URL: http://www.elsevier.nl/locate/entcs/volume22.html 33 pages

From Fairness to Chance

Luca de Alfaro

Department of EECS
University of California at Berkeley
Berkeley, CA 94720-1770, USA
dealfaro@eecs.berkeley.edu

Abstract

Fairness is a mathematical abstraction used in the modeling of a wide range of phenomena, including concurrency, scheduling, and probability. In this paper, we study fairness in the context of probabilistic systems, and we introduce *probabilistic fairness,* a novel notion of fairness that is itself defined in terms of probability. The definition of probabilistic fairness makes it invariant with respect to synchronous composition, and facilitates the design of model-checking algorithms for quantitative properties of probabilistic systems. We compare probabilistic fairness with other notions of fairness for probabilistic systems, and we provide algorithms that solve the verification problem for various classes of probabilistic properties on finite-state systems with fairness.

1 Introduction

The use of formal methods for the analysis and verification of systems requires a mathematical model of the system being studied. Many system models include nondeterminism, which enables the representation of interleaving concurrency, and the modeling of schedulers and of partially unknown or unspecified components. Fairness is a constraint on the resolution of the nondeterministic choices, and it has been introduced to represent a multiplicity of related phenomena, such as the progress of threads of computation, general environments, the behavior of probabilistic choice, and the impartiality of arbiters and schedulers. Several notions of fairness have been presented, each tailored to the modeling of some class of phenomena; [20,15,19] present general overviews.

[1] This research was supported in part by the NSF CAREER award CCR-9501708, by the DARPA (NASA Ames) grant NAG2-1214, by the DARPA (Wright-Patterson AFB) grant F33615-98-C-3614, by the ARO MURI grant DAAH-04-96-1-0341, and by the Gigascale Silicon Research Center.

In the context of non-probabilistic systems, a notion of fairness is usually defined by specifying the set φ of system paths that are considered fair, where a "path" is defined as an infinite sequence of states, or as an infinite sequence of alternated states and transitions. The semantics of the system is defined in terms of the subset φ of fair paths only: the paths outside φ are not interpreted as possible system behaviors. For example, consider a system in which at a state s the choice between two alternatives a and b is possible, and assume that this choice is required to be fair. The two alternative might represent the choice of servicing the requests coming from either one of two processes. According to the notion of *strong fairness,* the set φ of fair paths consists of all the paths that choose both a and b infinitely often, whenever s is visited infinitely often. In the example, strong fairness enables the study of the system under the assumption that the scheduling algorithm does not eventually cease to schedule the requests originating from one of the two processes. Other notions of fairness, such as *weak fairness* and *α-fairness,* are specified by providing different definitions for the set φ of fair paths [21,24].

In this paper, we study systems in which both probabilistic and nondeterministic behavior coexist; these systems will be called for brevity *probabilistic systems.* As in other types of systems, fairness in probabilistic systems is also a constraint on the resolution of the nondeterministic choices. However, fairness in probabilistic systems is defined differently than in purely nondeterministic systems, since the apparatus required to deal with both probabilistic and nondeterministic choice is more complex than the one required for nondeterminism alone.

Consider a system where nondeterministic choice coexists with probabilistic one, and assume that at a given state s the nondeterministic choice between two alternatives a and b is possible. Following [16,31], we model the resolution of the nondeterministic choice by a scheduler — that we call *policy* — which at s selects one of a, b. Unlike [22,16,31], however, we consider *randomized* policies rather than deterministic ones, following the customary approach in the theory of Markov decision processes [14], as well as the approach of [29,28]. Each time the system is at s, the (randomized) policy dictates the probabilities of choosing a and b, possibly as a function of the system's past. Since nondeterminism is resolved by the policies, in probabilistic systems fairness is usually expressed by specifying a set Φ of *fair policies.* Again, during the analysis of system properties, only fair policies are considered.

The notions of fairness that have been proposed so far for probabilistic systems are the direct counterparts of notions proposed for purely nondeterministic systems [16,31,17]. Given a notion of fairness for nondeterministic systems specified as a set φ of fair paths, the corresponding notion for probabilistic systems is obtained by defining a policy to be fair iff all the paths arising from the policy (except perhaps for a set of measure 0) belong to φ. Hence, each notion of fairness φ defined as a set of paths gives rise to a corresponding notion $\Phi(\varphi)$ defined on policies. Consider again our system where the alternatives a and b must be fairly chosen at a state s. According to the notion of fairness that corresponds to strong fairness, a policy is fair iff all the paths that arise from it (except perhaps for a set of measure 0) are such that, if s is visited infinitely, both a and b are chosen infinitely often. This is one of the notions of fairness described in [31,17].

