for Player Max is a function assigning to every history ending in a state $s \in S_1$ a distribution over the successors of s. Similarly, a strategy for Min is defined for histories ending in S_2. A strategy is *memoryless*, if it only depends on the last state of the history.

Fixing a pair of strategies, (σ, π), for Max and Min, respectively, we assign to every finite path, w, the product, $\rho^{\sigma,\pi}(w)$, of weights on the edges along w given by δ, σ, and π. Fixing also an initial state, s, we define a probability measure $\mathbb{P}_s^{\sigma,\pi}[\cdot]$ by $\mathbb{P}_s^{\sigma,\pi}[Run(w)] := 0$ for w not starting in s, $\mathbb{P}_s^{\sigma,\pi}[Run(w)] := \rho^{\sigma,\pi}(w)$ for w starting in s, and extending this to complement and union to satisfy the axioms of a probability measure. The uniqueness of this construction is a standard fact; see, e.g., [10, p. 30].

The definition of the *value*, $Val(\cdot)$, has been given in (1). For $\varepsilon \geq 0$, a strategy σ for Max is ε-*optimal* in a state s if $\mathbb{P}_s^{\sigma,\pi}[W] \geq Val(s) - \varepsilon$ for all strategies, π, for Min. The ε-optimal strategies for Min are defined analogously. We call 0-optimal strategies just *optimal*.

Remark 1. If we say that a strategy is optimal, omitting specifying in which state, then we mean: optimal in every state. For a reachability winning condition, if for every state there is a strategy optimal in it, then there is one strategy optimal in every state, because the behaviour of an optimal strategy between the initial state and the target set is optimal as well. Note that this is, in general, false both for ε-optimal strategies, and for more complicated objectives.

2 As noted in the introduction, we use the simple term "games" for our special kind of games, because we do not speak about other games here.

2.1. Technical assumptions

The following assumption states a property of a game $\mathcal{G}$. Whenever we will use it, we will explicitly refer to it.

Assumption 1. The game graph of the game $\mathcal{G}$ is a forest.

The following lemma essentially says that for all we do with games in this paper, Assumption 1 is without loss of generality. This is done by simulating $\mathcal{G}$ by a new game $\mathcal{G}'$, which is derived from $\mathcal{G}$ by encoding the history in the states. It is shown that all interesting properties of $\mathcal{G}$ are preserved, while the game tree of $\mathcal{G}'$ forms a forest.

Lemma 1. *For every game $\mathcal{G} = ((S, \rightarrow, \delta), W)$ with W open, there is a game $\mathcal{G}' = ((S', \hookrightarrow, \delta'), W')$ satisfying Assumption 1, and functions ϕ from paths in $G' = (S', \hookrightarrow, \delta')$ to paths in $G = (S, \rightarrow, \delta)$, and Φ from strategies in $\mathcal{G}$ to memoryless strategies in $\mathcal{G}'$ such that*

1. *G' is a forest.*
2. *$S \subseteq S'$.*
3. *For each $s \in S$, the restriction of ϕ to runs of G' starting in s is a probability preserving bijection to the runs of G starting in s.*
4. *For all measurable $A \subseteq Run(s)$ in $\mathcal{G}$, and all pairs (σ, π) of strategies in $\mathcal{G}$: $\mathbb{P}_s^{\sigma,\pi}[A] = \mathbb{P}_s^{\Phi(\sigma),\Phi(\pi)}\left[\phi^{-1}(A)\right]$.*
5. *For each strategy σ' of Max in $\mathcal{G}'$ there is a strategy σ of Max in $\mathcal{G}$ such that $\Phi(\sigma)(w) = \sigma'(w)$ for all histories w in $\mathcal{G}'$. This is also true for strategies of Min.*
6. *$W' = \phi^{-1}(W)$ is a reachability objective, and its target set T can be chosen closed on outgoing transitions.*
7. *$Val(s)$ is the same in $\mathcal{G}$ and $\mathcal{G}'$ for all $s \in S$, and the sets of all values in $\mathcal{G}$ and in $\mathcal{G}'$ are equal.*
8. *For each $\varepsilon \geq 0$, a strategy σ for Max is ε-optimal in $\mathcal{G}$ iff $\Phi(\sigma)$ is ε-optimal in $\mathcal{G}'$. This is true also for optimal strategies, and also for Player Min.*

Proof. We will construct game $\mathcal{G}'$ from $\mathcal{G}$ simply by keeping track of the history inside the states. More precisely, given a game $\mathcal{G} = (G, W)$, $G = (S, \rightarrow, \delta)$, consider a game $\mathcal{G}' = (G', W')$, $G' = (S', \hookrightarrow, \delta')$, where the states in S' are just histories from $\mathcal{G}$. We define $\hookrightarrow$ to be the smallest set satisfying: whenever $r \rightarrow s$ in $\mathcal{G}$ then $wr \hookrightarrow wrs$ in $\mathcal{G}'$. Clearly, both 1 and 2 are true. Projecting the states of S' to their last component induces a map ϕ from paths in G' to paths in G. We also define a map Φ from strategies in $\mathcal{G}$ to memoryless strategies in $\mathcal{G}'$, by setting for all histories w in $\mathcal{G}$: $\Phi(\sigma)(w) = \sigma(w)$. The partition of S', and the weight function δ' are both derived from S and δ by projecting states from S' to the last component.

It is easy to verify straight from the definitions that both 3 and 4 are true. Because no state in $\mathcal{G}'$ can be reached from another state on more than two paths, we obtain 5. The target set T for W' is simply the set of all histories w from $\mathcal{G}$ such that $Run(w) \subseteq W$. Clearly, if $w \in T$ then all its extensions (i.e., successors in G') are in T, proving 6. Finally, 7 is a consequence of 3, 4, 5, and the definition of W'; and 8 is a consequence of 7 and 4. $\square$

2.2. Known results for reachability games

We state here some known results to be used later. The following gives a characterisation of values, and allows us to characterise the existence of optimal strategies for Min.

Fact 1 (*cf.* [4, Theorem 3.1]). *Let $\mathcal{G} = (G, W)$, $G = (S, \rightarrow, \delta)$ be a reachability game, with winning condition W and target set T. The least fixed point of the following (Bellman) functional $\mathcal{V} : (S \rightarrow [0, 1]) \rightarrow (S \rightarrow [0, 1])$ exists and is equal to $Val(\cdot)$.*

$$\mathcal{V}(f)(s) = \begin{cases} 1 & \text{if } s \in T \\ \sup\{f(r) \mid s \rightarrow r\} & \text{if } s \in S_1 \setminus T \\ \inf\{f(r) \mid s \rightarrow r\} & \text{if } s \in S_2 \setminus T \\ \sum_{s \rightarrow r} \delta(s)(r) \cdot f(r) & \text{if } s \in S_0 \setminus T. \end{cases} \tag{2}$$

Corollary 1. *Let $\mathcal{G}$ be a game as in Fact 1, satisfying $(*)$. Let $s \in S_1 \cup S_2$. There is some $s \rightarrow r$ such that $Val(s) = Val(r)$, unless $s \in S_2$ and $Val(s) = 0$.*

Proof. First assume that $s \in S_1$. If $Val(s) = 0$ then by (2) for every $s \rightarrow r$ we have $0 \leq Val(r) \leq Val(s) = 0$. Now assume $Val(s) > 0$, and for proof by contradiction, also assume that $Val(r) < Val(s)$ for all $s \rightarrow r$. By (2), $Val(s) = \sup_{s \rightarrow r} Val(r)$, thus for every $\varepsilon > 0$ there is some $s \rightarrow r$ such that $Val(s) - \varepsilon < Val(r) < Val(s)$. In other words, $Val(s) > 0$ is an accumulation point, a contradiction with $(*)$. Thus, there must always be some $s \rightarrow r$ such that $Val(r) = Val(s)$. The proof for $s \in S_2$ and $Val(s) > 0$ is analogous. $\square$

Corollary 2 (*cf.* [4, Theorem 3.1]). *Let $\mathcal{G}$ be a game as in Fact 1. Let $G' = (S, \hookrightarrow, \delta)$ be a subgraph of G where $\hookrightarrow$ is a subset of $\rightarrow$, and if there is a pair $r, s \in S$ such that $r \rightarrow s$ and $r \not\hookrightarrow s$ then $r \in S_2$ and there is some $s' \in S$ such that $r \hookrightarrow s'$ and $Val(s') \leq Val(s)$ in $\mathcal{G}$. Let $\mathcal{G}' = (G', W)$. Then the values are the same in $\mathcal{G}$ and $\mathcal{G}'$.*

As a consequence, a strategy, π, for Min is optimal iff for all $r \in S_2$ it chooses with positive probability only successors $s \in S$ satisfying $Val(r) = Val(s)$.

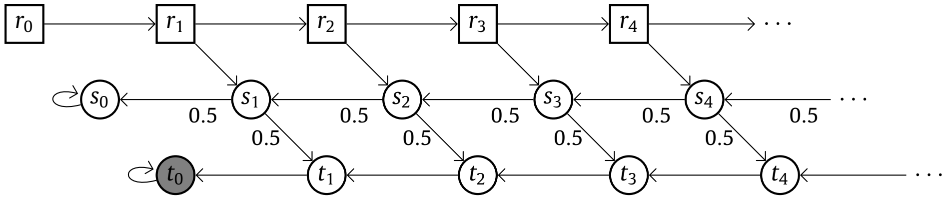

Fig. 3. A reachability game where Player Max (□ states) has no optimal strategy.

Proof. Let $\mathcal{V}'$ be the Bellman functional associated with $\mathcal{G}'$. Observe that the values in $\mathcal{G}$ form a fixed point of $\mathcal{V}'$, thus for all $s \in S$, $Val(s)$ in $\mathcal{G}'$ is equal to or less than $Val(s)$ in $\mathcal{G}$. Moreover, it cannot be less, because Player Max has the same set of strategies in $\mathcal{G}'$ as in $\mathcal{G}$, whereas Player Min does not get more strategies in $\mathcal{G}'$. For the consequence, remove edges leaving S_2 unused by π. □

Note that the situation is not symmetric for Player Max. Consider games without Player Min, and with out-degree and in-degree bounded by 2. In particular, this implies that every state, r, of Player Max has at least one successor, s, with $Val(r) = Val(s)$. Even in these games, Player Max may lack optimal strategies, as illustrated in the following known example (see, e.g., [2, p. 871], [11, Example 6]).

Example 1. Consider the reachability game from Fig. 3. Its game graph, G, has the set $S := \{r_i, s_i, t_i \mid i \geq 0\}$ of states, partitioned by $S_0 = \{s_i, t_i \mid i \geq 0\}$, $S_1 = \{r_i \mid i \geq 0\}$, and $S_2 = \emptyset$. Transitions are $s_0 \to s_0$, $t_0 \to t_0$, and $r_{i-1} \to r_i$, $r_i \to s_i$, $s_i \to s_{i-1}$, $s_i \to t_i$, and $t_i \to t_{i-1}$ for $i > 0$. Probabilities are always uniform. The target set is $T = \{t_0\}$. Clearly, $Val(s_i) = 1 - 2^{-i}$ for all $i \geq 0$. Thus $Val(r_i) = 1$ for all $i \geq 0$: for every $N > 0$, choosing the transition $r_i \to r_{i+1}$ for $i < N$, and the transition $r_i \to s_i$ for $i \geq N$, is a 2^{-N}-optimal strategy for Max. Yet Max has no optimal strategy in any r_i, $i \geq 0$: no strategy reaching some s_j is optimal, and, on the other hand, never reaching s_j means never reaching t_0.

3. Strong determinacy

In this section we prove Theorem 1. Note that a game with a closed winning condition becomes a game with an open winning condition if the roles of the players are swapped. To prove the theorem it is thus enough to assume that $\mathcal{G}$ is open. Moreover, due to Lemma 1 we may safely assume that $\mathcal{G}$ satisfies Assumption 1.

We divide Theorem 1 into two parts: Propositions 1 and 2, each of which reflects the situation where Player Min, or Player Max, respectively, is guaranteed to have an optimal strategy in each state of the game. This section starts with some auxiliary results in Section 3.1, and continues with stating and proving Propositions 1 and 2 in separate Sections 3.2 and 3.3.

3.1. Auxiliary lemmata

Lemma 2. *Let S be a countable set. There is an injection,* Rank, *of all finite sequences of elements of S into natural numbers, $\mathbb{N}$, so that if a sequence w is a prefix of a sequence w', then* $\mathrm{Rank}(w) \leq \mathrm{Rank}(w')$.

Proof. Clearly there is a bijection $f : S \to \mathbb{N}$, because S is countable. Let $p_0, p_1, p_2 \ldots$ be the sequence of all primes. Let $w = w(0) \cdots w(k)$ be a history of length k. We set $\mathrm{Rank}(w) := \prod_{j=0}^{k} p_j^{f(w(j))}$, which is an injection due to the unique prime decomposition of natural numbers. Observe that if w is a prefix of w' then $\mathrm{Rank}(w)$ divides, and therefore is less than or equals, $\mathrm{Rank}(w')$. □

The following property of games, named Assumption 2, is used in Lemmas 3–5.

Assumption 2. $\mathcal{G} = (G, W)$ is a reachability game satisfying Assumption 1. The graph G is a tree, rooted in $\hat{r} \in S$ and satisfying

$$\forall s \in S_1 : \forall s \to t : Val(s) = Val(t). \tag{3}$$

Each time a game $\mathcal{G}$ satisfies Assumption 2, and thus also Assumption 1, we only need to consider memoryless strategies (Lemma 1, part 5). This greatly simplifies the notation for conditional probability. Indeed, consider a path w, from state s to r, and a reachability condition W. Instead of writing $\mathbb{P}_s^{\sigma,\pi}[W \mid Run(w)]$ and worrying whether $\mathbb{P}_s^{\sigma,\pi}[Run(w)] > 0$, we can just measure $\mathbb{P}_r^{\sigma,\pi}[W]$. Because both strategies are memoryless, their behaviour will not change when the history w gets reduced to r.

Lemma 3. *Let $\mathcal{G}$ satisfy Assumption 2. Let $\delta > 0$ be a fixed number. A strategy σ for Max is optimal if for all strategies π of Min and $s \in S$:*

$$\mathbb{P}_s^{\sigma,\pi}[W] \geq \delta \cdot Val(s). \tag{4}$$

Proof. We set $\delta' := \delta/2 > 0$. Observe that if σ has the following property then it clearly is optimal:

$$\forall \pi : \forall s \in S : \forall i \geq 0 : \mathbb{P}_s^{\sigma,\pi}[W] \geq (1 - (1 - \delta')^i) \cdot Val(s). \tag{5}$$

In the rest of the proof we will prove (5) by induction on $i \geq 0$. The base case, $j = 0$ is trivial, because every probability is non-negative. Now assume that (5) is true for $j = k$. We will prove it for $j = k + 1$, an arbitrary but fixed $s \in S$, and an arbitrary fixed strategy π of Min.

Recall that W is a reachability condition, consisting exactly of paths visiting states from some target set, $T \subseteq S$. For $n \geq 0$ let W_n denote the restriction of W to paths reaching T within n steps. Because $\mathbb{P}_s^{\sigma,\pi}[W] = \lim_{n\to\infty} \mathbb{P}_s^{\sigma,\pi}[W_n]$, we know that there is some $n \geq 0$ such that

$$\mathbb{P}_s^{\sigma,\pi}[W_n] \geq \frac{\mathbb{P}_s^{\sigma,\pi}[W]}{2} \geq \frac{\delta \cdot Val(s)}{2} = \delta' \cdot Val(s). \tag{6}$$

The second inequality follows from (4).

Let U be the set of all histories of length n starting in s. Let $U_{\hat{T}} \subseteq U$ be the subset of all $u \in U$ which *do not* contain a state from T. Clearly,

$$\mathbb{P}_s^{\sigma,\pi}[W] = \mathbb{P}_s^{\sigma,\pi}[W_n] + \sum_{u\in U_{\hat{T}}} \mathbb{P}_s^{\sigma,\pi}[W \cap Run(u)]\,. \tag{7}$$

In the following equation we write r_u for the last state of any history u. Because all edges leaving states of player Max preserve the value, we know that there is some $\xi \geq 0$ such that

$$Val(s) = \mathbb{P}_s^{\sigma,\pi}[W_n] + \sum_{u\in U_{\hat{T}}} Val(r_u) \cdot \mathbb{P}_s^{\sigma,\pi}[Run(u)] - \xi\,. \tag{8}$$

By the inductive hypothesis (5) for $j = k$ and each $u \in U_{\hat{T}}$:

$$\begin{aligned} \mathbb{P}_s^{\sigma,\pi}[W \cap Run(u)] &= \mathbb{P}_{r_u}^{\sigma,\pi}[W] \cdot \mathbb{P}_s^{\sigma,\pi}[Run(u)] \\ &\geq (1 - (1 - \delta')^k) \cdot Val(r_u) \cdot \mathbb{P}_s^{\sigma,\pi}[Run(u)]\,. \end{aligned} \tag{9}$$

Finally, we put together the proof of (5) for $j = k + 1$, s and π:

$$\begin{aligned} Val(s) - \mathbb{P}_s^{\sigma,\pi}[W] &= \left(\sum_{u\in U_{\hat{T}}} \left(Val(r_u) \cdot \mathbb{P}_s^{\sigma,\pi}[Run(u)] - \mathbb{P}_s^{\sigma,\pi}[W \cap Run(u)]\right)\right) - \xi \text{ by (7) and (8)} \\ &\leq (1-\delta')^k \cdot \left(\sum_{u\in U_{\hat{T}}} Val(r_u) \cdot \mathbb{P}_s^{\sigma,\pi}[Run(u)]\right) - \xi \text{ by (9)} \\ &\leq (1-\delta')^k \cdot \left(\sum_{u\in U_{\hat{T}}} Val(r_u) \cdot \mathbb{P}_s^{\sigma,\pi}[Run(u)] - \xi\right) \text{ because } \xi \geq 0 \\ &= (1-\delta')^k \cdot \left(Val(s) - \mathbb{P}_s^{\sigma,\pi}[W_n]\right) \text{ by (8)} \\ &\leq (1-\delta')^{k+1} \cdot Val(s) \text{ by (6)}. \end{aligned}$$

The proof of the lemma is finished. □

Definition 3 (*Notation*)**.** Let $\mathcal{G}$ be a game satisfying Assumption 2. For each pair $s, r \in S$ of states, there is at most one path from s to r in the graph G. If there is one, we use $w_{s,r}$ to denote it, and we call s an *ancestor* of r, and r a *descendant* of s. The set of all ancestors of $s \in S$, including $\hat{r}$ and s, is called Pre_s. A Rank of a state, s, is defined as the Rank of the path $w_{\hat{r},s}$, which always exists, $\text{Rank}(s) := \text{Rank}(w_{\hat{r},s})$, where Rank is the function from Lemma 2.

In the following Lemmas 4 and 5, we will need the notion of a reset sequence. To give some intuition: we already know from Lemma 3 that if a strategy for Max is not optimal, it must start to significantly underperform in at least one state, delivering less than 1/3 of the attainable value in that state. We improve such a strategy by "resetting" it after that state is visited, to behave like some 2/3-optimal strategy. A reset sequence represents a chain of such resets performed on some strategy for Max.

Definition 4 (*Reset Sequence*)**.** Let $\mathcal{G}$ be a game satisfying Assumption 2. The following defines a property of a state $s \in S$ and a strategy σ for Max:

$$\Phi(s,\sigma) \quad \stackrel{\text{def}}{\equiv} \quad \exists \pi : \mathbb{P}_s^{\sigma,\pi}[W] < \frac{1}{3} \cdot Val(s).$$

A *reset sequence* is a sequence $\{(\sigma_i, s_i, U_i)\}_{i=0}^{\infty}$ of triples of strategies σ_i of Max, states s_i, and sets U_i of histories. We call σ_i *reset strategies*, and s_i *(the i-th) reset*. The sequence is defined inductively.

The root of G is the 0-th reset, $s_0 = \hat{r}$, σ_0 is some strategy $\frac{2}{3}$-optimal in s_0.

For $i \geq 0$, if σ_i is optimal in $\hat{r}$, then $\sigma_{i+1} = \sigma_i$ and $s_{i+1} = s_i$. Otherwise, by Lemma 3 there is at least one state s which satisfies $\Phi(s, \sigma_i)$. We set s_{i+1} to be the Rank-minimal among all such states. There is a strategy σ of Max which is $\frac{2}{3}$-optimal in s_{i+1}. We define σ_{i+1} to behave like σ after s_{i+1} is reached, and like σ_i in all histories w' which do not contain s_{i+1}.

Let $w = w(0)w(1)\cdots w(n)$ be a path starting in s_i. We call it $i{+}1$-*reset free* if for all k, $0 \leq k \leq n$: $\Phi(w(k), \sigma_i)$ is false. Recall that W is a reachability condition, of the form $W = \bigcup_w Run(w)$ where w ranges over all paths ending in a target set $T \subseteq S$. We define U_i to be the set of all paths from s_i ending in T which are also $i{+}1$-reset free.

Note that a reset sequence has a game $\mathcal{G}$ implicitly as its parameter, i.e., not reflected in the notation, but used in the definition. However, in all places where we use it, $\mathcal{G}$ is clear from the context.

Lemma 4. *Let $\mathcal{G}$ satisfy Assumption 2, and let $\{(\sigma_i, s_i, U_i)\}_{i=0}^{\infty}$ be one of its reset sequences. The following is true for all $i \geq 0$:*

1. *If $Val(s_0) = 0$ then $s_0 = s_i$ and $\sigma_0 = \sigma_i$. If $Val(s_0) > 0$ then $Val(s_i) > 0$.*
2. $\forall \pi : \mathbb{P}_{s_i}^{\sigma_i,\pi}[W] \geq \frac{2}{3} \cdot Val(s_i)$.
3. *For all $s \in S$ which are not descendants of s_{i+1}, and for all π of Min:*
$$\mathbb{P}_{\hat{r}}^{\sigma_{i+1},\pi}\left[Run(w_{\hat{r},s})\right] = \mathbb{P}_{\hat{r}}^{\sigma_i,\pi}\left[Run(w_{\hat{r},s})\right] \;\wedge\; \mathbb{P}_{s}^{\sigma_{i+1},\pi}[W] \geq \mathbb{P}_{s}^{\sigma_i,\pi}[W]\,. \tag{10}$$
4. $\forall s \in S : [(\forall r \in Pre_s : \neg\Phi(r, \sigma_i)) \implies (\forall j \geq i : \forall r \in Pre_s : \neg\Phi(r, \sigma_j))]$.
5. *For all $K \geq i$ and $j \leq K$, if for no k, $0 \leq k < K$, σ_k is optimal in $\hat{r}$ then:*
   - *If $s_i \in Pre_{s_j}$ then $i \leq j$.*
   - *If $i \neq j$ then $s_i \neq s_j$.*
6. *If σ_k is not optimal in $\hat{r}$ for any $k \geq 0$ then:*
   - *For all $j \geq 0$, if $s_i \in Pre_{s_j}$ then $i \leq j$.*
   - *For all $j \geq 0$, if $i \neq j$ then $s_i \neq s_j$.*
   - $\forall s \in S : \exists J \geq 0 : \forall j > J : s_j \notin Pre_s$.
7. $\forall s \in S : \forall j > \mathrm{Rank}(s) : \neg\Phi(s, \sigma_j)$.
8. $\forall s \in S : (\Phi(s, \sigma_i) \implies \exists j > i : s_j \in Pre_s)$.
9. *Let s be the last state of some $u \in U_i$. Then:*
$$\forall r \in Pre_s : \forall j \geq i : \forall \pi : \mathbb{P}_{\hat{r}}^{\sigma_i,\pi}\left[Run(w_{\hat{r},r})\right] = \mathbb{P}_{\hat{r}}^{\sigma_j,\pi}\left[Run(w_{\hat{r},r})\right]. \tag{11}$$
10. $\forall \pi : \mathbb{P}_{s_i}^{\sigma_i,\pi}\left[\bigcup_{w \in U_i} Run(w)\right] > \frac{1}{3} \cdot Val(s_i)$.

Proof. If $Val(s_0) = 0$ then σ_0 is optimal, and the first half of 1 follows directly from the definition of a reset sequence. Furthermore, $Val(s_0) > 0 \implies Val(s_i) > 0$ is an easy inductive consequence of $\forall j \geq 0 : Val(s_j) > 0 \implies Val(s_{j+1}) > 0$. The latter is true, because either there is no $s \in S$ such that $\Phi(s, \sigma_j)$ is true, but then σ_j is optimal in s_j and thus $s_{j+1} = s_j$, or there is such s, and in particular $\Phi(s_{j+1}, \sigma_j)$ is true. Observe however, that by definition of Φ, $\Phi(s_{j+1}, \sigma_j)$ can only be true if $Val(s_{j+1}) > 0$. Because the rest of the lemma is immediate to check in the case when $Val(s_0) = 0$ and $s_0 = s_i$ for all $i \geq 0$, we assume until the end of this proof that $Val(s_i) > 0$ for all $i \geq 0$.

Item 2 follows straight from the definition of the reset strategies. The equality in (10) is true because σ_{i+1} behaves by definition identically to σ_i until hitting s_{i+1}. The inequality in (10) follows from the equality, from item 2, and from $\Phi(s_{i+1}, \sigma_i)$. Let us prove 4 now. First we prove

$$\forall s \in S : [(\forall r \in Pre_s : \neg\Phi(r, \sigma_i)) \implies \neg\Phi(s, \sigma_{i+1})]. \tag{12}$$

Assume that $\Phi(r, \sigma_i)$ is false for all $r \in Pre_s$. This guarantees that s_{i+1} is not an ancestor of s, because $\Phi(s_{i+1}, \sigma_i)$ is true by definition of s_{i+1}. This means that by (10), for all π of Min we have $\mathbb{P}_s^{\sigma_{i+1},\pi}[W] \geq \mathbb{P}_s^{\sigma_i,\pi}[W]$, implying $\Phi(s, \sigma_{i+1}) \implies \Phi(s, \sigma_i)$. Because $s \in Pre_s$ and $\forall r \in Pre_s : \neg\Phi(r, \sigma_i)$ we have that $\Phi(s, \sigma_{i+1})$ is false. Using (12), we immediately have

$$\forall s \in S : [(\forall r \in Pre_s : \neg\Phi(r, \sigma_i)) \implies (\forall r \in Pre_s : \neg\Phi(r, \sigma_{i+1}))]. \tag{13}$$

by applying (12) to each $r \in Pre_s$ in place of s. Finally, 4 follows from (13) by an easy induction on $j - i$.

For proving 5, fix some K such that $i \leq K$ and assume that for no k, $0 \leq k < K$, σ_k is optimal in $\hat{r}$. First we prove that if $s_k = s_0$ for some $k \leq K$ then $k = 0$, by contradiction. Assume that $s_k = s_0$ and $k > 0$. By 4, $\Phi(s_k, \sigma_\ell)$ is false for all $\ell \geq 0$. On the other hand, from σ_{k-1} not being optimal in $\hat{r}$ we know that, by definition of σ_k, $\Phi(s_k, \sigma_{k-1})$ is true, a contradiction. Thus if $s_k = s_0$ then $k = 0$. Now we fix also j such that $j \leq K$, and prove the first part of 5 for it. If $i = 0$ we clearly have $i \leq j$. If $j = 0$ and $s_i \in Pre_{s_j}$ then $s_i = s_0$. We already proved that then $i = 0$, so $i \leq j$ again. The remaining case is when $i, j > 0$: assuming that s_j is an ancestor of s_i we have to prove that $i \leq j$. By Rank-minimality of s_j among states s for which $\Phi(s, \sigma_{j-1})$ is true, and by the fact that Rank preserves the ancestor order (Lemma 2), $\Phi(s, \sigma_{j-1})$ is false for every ancestor $s \neq s_j$ of s_j. By 4, $\Phi(s, \sigma_k)$ is false for every ancestor $s \neq s_j$ of s_j and every $k \geq j - 1$. By 4 again, and by 2, $\Phi(s, \sigma_k)$ is false for every ancestor s of s_j, including s_j, and all $k \geq j$. In particular, because $\Phi(s_i, \sigma_{i-1})$ is true, we have $i - 1 < j$. In other words, $i \leq j$. The next part of 5 follows directly from the first part: obviously if $s_i = s_j$, they are ancestors of each other and thus $i \leq j$ and $j \leq i$, yielding $i = j$.

To prove part 6, first two items, choose $K \geq \max(i, j)$ and apply 5. For the third item, observe that every $s \in S$ has only finitely many ancestors. Because $s_i \neq s_j$ for $i \neq j$, there are only finitely many resets among the ancestors of s.

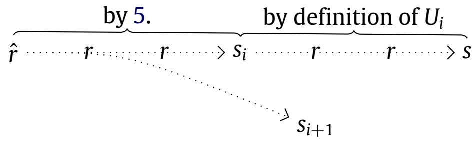

Fig. 4. Proving $\neg\Phi(r, \sigma_i)$ for ancestors r of s. Here s_{i+1} may also be a descendant of s_i.

We prove 7 by induction on $k := \text{Rank}(s) + 1 \geq 0$. For $k = 0$ the claim is immediate, because no state has negative Rank. Now assume $k = \ell + 1$ for some $\ell \geq 0$. If there is no s with $\text{Rank}(s) = \ell$ then it is enough to use the inductive hypothesis. Otherwise let us fix such s. Note that for each ancestor $r \neq s$ of s: $\text{Rank}(r) < \text{Rank}(s) = \ell$. Thus we may use the inductive hypothesis 7 for $k = \ell$ on all ancestors $r \neq s$ of s. Now $\Phi(s, \sigma_\ell)$ is either true or false. If false, then by 4 and the hypothesis for $k = \ell$, $\Phi(s, \sigma_j)$ is false for all $j \geq \ell$. If true, then $s = s_{\ell+1}$ by the hypothesis for $k = \ell$ and by construction of the resets, and thus $\Phi(s, \sigma_{\ell+1})$ is false. By 4 and the hypothesis, $\Phi(s, \sigma_j)$ is false for all $j \geq \ell + 1$.

We now show how 7 implies 8. If $\Phi(s, \sigma_i)$ is true then by 7, $\text{Rank}(s) \geq i > 0$, and $\Phi(s, \sigma_k)$ is false for all $k > \text{Rank}(s)$. This is only possible if for some j, $i < j \leq \text{Rank}(s) + 1$, and some ancestor r of s, $\sigma_j(r) \neq \sigma_{j-1}(r)$. Such r with a minimal Rank has become the j-th reset, $r = s_j$.

Now we prove 9. Let us fix s, the last state of some $u \in U_i$. If σ_i is optimal in $\hat{r}$ then $\sigma_j = \sigma_i$ for all $j \geq i$ and so 9 is obviously true. Otherwise we can use 5 for $K = i + 1$, and get that s_{i+1} is not an ancestor of s_i. Because s_{i+1} is the Rank-minimal state r for which $\Phi(r, \sigma_i)$ is true, and Rank preserves the ancestor order (Lemma 2), $\Phi(r, \sigma_i)$ is false for every ancestor r of s_i. By definition of U_i, $\Phi(r, \sigma_i)$ is also false for every r which is a descendant of s_i and an ancestor of s. Therefore $\Phi(r, \sigma_i)$ is false for every ancestor r of s. By 4, $s_j \neq r$ for all ancestors r of s and all $j \geq i$. See also Fig. 4 for a schema. An induction on $j - i$ using the first part of (10) finishes the proof of (11), as follows: for $j - i = 0$ (11) is obviously true. Let us assume that (11) is true for some $j = k$, $k \geq i$, we prove it for $j = k + 1$. We already proved that s_{k+1} is not an ancestor of any $r \in Pre_s$. Therefore we can use the equality in (10) with k substituted for i to derive (11) for $j = k + 1$.

Now we prove 10. Runs from s_i which belong to W but do not have a prefix in U_i have to hit a state s, before reaching the target set of W, such that $\Phi(s, \sigma_i)$ is true. This means that every π can be modified to π' so that

$$\mathbb{P}^{\sigma_i,\pi'}_{s_i}\left[W \smallsetminus \bigcup_{u \in U_i} Run(u)\right] < Val(s_i) \cdot \frac{1}{3}, \tag{14}$$

and still

$$\mathbb{P}^{\sigma_i,\pi}_{s_i}\left[\bigcup_{u \in U_i} Run(u)\right] = \mathbb{P}^{\sigma_i,\pi'}_{s_i}\left[\bigcup_{u \in U_i} Run(u)\right]. \tag{15}$$

The modification is as follows: π' behaves like π until hitting some s, which satisfies $\Phi(s, \sigma_i)$. After that point, by definition of Φ, σ_i is performing so badly that it only can assure less than $\frac{1}{3}$ of the value of s. Thus there is a strategy π_s such that indeed $\mathbb{P}^{\sigma_i,\pi_s}_{s}[W] < \frac{1}{3} \cdot Val(s)$. After reaching s, π' starts to behave like π_s. Now (14) follows from the definition of π' and π_s, and (15) from the definition of U_i, which ensures that π' never switches from π to some π_s on a path from U_i.

Let us fix an arbitrary π. We set $p := \mathbb{P}^{\sigma_i,\pi}_{s_i}\left[\bigcup_{u \in U_i} Run(u)\right]$. By (15) we know that $p = \mathbb{P}^{\sigma_i,\pi'}_{s_i}\left[\bigcup_{u \in U_i} Run(u)\right]$. Using 2 with π' substituted for π, we get that $p + \mathbb{P}^{\sigma_i,\pi'}_{s_i}\left[W \smallsetminus \bigcup_{u \in U_i} Run(u)\right] \geq \frac{2}{3} \cdot Val(s_i)$. Substituting (14) in the left part we obtain $p + \frac{1}{3} \cdot Val(s_i) > \frac{2}{3} \cdot Val(s_i)$, which translates to $p > \frac{1}{3} \cdot Val(s_i)$. Because π was chosen arbitrarily, the proof of 10 is finished. $\square$

Lemma 5. *Max has an optimal strategy in every $\mathcal{G}$ satisfying Assumption 2.*

Proof. Let $\{(\sigma_i, s_i, U_i)\}_{i=0}^{\infty}$ be some reset sequence of $\mathcal{G}$. If σ_i is optimal in s_0 for some $i \geq 0$ we are done. Assume that no σ_i is optimal in s_0. We first define a strategy, σ, and then we prove it is optimal. Let s be an arbitrary state. Lemma 4, part 6 gives us a function $J(s)$, which is the least number $J \geq 0$ such that no ancestor of s is the j-th reset for any $j > J$. As a consequence of the construction of reset strategies, this implies that $\sigma_j(s) = \sigma_{J(s)}(s)$ for all $j \geq J(s)$. We set $\sigma(s) := \sigma_{J(s)}(s)$.

Let s be an arbitrary state. In the rest of the proof we will show that

$$\inf_{\pi} \mathbb{P}^{\sigma,\pi}_{s}[W] \geq \frac{1}{3} \cdot Val(s). \tag{16}$$

By Lemma 3 with $\delta = \frac{1}{3}$, σ satisfying (16) is optimal.

Observe that $J(s_j) = j$ for each $j \geq 0$, because s_j is its own ancestor, and by Lemma 4, part 6, no s_k is an ancestor of s_j for $k > j$. From Lemma 4, part 9 we know that all σ_k, $k \geq j$ behave like σ_j on prefixes of runs from U_j. Thus for each π of Min: $\mathbb{P}^{\sigma,\pi}_{s_j}\left[\bigcup_{u \in U_j} Run(u)\right] = \mathbb{P}^{\sigma_j,\pi}_{s_j}\left[\bigcup_{u \in U_j} Run(u)\right]$. Applying Lemma 4, part 10 to the latter number we obtain

$$\forall j \geq 0 : \inf_{\pi} \mathbb{P}^{\sigma,\pi}_{s_j}[W] > \frac{1}{3} \cdot Val(s_j). \tag{17}$$

For each $j > J(s)$ let $\mathcal{R}_j \subseteq Run(s)$ be the set of all runs starting in s which hit s_j and do not hit any other reset before s_j. By the definition of resets and by using Lemma 4, part 3, an easy induction on $j - J(s)$ shows

$$\forall j > J(s), \quad \mathcal{R}_j \neq \emptyset : \forall k, \quad J(s) \le k < j : \inf_{\pi} \mathbb{P}_{s_j}^{\sigma_k,\pi}[W] < \frac{1}{3} \cdot Val(s_j). \tag{18}$$

In particular, (17) and (18) together yield:

$$\forall j > J(s), \quad \mathcal{R}_j \neq \emptyset : \inf_{\pi} \mathbb{P}_{s_j}^{\sigma, \quad \pi}[W] \ge \inf_{\pi} \mathbb{P}_{s_j}^{\sigma_{J(s)},\pi}[W]\,. \tag{19}$$

By Lemma 4, part 8, we know that $\Phi(s, \sigma_{J(s)})$ is false. Thus:

$$\inf_{\pi} \mathbb{P}_{s}^{\sigma_{J(s)},\pi}[W] \ge \frac{1}{3} \cdot Val(s). \tag{20}$$

We now derive (16) from (20) and (19) by partitioning the runs leaving s. First let us consider runs from $Run(s) \setminus \bigcup_{j>J(s)} \mathcal{R}_j$. By definition of σ and $\mathcal{R}_j$, we immediately have that $\sigma(w) = \sigma_{J(s)}(w)$ as long as w is a prefix of a run from $Run(s) \setminus \bigcup_{j>J(s)} \mathcal{R}_j$. Now observe that all the states where σ starts to differ from $\sigma_{J(s)}$ are of the form s_j for some $j > J(s)$. From there, by (19), the probability of winning under σ is no worse than under $\sigma_{J(s)}$. In total, σ performs in s at least as good as $\sigma_{J(s)}$, thus we derive (16) from (20). $\square$

3.2. When Min has optimal strategies

Proposition 1. *Let $\mathcal{G} = (G, W)$ be a reachability game satisfying Assumption 1 where player Min has a strategy optimal in each state. Then $\mathcal{G}$ is strongly determined.*

Proof. For the reachability condition W, every state $s \in S$, every threshold p, $0 \le p \le 1$ and for both $\lhd \in \{<, \le\}$ we have to prove that either

$$\exists \sigma \text{ for Max} : \forall \pi \text{ for Min} : \mathbb{P}_s^{\sigma,\pi}[W] \not\lhd p \tag{21}$$

or

$$\exists \pi \text{ for Min} : \forall \sigma \text{ for Max} : \mathbb{P}_s^{\sigma,\pi}[W] \lhd p. \tag{22}$$

Note that if $p > Val(s)$, then $\varepsilon := \frac{p-Val(s)}{2} > 0$ and the existence of an ε-optimal strategy for Max proves that, in fact, (21) always holds. Similarly, we obtain that (22) is always true for $p < Val(s)$, and it is also true when $p = Val(s)$ and $\lhd \equiv \le$, because we assumed that Min always has an optimal strategy, and thus (22) is true.

For the rest of the proof, assume that $p = Val(s)$, and $\lhd \equiv <$ in both (21) and (22). Our aim is to prove that if (22) is false then (21) is true. In other words, to prove Proposition 1 it only remains to prove for all $s \in S$:

$$\forall \pi \text{ for Min} : \exists \sigma \text{ for Max} : \mathbb{P}_s^{\sigma,\pi}[W] \ge Val(s) \implies \exists \sigma \text{ for Max} : \forall \pi \text{ for Min} : \mathbb{P}_s^{\sigma,\pi}[W] \ge Val(s). \tag{23}$$

To prove this we will modify the game $\mathcal{G}$ to a game $\mathcal{H}$, use Lemma 5 on $\mathcal{H}$ to obtain an optimal strategy for Max in $\mathcal{H}$, and lift this strategy back to $\mathcal{G}$.

In this proof it will be key to distinguish in which game we are measuring the value. We will thus extend our notation and note the game as an index of the value, so that, e.g., the value of a state r in a game $\mathcal{H}$ is denoted by $Val_{\mathcal{H}}(r)$. We also need the notion of a *good* state and of a *value preserving* edge.

We call a state $s \in S$ *good* if it satisfies in $\mathcal{G}$: $\forall \pi$ for Min : $\exists \sigma$ for Max : $\mathbb{P}_s^{\sigma,\pi}[W] \ge Val_{\mathcal{G}}(r)$. In other words, a state s is good if (22) is false for it. An edge $r \to s$ in G is *value preserving* if either $r \in S_0$, or $Val_{\mathcal{G}}(r) = Val_{\mathcal{G}}(s)$.

We know that $\mathcal{G}$ satisfies Assumption 1. We will use $T \subseteq S$ for the target set of W, meaning that W consists of exactly all runs hitting T. If $T = \emptyset$ then $W = \emptyset$ and thus (23) is clearly true. Thus from now we will assume that $T \neq \emptyset$. We will also assume that $\mathcal{G}$ contains a state $sink \in S \setminus T$, such that no run from $sink$ ever hits T. If $\mathcal{G}$ does not contain such a state, we can add new states $sink_i$ for all $i \ge 0$ with the only outgoing transition being $sink_i \to sink_{i+1}$, and set $sink := sink_0$.[3] This does not change validity of (23) for the original states, so this assumption is without loss of generality.

We now define the game $\mathcal{H}$. Let us denote by $G = (S, \to, \delta_G)$ the game graph of $\mathcal{G}$. We will now define the game graph $H = (S', \hookrightarrow, \delta_H)$ of $\mathcal{H}$. The set S' is the union of S and the set $C = \{r_c \mid r \in S_2\}$.[4] Let $\hat{t}$ be an arbitrary but fixed state from T. For $r, s \in S'$, there is an edge $r \hookrightarrow s$ iff one of the following holds true:

- $r \in S_0$ and $r \to s$;
- $r \in S_1$ is good, s is good, and $r \to s$ is value preserving;
- $r \in S_1$ is not good and $r \to s$;
- $r \in S_2$ and either $r \to s$ is value preserving, or $s = r_c \in C$;
- $r \in C$ and $s = sink$ or $s = \hat{t}$.

[3] We have to add a whole sequence of such states instead of just one absorbing state to keep Assumption 1 true.

[4] The "C" stands for adding a choice for Min, see the definition of $\hookrightarrow$.

We also partition S' as follows: $S'_1 = S_1, S'_2 = S_2, S'_0 = S_0 \cup C$. Finally, $\delta_H(s) = \delta_G(s)$ for $s \in S_0$, and $\delta_H(r_c)$ assigns probability $1 - Val_{\mathcal{G}}(r)$ to the edge $r_c \to sink$, and probability $Val_{\mathcal{G}}(r)$ to the edge $r_c \to \hat{t}$. The definition of $\mathcal{H} = (H, \bar{W})$ is completed by setting $\bar{W}$ to be all runs hitting T in $\mathcal{H}$. Note that $\bar{W} \setminus W$ are exactly the runs visiting T by going through C.

We will now prove the following claims.

1. Every good $s \in S_1$ has at least one good successor in H. All successors of a good $s \in S_2 \cup S_0$ in H are good.
2. $\mathcal{H}$ is a game.
3. For every good $s \in S$, $Val_{\mathcal{G}}(s) = Val_{\mathcal{H}}(s)$.

Proof of 1. We proceed by contradiction. First assume that $r \in S_1$, r is good, and no successor s of r in H is good. Thus, starting in any s, such that $r \hookrightarrow s$, Player Min has a strategy π in $\mathcal{G}$ such that $\forall\sigma$ for Max : $\mathbb{P}^{\sigma,\pi}_s[W] < Val_{\mathcal{G}}(s) = Val_{\mathcal{G}}(r)$. Further, for every $r \to s$ such that $r \not\hookrightarrow s$ we know that $Val_{\mathcal{G}}(r) > Val_{\mathcal{G}}(s)$, and thus in $\mathcal{G}$ Min has a strategy π such that $\forall\sigma$ for Max : $\mathbb{P}^{\sigma,\pi}_s[W] \leq Val_{\mathcal{G}}(s) < Val_{\mathcal{G}}(r)$. Putting these strategies together, in $\mathcal{G}$, Min now has a strategy π such that $\forall\sigma$ for Max : $\mathbb{P}^{\sigma,\pi}_r[W] < Val_{\mathcal{G}}(r)$. This is a contradiction with r being good.

Now assume that $r \in S_2$ and that there is some successor s of r in H which is not good. This means that Min has a strategy π in $\mathcal{G}$ such that $\forall\sigma$ for Max : $\mathbb{P}^{\sigma,\pi}_s[W] < Val_{\mathcal{G}}(s) = Val_{\mathcal{G}}(r)$. Modifying this strategy so that it chooses the edge $r \to s$ proves that r is not good either, a contradiction.

Finally, assume that $r \in S_0$ and that there is some successor s of r in H, and hence also in G, which is not good. This means that Min has a strategy π in $\mathcal{G}$ such that $\forall\sigma$ for Max : $\mathbb{P}^{\sigma,\pi}_s[W] < Val_{\mathcal{G}}(s)$. Min has also an optimal strategy in every other successor s of r, so that $\forall\sigma$ for Max : $\mathbb{P}^{\sigma,\pi}_s[W] \leq Val_{\mathcal{G}}(s)$. We know that $Val_{\mathcal{G}}(r) = \sum_{r\to s} \delta(r)(s) \cdot Val_{\mathcal{G}}(s)$, and that $\delta(r)(s) > 0$ for all $r \to s$. Thus, starting in r and using the strategy π for the non-good successor, and the optimal strategy for all other successors, Min can always achieve probability of W being strictly less than $Val_{\mathcal{G}}(r)$. This contradicts r being good. $\square$

Proof of 2. By 1, the definition of $\hookrightarrow$, and because G is a game graph, every state in H has a successor. Also for every $s \in S_0$ the sets of successors in G and in H are equal, thus δ_H is well defined. Consequently, H is a game graph and $\mathcal{H}$ is a game. $\square$

Proof of 3. Min has an optimal strategy in $\mathcal{G}$. By Corollary 2, this strategy only uses value preserving edges, which are preserved in $\mathcal{H}$. Thus we know that $Val_{\mathcal{G}}(s) \geq Val_{\mathcal{H}}(s)$ for all $s \in S$.

We will prove $Val_{\mathcal{G}}(r) \leq Val_{\mathcal{H}}(r)$ by contradiction. Assume that there is a good $s \in S$ such that $\varepsilon := Val_{\mathcal{G}}(s) - Val_{\mathcal{H}}(s) > 0$. We will construct a strategy π for Min in $\mathcal{G}$ witnessing that s is not good, yielding the desired contradiction. We will put it together from the following strategies:

- $\bar{\pi}_{\mathcal{G}}$ is an optimal strategy of Min in $\mathcal{G}$.
- $\bar{\pi}_{\mathcal{H}}$ is a strategy of Min in $\mathcal{H}$, which is $\frac{\varepsilon}{2}$-optimal in s.[5]
- For all $r \in S$ which are *not* good, $\hat{\pi}_r$ is a strategy of Min in $\mathcal{G}$ satisfying: $\forall\sigma$ of Max : $\mathbb{P}^{\sigma,\hat{\pi}}_r[W] < Val_{\mathcal{G}}(r)$.

Now we define π. In this paragraph, a state called r is always assumed to be of Player Min. First consider a situation when only good states have been visited so far, r being the current state. Denoting $p := \bar{\pi}_{\mathcal{H}}(r)(r_c)$, we set for all $r \to t$: $\pi(r)(t) := \bar{\pi}_{\mathcal{H}}(r)(t) + p \cdot \bar{\pi}_{\mathcal{G}}(r)(t)$. Second, consider the situation where the last transition was $\bar{r} \to r$, $\bar{r} \in S_1$ and $Val_{\mathcal{G}}(\bar{r}) > Val_{\mathcal{G}}(r)$. Then $\pi(t) := \bar{\pi}_{\mathcal{G}}(t)$ for all states t visited after r, including r. Third, the remaining situation must be where the last transition was $\bar{r} \to r$, $Val_{\mathcal{G}}(\bar{r}) = Val_{\mathcal{G}}(r)$, but $\bar{r}$ being good, while r not. Then $\pi(t) := \hat{\pi}_r(t)$ for all states t visited after r, including r. Observe that a switch from a good to a non-good state can only be made from a state of Player Max. Thus, we just defined the behaviour of π both for all histories consisting of only good states, and for histories starting in a good state but switching to a non-good one at some point. Because the initial state s is good, the definition of π is complete.

Now we argue that $\mathbb{P}^{\sigma,\pi}_s[W] < Val_{\mathcal{G}}(s)$, contradicting that s is good. We call the states r from the second and the third situation above *breakpoints*, and denote by B the set of all breakpoints. Observe that if a path from s in G never hits a breakpoint then it is also a path in H, and it only uses value preserving edges. In particular, for all σ of Max:

$$Val_{\mathcal{G}}(s) \geq \mathbb{P}^{\sigma,\pi}_s[W \cap \neg Reach(B)] + \sum_{r\in B} \mathbb{P}^{\sigma,\pi}_s[Reach(r)] \cdot \mathbb{P}^{\sigma,\pi}_r[W] .$$

Further, by our choice of breakpoints, and the strategies $\bar{\pi}_{\mathcal{G}}$ and $\hat{\pi}_r$, for every breakpoint r: $\mathbb{P}^{\sigma,\pi}_r[W] < Val_{\mathcal{G}}(r)$. Thus if $\mathbb{P}^{\sigma,\pi}_s[Reach(B)] > 0$ we have $\mathbb{P}^{\sigma,\pi}_s[W] < Val_{\mathcal{G}}(s)$. On the other hand, if $\mathbb{P}^{\sigma,\pi}_s[Reach(B)] = 0$ then all runs stay in H and π behaves very similarly to $\bar{\pi}_{\mathcal{H}}$. The only difference is that where $\bar{\pi}_{\mathcal{H}}$ would chose $r \hookrightarrow r_c$, π behaves like an optimal strategy in $\mathcal{G}$. Because by construction of $\mathcal{H}$, $Val_{\mathcal{H}}(r_c) = Val_{\mathcal{G}}(r)$, we get $\mathbb{P}^{\sigma,\pi}_s[W] = \mathbb{P}^{\sigma,\pi}_s\left[\bar{W}\right] \leq \mathbb{P}^{\sigma,\bar{\pi}_{\mathcal{H}}}_s\left[\bar{W}\right] \leq Val_{\mathcal{H}}(s) + \frac{\varepsilon}{2} < Val_{\mathcal{G}}(s)$. For the last inequality, recall that $\varepsilon > 0$ is the difference $Val_{\mathcal{G}}(s) - Val_{\mathcal{H}}(s)$. The proof of claim 3 is finished. $\square$

Finally, we prove (23) by giving, for every good $r \in S$, a strategy for Max which is optimal in r in $\mathcal{G}$. Observe that the game graph of $\mathcal{H}$ is a forest. Given a state, $r \in S$, we denote by $\mathcal{H}_r$ the sub-game spanned by the tree rooted in r.

[5] Note that unlike in $\mathcal{G}$, we cannot assume that Min has an optimal strategy in $\mathcal{H}$.

Let us fix a good state, $r \in S$. By definition of $\mathcal{H}$ and by Claim 1, all states from S included in $\mathcal{H}_r$ are good, and only value preserving edges are included in $\mathcal{H}_r$. Because by 3 the values in $\mathcal{G}$ and $\mathcal{H}$ are the same for every good state, we have that $\mathcal{H}_r$ satisfies Assumption 2. Lemma 5 applied to $\mathcal{H}_r$ yields a strategy σ for Max which is optimal in r in $\mathcal{H}$. We now lift σ to $\mathcal{G}$ so that it stays optimal in r.

We need to define σ in $\mathcal{G}$ on non-good successors of good states. Let $s \to t$ be an edge in G such that $s \in S$ is a good state, and $t \in S_1$ is not. We need to define σ from t onwards. This cannot happen for $s \in S_0$. If $s \in S_1$ there is no need to define σ for t, because t is never entered under σ (recall that G is a forest). If $s \in S_2$ then $\varepsilon := Val_{\mathcal{G}}(t) - Val_{\mathcal{G}}(s) > 0$. We let σ behave like a $\frac{\varepsilon}{2}$-optimal strategy from t on.

The last task is now to prove that this σ is optimal in r in $\mathcal{G}$. Let π be an arbitrary strategy of Min for $\mathcal{G}$. We define a strategy π' of Min in $\mathcal{H}$ such that $\mathbb{P}_r^{\sigma,\pi}[W] \geq \mathbb{P}_r^{\sigma,\pi'}\left[\bar{W}\right]$, where the former is measured in $\mathcal{G}$, and the latter in $\mathcal{H}$. Because σ is optimal in r in $\mathcal{H}$ and $Val_{\mathcal{G}}(r) = Val_{\mathcal{H}}(r)$, and because π was chosen arbitrarily, this proves that σ is optimal in r in $\mathcal{G}$.

The strategy π' copies the behaviour of π as far as value preserving edges are chosen. More precisely, $\pi'(s)(t) = \pi(s)(t)$ for all $s \to t$ where $Val_{\mathcal{G}}(s) = Val_{\mathcal{G}}(t)$, and $\pi'(s)(s_c) = \sum\{\pi(s)(t) \mid s \to t, Val_{\mathcal{G}}(s) < Val_{\mathcal{G}}(t)\}$.

By construction of $\mathcal{H}$, $\mathbb{P}_{s_c}^{\sigma,\pi'}\left[\bar{W}\right] = Val_{\mathcal{G}}(s)$. Thus by switching to s_c, the strategy π' only decreased the probability of Max winning, because the corresponding move under π resulted in a state t with $Val_{\mathcal{G}}(t) > Val_{\mathcal{G}}(s)$ and σ was chosen to ensure winning with probability at least $Val_{\mathcal{G}}(t) - \frac{Val_{\mathcal{G}}(t) - Val_{\mathcal{G}}(s)}{2} > Val_{\mathcal{G}}(s)$. Thus $\mathbb{P}_r^{\sigma,\pi}[W] \geq \mathbb{P}_r^{\sigma,\pi'}\left[\bar{W}\right]$, and the proof of the proposition is done. $\square$

3.3. When Max has optimal strategies

Proposition 2. *Let $\mathcal{G}$ be a reachability game satisfying Assumption 1 where player Max has a strategy optimal in each state. Then $\mathcal{G}$ is strongly determined.*

Proof. Almost identically to the proof of Proposition 1, we first eliminate the easy cases. For the reachability condition W, every state $s \in S$, every threshold p, $0 \leq p \leq 1$ and for both $\lhd \in \{<, \leq\}$ we have to prove that either

$$\exists\sigma \text{ for Max} : \forall\pi \text{ for Min} : \mathbb{P}_s^{\sigma,\pi}[W] \ntriangleleft p \tag{24}$$

or

$$\exists\pi \text{ for Min} : \forall\sigma \text{ for Max} : \mathbb{P}_s^{\sigma,\pi}[W] \lhd p. \tag{25}$$

If $p < Val(s)$, then $\varepsilon := \frac{Val(s)-p}{2} > 0$ and the existence of an ε-optimal strategy for Min proves that, in fact, (25) always holds. Similarly, we obtain that (24) is always true for $p > Val(s)$, and it is also true when $p = Val(s)$ and $\lhd \equiv <$, because we assumed that Max always has an optimal strategy.

For the rest of the proof, assume that $p = Val(s)$, and $\lhd \equiv \leq$ in both (24) and (25). Our aim is to prove that if (25) is false then (24) is true. In other words, to prove Proposition 2 it only remains to prove for all $s \in S$:

$$\forall\pi \text{ for Min} : \exists\sigma \text{ for Max} : \mathbb{P}_s^{\sigma,\pi}[W] > Val(s) \implies \exists\sigma \text{ for Max} : \forall\pi \text{ for Min} : \mathbb{P}_s^{\sigma,\pi}[W] > Val(s). \tag{26}$$

Claim 1. *Let $s \in S$. Assume that $\forall\pi$ for Min : $\exists\sigma$ for Max : $\mathbb{P}_s^{\sigma,\pi}[W] > Val(s)$. Then also for all π for Min there is σ for Max such that both $\mathbb{P}_s^{\sigma,\pi}[W] > Val(s)$, and $Val(r) = Val(t)$ whenever $r \in S_1$ and $\sigma(r)(t) > 0$.*

Proof. Let $R \subseteq S_1$ be the set of all states r such that there is a path from s to r in G, and if an edge $t \to t'$ is included in this path, which is unique by Assumption 1, then $Val(t) = Val(t')$ or $t \notin S_1$.

Let π be an arbitrary strategy of Min. We obtain a strategy π' by modifying π: whenever a transition $r \to t$ with $r \in R$ and $\varepsilon := Val(r) - Val(t) > 0$ is taken, π' behaves like a $\frac{\varepsilon}{2}$-optimal strategy after visiting t. Otherwise, it behaves like π. Let σ' be a strategy of Max such that $\mathbb{P}_s^{\sigma',\pi'}[W] > Val(s)$. We now define a strategy σ by modifying σ': let $r \in R$ and $p := \sum\{\sigma'(r)(t) \mid r \to t, Val(r) > Val(t)\}$. If $p = 0$ then $\sigma(r) = \sigma'(r)$. Otherwise let $\bar{\sigma}$ be a strategy optimal in r; for every $r \to t$ we set $\sigma(r)(t) = \sigma'(r)(t) + p \cdot \bar{\sigma}(r)(t)$ whenever $Val(r) = Val(t)$, and $\sigma(r)(t) = 0$ otherwise. By construction of π' and σ, we know that $\mathbb{P}_s^{\sigma,\pi}[W] = \mathbb{P}_s^{\sigma,\pi'}[W]$, and that $\mathbb{P}_s^{\sigma,\pi'}[W] \geq \mathbb{P}_s^{\sigma',\pi'}[W] > Val(s)$. In particular, $\mathbb{P}_s^{\sigma,\pi}[W] > Val(s)$. $\square$

Due to Claim 1, in proving (26) we can safely assume that for all $r \to t$ such that $r \in S_1$ we have $Val(r) = Val(t)$. In general, the graph G is a forest. When proving (26) for a fixed $s \in S$ we may, however, restrict our attention to only the states reachable from s in G. Note that the sub-game of $\mathcal{G}$ spanned by such states satisfies Assumption 2.

Let us fix an arbitrary $s \in S$. As we just observed, to finish the proof of Proposition 2 it is enough to prove (26) for s, under the assumption that G is a tree rooted in s, and $\mathcal{G}$ satisfies Assumption 2.

We start with a useful notion of a critical state. Let π be a strategy of Min. Due to our assumption there is a σ such that $\mathbb{P}_s^{\sigma,\pi}[W] > Val(s)$. Moreover, Corollary 2 allows us to derive from this inequality that there must be a state, $r_\pi \in S_2$, with a successor $t \in S$, such that $Val(r_\pi) < Val(t)$ and yet $\pi(r_\pi)(t) > 0$. Indeed, if there was no such state then by Corollary 2, π is optimal, a contradiction with our assumption. If r_π is the only state on the (unique) path from s to r_π with the above property, we call such a state r_π a *critical state* associated with π.

We will now construct a strategy σ for Max such that $\forall\pi$ for Min : $\mathbb{P}_s^{\sigma,\pi}[W] > Val(s)$, thus proving the right-hand part of (26).

The idea is to use some strategy of Max, which is optimal in s, but at the same time to reach all the states r_π for all possible π. We achieve this by employing two strategies, an optimal σ_O, and an "exploratory" σ_E which explores all paths to possible critical states. To ensure the resulting strategy is optimal again, we will mix the decisions of both strategies so that the weight put on the latter, "exploratory", one tends to 0 as the number of steps from s increases, but stays positive forever.

We start with defining the two strategies. The strategy σ_O is optimal in every state of $\mathcal{G}$. We know that such a strategy exists. The strategy σ_E is an arbitrary strategy satisfying $\sigma_E(r)(t) > 0$ iff $r \to t$ and $r \in S_1$. Such a strategy obviously exists, even if the game is countably infinitely branching.

Now we define σ. Recall that G is a tree rooted in s. Thus for every state r there is exactly one path from s to r in G. Its length is finite, we denote it by $n_r > 0$. In particular, $n_s = 1$. For every state $r \in S_1$ we define σ by

$$\sigma(r) := 2^{-(n_r+1)} \cdot \sigma_E(r) + (1 - 2^{-(n_r+1)}) \cdot \sigma_O(r).$$

Because the probability that, starting in some r, σ behaves forever like σ_O is at least $1 - \sum_{j=n_r+1}^{\infty} 2^{-j} = 1 - 2^{-n_r} \geq \frac{1}{2}$, in every $r \in S$ we have

$$\inf_\pi \mathbb{P}_r^{\sigma,\pi}[W] \geq \frac{1}{2} \cdot Val(r).$$

By Lemma 3, σ is thus optimal, in every $r \in S$. Moreover, let π be an arbitrary strategy of Min, and r_π a critical state associated with π. Because $\mathbb{P}_s^{\sigma_E,\pi}[Reach(r_\pi)] > 0$ we also have $\mathbb{P}_s^{\sigma,\pi}[Reach(r_\pi)] > 0$. Further, $\mathbb{P}_{r_\pi}^{\sigma,\pi}[W] > Val(r_\pi)$ by the definition of r_π, and we showed already that σ is optimal in every $r \in S$. As a consequence, $\mathbb{P}_s^{\sigma,\pi}[W] > Val(s)$. The proposition is proved. □

4. Optimal strategies

In this section we prove Theorem 2. We do this in two steps, first for games without Player Min (also known as Markov decision processes, see, e.g., [10] for general reference and [2,5] for specific context of reachability objectives), and then for the general case.

4.1. Games without Player Min

Proposition 3. *Theorem 2 is true for all games $\mathcal{G} = (G, W)$, $G = (S, \to, \delta)$ where $S_2 = \emptyset$.*

This whole subsection is devoted to proving the above proposition. We fix the game $\mathcal{G}$ in the rest of the subsection. By Lemma 1 we can assume that $\mathcal{G}$ satisfies Assumption 1. In particular, G is a forest, and there is $T \subseteq S$ such that $W = \bigcup\{Run(w) \mid w \text{ ends in } T\}$ and T is closed on outgoing transitions. The proof is by contradiction, in three steps. First, we prove that if there is a state with no optimal strategy, then there must be a state from which winning with probability sufficiently close to the optimum implies the need to use some value decreasing transition. A transition $r \to s$ is value decreasing if $Val(r) > Val(s)$. Second, we will argue that the potential "damage" caused by this transition is positive and bounded away from 0, independently of the actual strategy. Third, we show that $(*)$ implies that the potential "damage" factor is indeed bounding the probability of reaching T away from the value of the initial state, which is a contradiction with the definition of the value.

We introduce a random variable, L (for "loss"). For a run, ω, a *losing* index is every i, such that $\omega(i) \in S_1$ and $Val(\omega(i)) > Val(\omega(i+1))$. If there is no losing index for ω, we set $L(\omega) := 0$. Otherwise, there is the least losing index, i, and we set $L(\omega) := Val(\omega(i)) > 0$. Finally, we say that a state $s \in S$ is *losing* if there is some $\delta_s > 0$ such that for every δ_s-optimal strategy, σ, in s, we have $\mathbb{P}_s^\sigma[L > 0] > 0$.

Lemma 6. *Assume $(*)$. If $\exists s \in S$ such that $\forall \sigma : \mathbb{P}_s^\sigma[W] < Val(s)$ then there is also some losing state.*

Proof. By contradiction. Assuming there is no losing state, we construct an optimal strategy in every state. We define a subset $\hookrightarrow$ of the transition relation $\to$ of $\mathcal{G}$, by setting for every pair $r, s \in S$: $r \hookrightarrow s$ iff $r \to s$ and either $r \in S_0$, or $Val(r) = Val(s)$. By Corollary 1, for all $r \in S_1$ there is at least one s such that $r \to s$ and $Val(r) = Val(s)$. Thus $\hookrightarrow$ is total and $G' = (S, \hookrightarrow, \delta)$ is a game graph. Without losing states, for every $r \in S$ and every $\varepsilon > 0$ there is some ε-optimal strategy, σ, such that $\mathbb{P}_s^\sigma[L > 0] = 0$, i.e., σ does not use value-decreasing transitions. This strategy works in $\mathcal{G}' = (G', W)$ as well, winning with the same probability, as in $\mathcal{G}$. The values in $\mathcal{G}$ and $\mathcal{G}'$ are thus the same.

Consider now $\mathcal{G}'$. For each $s \in S$, denote by $FP_k(s)$ the set of all finite paths of length k starting in s. Because T is closed on outgoing transitions, and because $\hookrightarrow$ preserves value, the following is true in $\mathcal{G}'$:

$$\forall k \geq 0 : \forall \sigma : \forall s \in S : Val(s) = \sum_{w \in FP_k(s)} \mathbb{P}_s^\sigma[Run(w)] \cdot Val(w(k)). \tag{27}$$

For all $s \in S$ fix a strategy σ_s which is $\frac{1}{4} \cdot Val(s)$-optimal in s. After some $n_s \geq 0$ of steps, T must be reached from s under σ_s with probability at least $\frac{1}{2} \cdot Val(s)$, because $\mathbb{P}_s^{\sigma_s}[W] = \lim_{k\to\infty} \mathbb{P}_s^{\sigma_s}[\{Run(w) \mid len(w) \leq k \wedge w(k) \in T\}]$.

For all $s \in S$ we finally construct a strategy σ for $\mathcal{G}'$, optimal in s. Because the values are the same in $\mathcal{G}$ and $\mathcal{G}'$, and every strategy for $\mathcal{G}'$ is also a strategy for $\mathcal{G}$, this will finish the proof of the lemma. The strategy σ starts in s according to σ_s, and follows it for n_s steps. After that, having arrived to some state r, it switches to σ_r and follows it for other n_r steps. This is repeated ad infinitum. The invariant (27), and the choice of n_r and σ_r for $r \in S$, guarantee that after the m-th stage of the above repetitive process, T has actually been reached with probability $(1-2^{-m}) \cdot Val(s)$. Because $\lim_{m\to\infty}(1-2^{-m}) = 1$, σ is optimal. □

For every losing state, $s \in S$, and every constant $\varepsilon > 0$ we define $\ell_s^\varepsilon := \inf\{\mathbb{E}_s^\sigma[L] \mid \sigma \text{ is } \varepsilon\text{-optimal in } s\}$. Since $\ell_s^\varepsilon \le \ell_s^\zeta \le 1$ for $\varepsilon \ge \zeta$, the limit $\ell_s := \lim_{\varepsilon\to 0} \ell_s^\varepsilon$ exists.

Lemma 7. *Assume* $(*)$*. For every losing state, s, in* $\mathcal{G}$ *we have* $\ell_s > 0$.

Proof. We fix some $s \in S$ and proceed by contradiction: assume that s is losing and $\ell_s = 0$. To every strategy σ which may possibly use value-decreasing transitions $r \to r'$ where $Val(r) > Val(r')$ we consider a strategy $\bar{\sigma}$, which copies the moves of σ until a value-decreasing transition is chosen. From that point on, just before the value-decreasing transition, the strategy $\bar{\sigma}$ keeps choosing arbitrary successors with the only requirement that they preserve the value, i.e., whenever $\bar{\sigma}$ chooses a transition $t \to t'$, $t, t' \in S$ with a positive probability then $Val(t) = Val(t')$. Such a choice always exists, by Corollary 1. Observe that for every σ, $\mathbb{P}_s^\sigma[W] - \mathbb{P}_s^{\bar{\sigma}}[W] \le \mathbb{E}_s^\sigma[L]$. As a consequence, due to $\ell_s = 0$, $Val(s) = \sup\{\mathbb{P}_s^{\bar{\sigma}}[W] \mid \sigma \text{ is some strategy}\}$. This contradicts s being losing, since $\mathbb{P}_s^{\bar{\sigma}}[L > 0] = 0$ for every σ. □

Now we prove Proposition 3, by contradiction. Assume $(*)$, and that there is some $r \in S$ with no strategy optimal in r. By Lemma 6, there is a losing state, $s \in S$, i.e., there is a state s such that whenever an ε-optimal strategy is used from s for a small enough ε, then with a positive probability that strategy chooses a value decreasing transition.

This is not precise enough, we need to choose $\varepsilon > 0$ and find a fixed positive threshold such that if a strategy is ε-optimal, then the decreasing transition is taken from a state with value above the threshold with a probability bounded from below from zero. Once we have that we can use the property $(*)$ to deduce a lower bound on how much the strategy actually loses by choosing the value decreasing transition, and show that this contradicts the existence of strategies arbitrarily approaching optimality.

Let us start with the choice of ε. By Lemma 7, $\ell_s > 0$, so there is some $\varepsilon > 0$ such that $\ell_s^\varepsilon \ge \ell_s/2 > 0$. Recall that ℓ_s^ε is a lower bound on expectations, and if an expectation of a random variable is above a constant, then also with positive probability the random value is above that constant. Here this means that starting in s, under every ε-optimal strategy, σ, with some positive probability, $p_\sigma > 0$, a state $r \in S_1$ with $Val(r) \ge \ell_s^\varepsilon$ is visited, and some transition $r \hookrightarrow r'$ with $Val(r') < Val(r)$ is taken.

Let us step back for a moment: we have chosen our ε, and we also have our positive threshold: ℓ_s^ε. But we have not yet separated from zero the probability that this threshold is exceeded, because the probability, p_σ, although always positive, depends on σ and could perhaps be approaching 0. Now we need to show that it is in fact bounded away from 0 for all ε-optimal strategies σ.

Observe that $(*)$ gives us the following "value-gap":

$$\delta := \inf\{|Val(r) - Val(r')| \mid r, r' \in S, Val(r) \ne Val(r'), Val(r) \ge \ell_s^\varepsilon\} > 0.$$

This allows us to bound p_σ independently of σ, as follows. If $p_\sigma \ge \ell_s^\varepsilon$ for all σ then we are done. Otherwise the set $K := \{Val(r) \mid r \in S, Val(r) < \ell_s^\varepsilon\}$ is non-empty, and thus $\xi := \sup K$ well defined and, by $(*)$, $\xi < \ell_s^\varepsilon$. We know that the loss incurred in states with value greater than ℓ_s^ε is bounded from above by 1, and the other losses are at most ξ. Thus we know

$$\ell_s^\varepsilon \le \mathbb{E}_s^\sigma[L] \le p_\sigma \cdot 1 + (1 - p_\sigma)\xi.$$

Hence, we have a positive lower bound, p, on p_σ, independent of σ:

$$p_\sigma \ge p \ge \min\left(\frac{\ell_s^\varepsilon - \xi}{1 - \xi}, \ell_s^\varepsilon\right) > 0.$$

Thus for every strategy, σ, we have that $Val(s) - \mathbb{P}_s^\sigma[W] \ge \min\{\varepsilon, \delta \cdot p\} > 0$. This clearly contradicts the definition of $Val(s)$. The proof is finished.

4.2. General case

In this subsection we prove Theorem 2 without restrictions on the set of states of Player Min. Let us fix a game $\mathcal{G} = (G, W)$, where $G = (S, \to, \delta)$ and W is open. Due to Lemma 1, we may assume that $\mathcal{G}$ satisfies Assumption 1. Again, by $T \subseteq S$ we denote the target set for the reachability objective $W = \bigcup\{Run(w) \mid w \text{ ends in } T\}$. We call a state s *safe* if $\forall\sigma$ for Max : $\exists\pi_\sigma$ for Min : $\mathbb{P}_s^{\sigma,\pi_\sigma}[W] = 0$. The following lemma states the strong determinacy restricted to states with value 0, and will be useful in proving each of both theorems. Note that we cannot use Theorem 1 because we did not prove yet that Max has always an optimal strategy in $\mathcal{G}$.

Lemma 8. *For every safe* $s \in S$*:* $\exists\pi$ *for Min* : $\forall\sigma$ *for Max* : $\mathbb{P}_s^{\sigma,\pi}[W] = 0$.

Proof. We cut off some choices for Min in the game graph G of $\mathcal{G}$, and obtain its sub-graph G', so that all states reachable in G' from s have value 0 in $\mathcal{G}' = (G', W)$. In particular, no run can satisfy W. Because the choices of Max remain unrestricted in G', this ensures that the probability of W is 0 in $\mathcal{G}$ as well. Let us proceed in more detail.

Observe that every safe state has value 0, so no safe state is in T. Also, observe that for every safe $r \in S_0 \cup S_1$ and $s \in S$, if $r \to s$ then s is safe. Likewise, if $r \in S_2$ is safe, then there must be a safe s such that $r \to s$. Fix a safe s, and define G' as the smallest sub-graph of G containing s and satisfying that for every state r in G', every safe successor r' of r is also in G. As shown above, G' is a game graph, the probability assignment δ from G is valid in G' as well, and all states in G' are safe in $\mathcal{G}$. Hence, no paths in G' visit T, and the value of every state in $\mathcal{G}'$ is 0. Fix an arbitrary strategy π for Min in $\mathcal{G}' = (G', W)$, then $\mathbb{P}_s^{\sigma,\pi}[W] = 0$ for all σ of Max in $\mathcal{G}'$. All transitions out of safe states of Max were preserved in G', and π is also a strategy in $\mathcal{G}$, so we have $\mathbb{P}_s^{\sigma,\pi}[W] = 0$ also for every σ of Max in $\mathcal{G}$. $\square$

Lemma 9. *If $\mathcal{G}$ satisfies $(*)$, then for all $s \in S$ we have: $\forall \pi$ for Min : $\exists \sigma$ for Max : $\mathbb{P}_s^{\sigma,\pi}[W] \geq Val(s)$.*

Proof. Lemma 1, part 5 allows us to restrict our attention to memoryless strategies. For every memoryless strategy π of Player Min, we denote by $\mathcal{G}_\pi$ the game where the choices of Player Min are resolved using π. Formally, $\mathcal{G}_\pi = (G', W)$, where $G' = (S', \hookrightarrow, \delta')$, and (1) $S' = S$ but comes with a different partition: $S'_0 = S_0 \cup S_2$, $S'_1 = S_1$, $S'_2 = \emptyset$, (2) the relation $\hookrightarrow \subseteq \to$ is given by $r \hookrightarrow s$ iff $r \to s$ and either $r \in S_0 \cup S_1$, or $r \in S_2$ and $\pi(r)(s) > 0$, and (3) $\delta' = \delta \cup \pi$. For every strategy σ for Player Max, and every $s \in S$ the measure $\mathbb{P}_s^{\sigma,\pi}[\cdot]$ in $\mathcal{G}$ obviously coincides with $\mathbb{P}_s^{\sigma}[\cdot]$ in $\mathcal{G}_\pi$. Thus we may apply Proposition 3 to all $\mathcal{G}_\pi$ to derive the lemma. $\square$

Consider now the following game $\mathcal{H} = (H, W)$, which is a slight modification of $\mathcal{G}$. The set of states of $H = (S, \hookrightarrow, \delta_H)$ is S, the same as in G, and with the same partition. There is a transition $r \hookrightarrow s$ iff exactly one of these three situations occurs: $Val(r) = 0$ in $\mathcal{G}$, and $s = r$; or $Val(r) > 0$, $r \in S_0$ and $r \to s$; or $Val(r) > 0$, $r \notin S_0$, $r \to s$, and $Val(r) = Val(s)$ in $\mathcal{G}$. In other words, in H we made all states with value 0 absorbing, and only left value preserving transitions for players. Finally, δ_H is the only probability weight function which coincides with δ on stochastic states with positive value.

Lemma 10. *If $\mathcal{G}$ satisfies $(*)$, then H is a game graph, and the values are the same in $\mathcal{G}$ and $\mathcal{H}$.*

Proof. We refine the modifications from above into three steps, obtaining game graphs $G = H_0$, H_1, H_2, and $H_3 = H$. We will show for each $i \in \{1, 2, 3\}$ that H_i is a game graph, and that the values are the same in $\mathcal{H}_i = (H_i, W)$ as they are in $\mathcal{G}$. All the graphs H_i constructed for $i \geq 1$ have the same set of states, S, and the same partition, as G, and the same weight function, δ_H, as H.

$H_1 = (S, \mapsto, \delta_H)$, and $r \mapsto s$ iff $Val(r) = 0$ in $\mathcal{G}$, and $s = r$, or $Val(r) > 0$ and $r \to s$. H_1 is clearly a game graph, because $\mapsto$ is total. The values do not change, because each absorbing loop outside T has value 0. Moreover, every $r \in S_2$ always has a successor with the same value. Indeed, if $Val(r) = 0$ then r itself is its own successor in H_1; if $Val(r) > 0$ then this follows from Corollary 1. By Corollary 2, Min has optimal strategies in $\mathcal{H}_1$.

$H_2 = (S, \rightsquigarrow, \delta_H)$, and $r \rightsquigarrow s$ iff $r \mapsto s$ and either $Val(r) = 0$ in $\mathcal{G}$, or $r \notin S_2$, or (if $Val(r) > 0$ and $r \in S_2$) $Val(r) = Val(s)$ in $\mathcal{G}$. Because Min has always value-preserving transitions in $\mathcal{H}_1$, H_2 is clearly a game graph, and by Corollary 2 all strategies of Min in $\mathcal{H}_2$ are optimal. Fix one such π for Min, and an arbitrary $s \in S$. By Lemma 9 there is a σ for Max in $\mathcal{G}$ (and thus also in $\mathcal{H}_2 = (H_2, W)$) such that $\mathbb{P}_s^{\sigma,\pi}[W] \geq Val(s)$. We will now show that the choices of σ made in $\mathcal{H}_2$ are also valid in the game graph H_3: because π is optimal in $\mathcal{H}_2$, σ cannot choose value-decreasing transitions in $\mathcal{H}_2$. In other words, σ is restricted to use only edges in $\hookrightarrow$, i.e., from $H_3 = H$. This proves that H_3 is a game graph. Now for values, observe that only the choices of Player Max were restricted in H_3, compared to H_2. Choices of Player Min remain the same. So the values could only potentially drop in $\mathcal{H}$ compared to $\mathcal{H}_2$. However, by the choice of σ, we know that $\mathbb{P}_s^{\sigma,\pi}[W]$ measured in $\mathcal{H}$ is at least $Val_{\mathcal{H}_2}(s)$. Thus also the values in $\mathcal{H}$ and $\mathcal{H}_2$, and thus also in $\mathcal{G}$, are the same. $\square$

Lemma 11. *If $\mathcal{G}$ satisfies $(*)$, then Max has an optimal strategy, σ, in $\mathcal{H}$.*

Proof. We show that this is an easy application of Lemma 5. Although $\mathcal{H}$ does not satisfy Assumption 1, by Lemma 1 there is a game $\mathcal{H}'$ containing all states of $\mathcal{H}$, such that for each state r of $\mathcal{H}$, $Val_{\mathcal{H}}(r) = Val_{\mathcal{H}'}(r)$, and there is an optimal strategy σ for Max in $\mathcal{H}$ if there is one in $\mathcal{H}'$. Because $\mathcal{G}$ satisfies $(*)$, Lemma 10 guarantees that, $\mathcal{H}$ is a well defined game and all transitions leaving the states of Player Max in particular, preserve value. Thus, $\mathcal{H}'$ satisfies Assumption 2, and by Lemma 5, Max has an optimal strategy in $\mathcal{H}'$. By Lemma 1, Max also has an optimal strategy in $\mathcal{H}$. $\square$

Now we prove Theorem 2. Consider the strategy σ from Lemma 11. It partially defines a strategy in $\mathcal{G}$. To complete its definition, we now specify it for histories containing a transition of the form $r \to s$, where $r \in S_2$ and $Val(s) > Val(r)$, by requiring σ to behave as a $1/2 \cdot (Val(s) - Val(r))$-optimal strategy after that point. Fix an initial state, s, and consider an arbitrary strategy, π, of Min. If π is optimal, then it is also valid in $\mathcal{H}$, and $\mathbb{P}_s^{\sigma,\pi}[W] = Val(s)$ by Lemmas 10 and 11. For a non-optimal π it is easy to verify that $\mathbb{P}_s^{\sigma,\pi}[W] > Val(s)$ by both the definition of σ, and Lemmas 10 and 11. (A formal proof of the previous sentence can be done in the same way as proving that σ is optimal in the very last part of the proof of Proposition 1. Because that construction is technical and brings nothing new, we do not repeat it here.)

5. One Counter games

One Counter stochastic games (OC-SSGs), see, e.g., [5,3,2], are games played on transition graphs of one-counter automata. Such automata have a finite *control-state* unit, Q, and a set of rules, which are triples of the form (r, k, s) with $r, s \in Q$ and $k \in \{-1, 0, +1\}$. States of an OC-SSG, also called *configurations*, are then of the form $\langle s, n \rangle$ where $s \in Q$ is a *control state*, and $n \geq 0$ is an integer, representing the *counter value*. Transitions are generated by setting $\langle r, i \rangle \to \langle s, j \rangle$ if there is a rule

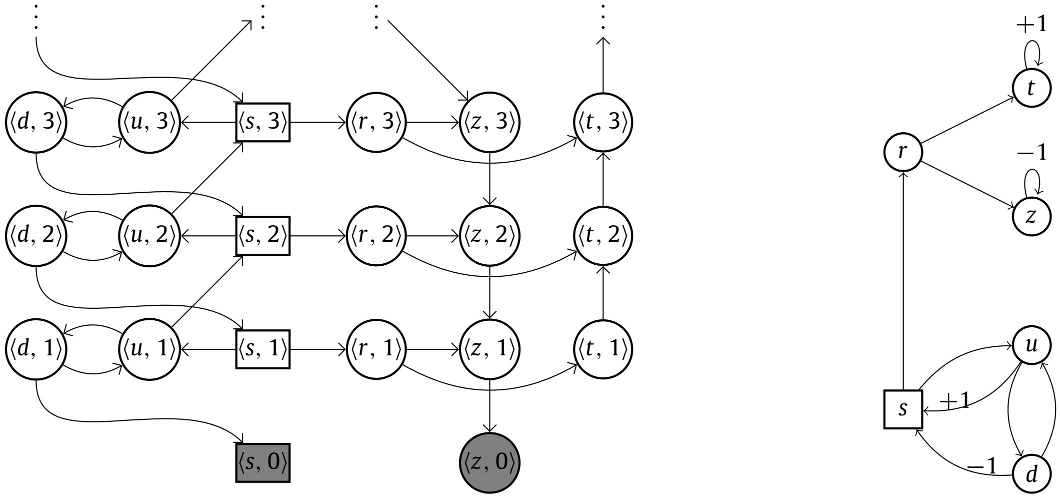

Fig. 5. Left: A game, $\mathcal{G}$, where player Max (□) does not have optimal strategies. All stochastic (◯) states have uniform distribution on outgoing transitions. Right: A One Counter description of $\mathcal{G}$. Signed numbers represent counter increments.

$(r, j - i, s)$. The partition of states is induced by a partition of Q, and the probabilities of transitions out of stochastic states, are induced by probabilities on rules. OC-SSGs come with an implicit reachability objective, the set to be reached is the set $\{\langle s, n\rangle \mid s \in Q, n \leq 0\}$ of states with counter at most 0. Because one counter machines are supposed to halt in 0 (although this is not reflected here in the transition relation) this winning condition is called *termination*.

Example 2. On the right-hand part of Fig. 5 we give the one-counter automaton with the set $Q = \{s, u, d, r, z, t\}$ of control states. An unlabelled edge, like $s \to u$, represents a 0-rule, e.g., $(s, 0, u)$. A label (± 1) represents the counter change, e.g., the loop $t \to t$ represents $(t, +1, t)$. The square-state s belongs to Max, other states are stochastic. The distributions on outgoing transitions are implicitly uniform in this example. On the left-hand half is a part of the generated game graph. Grey states are to be reached, and so we do not bother showing edges leaving them. Later in this section we will show that $Val(\langle s, i\rangle) = \frac{2^i+1}{2^{i+1}}$, but no strategy of Player Max is optimal in $\langle s, i\rangle$. Observe that $1/2 = \lim_{i\to\infty} \frac{2^i+1}{2^{i+1}}$ is an accumulation point in the set of all values.

Note that every OC-SSG has bounded out-degree and in-degree, in particular it is finitely branching. Thus, by Fact 1, configurations of Min always have at least one value-preserving outgoing transition, and by Corollary 2, Min thus has always optimal strategies in OC-SSGs. However, they may not always satisfy $(*)$, and Example 2 shows that in OC-SSGs, Max may have no optimal strategies. On the other hand, the structure of the accumulation points in the set of all values is well understood for OC-SSGs. To describe it, we need to introduce another winning condition, called $[LimInf = -\infty]$. A run satisfies $[LimInf = -\infty]$ if the lim inf of the counter value appearing in the visited configurations is $-\infty$. Note that $\inf_\pi \sup_\sigma \mathbb{P}^{\sigma,\pi}_{\langle s,n\rangle}[[LimInf = -\infty]]$ is independent of n. In what follows, we thus call the above value the $[LimInf = -\infty]$-*value of* s, as opposed to the different termination values $Val(\langle s, n\rangle)$ for different n.

In [3,2] it was shown that both players always have very simple optimal strategies with respect to the $[LimInf = -\infty]$ objective: ones that always choose one successor with probability 1, are memoryless, and in fact, only need to know the control state and ignore the counter value. It was also shown there that the optimal value is always rational and computable. Observe that the termination values, $Val(\langle s, n\rangle)$, for a fixed $s \in Q$, are non-increasing with increasing n. Thus their limit exists, and, in fact, it is an easy exercise to employ the results of [3,2] to prove that the $[LimInf = -\infty]$-value of a control state, s, equals $\lim_{n\to\infty} Val(\langle s, n\rangle)$, see also [5, Section 3]. Intuitively this is because when increasing the initial counter, n, the objective of reaching 0 or below becomes more and more similar to the $[LimInf = -\infty]$ objective. Thus the set of $[LimInf = -\infty]$-values of all states $s \in Q$ contains the set of all accumulation points of the termination values. It is also possible to decide in time polynomial in $|Q|$ whether a $[LimInf = -\infty]$-value, p, actually is an accumulation point of the termination values, i.e., whether for all states s with $[LimInf = -\infty]$-value p we have $Val(\langle s, n\rangle) > p$ or not.

Corollary 3. *Let $\mathcal{G}$ be an OC-SSG with the set Q of control states. If for every $s \in Q$ the $[LimInf = -\infty]$-value of s is 1 or 0, then Player Max has an optimal strategy for termination in $\mathcal{G}$.*

Proof. The limits of termination values are approached from above, because $Val(\langle s, n\rangle) \geq Val(\langle s, n + 1\rangle)$ for all $s \in Q$ and all n. Thus, 1 is not an accumulation point, and we may apply Theorem 2. □

Note that the class of OC-SSGs satisfying the condition of Corollary 3 involves all OC-SSGs where the finite graph of control states and transition rules is strongly connected, and one of the players is missing. This is because $[LimInf = -\infty]$ is a prefix independent objective, and strong connectivity allows the only player to reach each control state almost surely, thus all control states have the same [$LimInf = -\infty$]-value. By results of [12, Theorem 3.2], such a common value can only be 0 or 1.

One particular case where such "strongly-connected" OC-SSGs can be used is modelling the gambler's ruin: a gambler has an initial positive amount of money, and in each step chooses one of finitely many actions. Each action is associated with a distribution on a finite set of integers. A number from this set is then sampled, and added to the sum of money owned by the gambler (it can be, however, negative), and the process ends only when the wealth becomes ≤ 0. It is easy to encode this in an OC-SSG (see [2]) with only one player, representing the gambler, and the stochastic vertices used to model the probability distributions on outcomes. If we are interested in maximising or minimising the probability of the gambler going bankrupt, we are just asking for maximising or minimising the termination value in the resulting OC-SSG. The case of minimising the probability of going bankrupt was studied as Solvency games in [7]. Many interesting and deep questions arise there, but the question of the existence of optimal strategies for minimising the probability of going bankrupt is easy – there must always be such strategies because the obtained OC-SSG is always finitely branching.

However, the situation where the probability of going bankrupt is maximised is, at least in theory, interesting as well, and here Corollary 3 gives a guarantee that optimal strategies exist again. We are not aware of any result prior to this which would indicate the existence of optimal strategies for Player Max.

5.1. Analysis of Example 2

Consider an arbitrary $n \geq 1$. It is easy to see that $Val(\langle r, n\rangle) = \frac{1}{2}$. Observe that starting in $\langle u, n\rangle$, $\langle s, n+1\rangle$ is visited with probability $\sum_{i=0}^{\infty} 2^{-1-2i} = \frac{2}{3}$, and $\langle s, n-1\rangle$ with probability $\frac{1}{3}$.

Lemma 12. *For the unique strategy σ which does not use transitions $\langle s, n\rangle \to \langle r, n\rangle,\ n \geq 1$, we have $\mathbb{P}^{\sigma}_{\langle s,i\rangle}[W] = 2^{-i}$.*

Proof. Clearly $\mathbb{P}^{\sigma}_{\langle s,0\rangle}[W] = 1 = 2^{-0}$. Further, the assignment $x := \mathbb{P}^{\sigma}_{\langle s,1\rangle}[W]$ is the least non-negative solution of the equation $x = \frac{1}{3} + \frac{2}{3} \cdot x^2$, see, e.g., [13, Theorem 3.4] or [14, Theorem 1], which is $\frac{1}{2}$. Solving the recurrence $\mathbb{P}^{\sigma}_{\langle s,i\rangle}[W] = \frac{1}{3} \cdot \mathbb{P}^{\sigma}_{\langle s,i-1\rangle}[W] + \frac{2}{3} \cdot \mathbb{P}^{\sigma}_{\langle s,i+1\rangle}[W]$, given the initial conditions for $i = 0, 1$, yields $\mathbb{P}^{\sigma}_{\langle s,i\rangle}[W] = 2^{-i}$. $\square$

Lemma 13. $Val(\langle s, 1\rangle) = \frac{3}{4}$.

Proof. First we prove $Val(\langle s, 1\rangle) \geq \frac{3}{4}$. For any n consider the memoryless strategy, σ_n, given by $\sigma_n(\langle s, i\rangle)(\langle u, i\rangle) = 1$ if $i < n$ and $\sigma_n(\langle s, i\rangle)(\langle r, i\rangle) = 1$ if $i \geq n$. Set $p_i := \mathbb{P}^{\sigma_i}_{\langle s,1\rangle}[\text{Reach } \langle s, i\rangle]$. Observe that p_i does not change if we define it using any σ_n with $n \geq i$, and that $1 - p_i = \mathbb{P}^{\sigma_n}_{\langle s,1\rangle}[W \wedge \neg\text{Reach } \langle s, i\rangle]$ for $n \geq i$. Moreover, $p_1 = 1$ and $p_{i+1} := \frac{2}{3} \cdot (p_i + (1 - p_i) \cdot p_{i+1})$. This uniquely determines that $p_i = \frac{2^{i-1}}{2^i - 1}$. Finally, observe that $\mathbb{P}^{\sigma_n}_{\langle s,1\rangle}[W] = (1-p_n)+p_n \cdot \frac{1}{2}$, thus $Val(\langle s, 1\rangle) \geq \lim_{n\to\infty}(1 - p_n) + p_n \cdot \frac{1}{2} = \frac{3}{4}$.

Now we prove that $Val(\langle s, 1\rangle) \leq \frac{3}{4}$ by proving $\mathbb{P}^{\sigma}_{\langle s,1\rangle}[W] \leq \frac{3}{4}$ for all σ. Consider the following probabilities: $p_a := \mathbb{P}^{\sigma}_{\langle s,1\rangle}[W \wedge \neg\text{Reach some } \langle r, j\rangle]$, $p_b := \mathbb{P}^{\sigma}_{\langle s,1\rangle}[W \wedge \text{Reach some } \langle r, j\rangle]$, $p_c := \mathbb{P}^{\sigma}_{\langle s,1\rangle}[\text{Reach some } \langle r, j\rangle]$. Clearly $p_b = \frac{p_c}{2}$. Due to Lemma 12 applied to $i = 1$ we also have that $p_a \leq \frac{1}{2}$. Finally, $p_a + p_c \leq 1$ since the events are disjoint. We conclude that $\mathbb{P}^{\sigma}_{\langle s,1\rangle}[W] = p_a + p_b \leq p_a + \frac{1}{2} \cdot (1 - p_a) = \frac{1}{2} \cdot p_a + \frac{1}{2} \leq \frac{3}{4}$. $\square$

Lemma 14. $Val(\langle s, i\rangle) = \frac{2^i+1}{2^{i+1}}$ *for all $i \geq 0$.*

Proof. The case $i = 0$ is trivial, and $i = 1$ is Lemma 13. Solving the recurrence $Val(\langle s, i\rangle) = \frac{1}{3} \cdot Val(\langle s, i-1\rangle) + \frac{2}{3} \cdot Val(\langle s, i+1\rangle)$, given the initial conditions for $i = 0, 1$, yields $Val(\langle s, i\rangle) = \frac{2^i+1}{2^{i+1}}$. $\square$

In particular, for all $i \geq 1$, $Val(\langle s, i\rangle) > Val(\langle r, i\rangle)$, thus no optimal strategy may use transitions $\langle s, n\rangle \to \langle r, n\rangle,\ n \geq 1$. By Lemma 12, there are no optimal strategies in $\langle s, i\rangle$.

Acknowledgement

The author is grateful to an anonymous reviewer for pointing out various errors and for helping suggestions for the clarity of the text.

References

[1] D.A. Martin, The determinacy of blackwell games, J. Symb. Log. 63 (4) (1998) 1565–1581.
[2] T. Brázdil, V. Brožek, K. Etessami, A. Kučera, D. Wojtczak, One-counter markov decision processes, in: ACM-SIAM SODA, 2010, pp. 863–874, full tech report: CoRR, abs/0904.2511, 2009. http://arxiv.org/abs/0904.2511.
[3] T. Brázdil, V. Brožek, K. Etessami, One-counter simple stochastic games, in: Proc. of FSTTCS'10, 2010, pp. 108–119. http://dx.doi.org/10.4230/LIPIcs.FSTTCS.2010.108.
[4] T. Brázdil, V. Brožek, A. Kučera, J. Obdržálek, Qualitative reachability in stochastic bpa games, Inform. and Comput. 209 (8) (2011) 1160–1183. http://dx.doi.org/10.1016/j.ic.2011.02.002.

[5] T. Brázdil, V. Brožek, K. Etessami, A. Kučera, Approximating the termination value of one-counter mdps and stochastic games, in: ICALP (2), Springer, 2011, pp. 332–343. http://dx.doi.org/10.1007/978-3-642-22012-8_26. Full tech report: CoRR, abs/1104.4978. http://arxiv.org/abs/1104.4978.
[6] T. Brázdil, V. Brožek, A. Kučera, J. Obdržálek, Qualitative reachability in stochastic bpa games, in: Proc. 26th STACS, 2009, pp. 207–218. http://dx.doi.org/10.4230/LIPIcs.STACS.2009.1837.
[7] N. Berger, N. Kapur, L.J. Schulman, V. Vazirani, Solvency games, in: Proc. of FSTTCS'08, 2008, pp. 61–72. http://dx.doi.org/10.4230/LIPIcs.FSTTCS.2008.1741.
[8] V. Brožek, Optimal strategies in infinite-state stochastic reachability games, in: GandALF, in: EPTCS, vol. 54, 2011, pp. 60–73. http://dx.doi.org/10.4204/EPTCS.54.5.
[9] A. Condon, The complexity of stochastic games, Inform. and Comput. 96 (1992) 203–224.
[10] M.L. Puterman, Markov Decision Processes, J. Wiley and Sons, 1994.
[11] T. Brázdil, V. Brožek, V. Forejt, A. Kučera, Reachability in recursive markov decision processes, Inform. and Comput. 206 (5) (2008) 520–537. http://dx.doi.org/10.1016/j.ic.2007.09.002.
[12] H. Gimbert, F. Horn, Solving simple stochastic tail games, in: ACM-SIAM Symposium on Discrete Algorithms, SODA10, 2010, pp. 847–862.
[13] J. Esparza, A. Kučera, R. Mayr, Model-checking probabilistic pushdown automata, in: Proc. of 19th LICS'04, IEEE, 2004, pp. 12–21.
[14] K. Etessami, M. Yannakakis, Recursive markov chains, stochastic grammars, and monotone systems of nonlinear equations, in: Proc. of 22nd STACS'05, Springer, 2005, pp. 340–352.