In this paper we introduce a novel notion of fairness, called *probabilistic fairness.* Unlike previous notions of fairness, probabilistic fairness is a *local* notion of fairness: it is expressed directly in terms of the behavior of the policies at the various states, and it has no counterpart as a requirement on paths. According to probabilistic fairness, a policy is fair iff there is an $\varepsilon > 0$ such that all fair alternatives are chosen with probability at least ε by the policy. In our previous example, a policy is fair iff the probability with which the alternatives a and b are chosen at s is bounded below by $\varepsilon > 0$. We note that, while ε can vary from one policy to the other, it must be constant for each policy, rather than dependent on the state of the system or on its past history. Probabilistic fairness entails several benefits over previous notions of fairness for probabilistic systems. These benefits are both semantical, concerning the modeling of systems, and algorithmic, concerning the algorithms for system verification.

1.1 Semantical benefits

Probabilistic fairness offers three semantical benefits: it provides a simple way of representing probabilistic choice while abstracting from the numerical values of probability; it exhibits a simple form of invariance with respect to synchronous composition; and it enables the representation of threads of computation in which the ratios between the speeds of computation is unknown, but bounded.

Representation of probabilistic choice

Representing the qualitative properties of probabilistic choice, while abstracting from the values of the transition probabilities, has two purposes. First, it enables the modeling of probabilistic behavior in the cases in which the probabilities of some alternatives are not known, except for the fact that they are positive. This can be useful whenever the probabilities have not been measured accurately, or when the portion of the system giving rise to the probabilistic behavior has not been designed yet. Second, probability provides a reference model for schedulers that are completely impartial with respect to the incoming requests. Indeed, several fairness notions that have been introduced to model schedulers, such as *strong fairness, event and process fairness,* and *interaction fairness,* exclude the set of paths that have 0 probability under the purely probabilistic scheduling of the steps, events, or process interactions that occur along the paths [15,19].

The problem of finding a notion of fairness that corresponds to the qualitative properties of probabilistic choice was considered already in [22]. With respect to the verification of linear-time temporal logic properties (and more generally, membership in ω-regular languages), the problem was settled with the introduction of *α-fairness,* a fairly complex notion of fairness [24]. Probabilistic fairness offers a straightforward solution to this problem, since it is defined directly in terms of probabilities. While the adoption of probabilistic fairness seems to contradict the goal of eliminating probability from the system model, we will show that the model-checking algorithms for probabilistic fairness do not incur any additional complexity due to its probability-based definition.

Synchronous composition

Synchronous composition is a basic step in the modeling and verification of systems: it can be used to construct the complete system from smaller component systems, and the synchronous composition of the system with an automaton derived from the specification is at the heart of several verification algorithms [31,23,5,6]. Probabilistic fairness exhibits a simple invariance property with respect to synchronous composition.

> *If two systems $\mathcal{P}$ and $\mathcal{Q}$ are non-interacting, and if a policy $\pi_{\mathcal{P}}$ for $\mathcal{P}$ is probabilistically fair, then the policy $\pi_{\mathcal{P}\|\mathcal{Q}}$ obtained by projecting $\pi_{\mathcal{P}}$ onto the synchronous composition $\mathcal{P}\|\mathcal{Q}$ of $\mathcal{P}$ and $\mathcal{Q}$ is also probabilistically fair.*

This invariance property states that the fairness of a policy for a given system does not depend on whether the system is considered in isolation, or together with other non-interacting systems. While some notions of fairness satisfy the above invariance (notably α-fairness), this is not the case for some of the most common notions, such as weak and strong fairness [21]. The fact that probabilistic fairness satisfies this invariance property is a direct consequence of the local nature of its definition.

Progress of independent threads of computation

Probabilistic fairness enables the modeling of the progress of independent threads of computation, in which the ratio between the speeds of computation is unknown, but bounded. In the context of timed probabilistic systems, probabilistic fairness also enables the modeling of transitions having finite, but unknown, average delay, as discussed in detail in [11]. In these respects, probabilistic fairness is related to *finitary fairness,* a (non-probabilistic) notion of fairness proposed for reasoning about distributed algorithms [1].

1.2 Algorithmic benefits

The solution of many verification problems for probabilistic systems consists in determining a policy that is optimal (or pessimal) with respect to the property of interest, and in checking whether the property holds for this optimal or pessimal policy. When fairness is introduced in the system model, the optimal (or pessimal) policy must be chosen from the set of fair policies, rather than from the set of all policies. However, the optimization methods available from the theory of Markov decision processes compute the optimal and pessimal policies in the set of all policies, and they cannot be easily adapted to conduct the optimization in the smaller set of fair policies [14,3]. To show that the (unconstrained) solution of an optimization problem can be used in the verification of fair probabilistic systems, we have to show that the optimal values of the quantities of interest can be realized or at least approximated by a set of fair policies, following the idea of [22,17].

The local definition of probabilistic fairness facilitates the construction of such approximating policies, by ensuring that the convex combination of a generic policy and a fair policy is a fair policy. To illustrate this point, assume that the policies are *memoryless,* i.e. that the probabilities with which the alternatives are chosen depend only on the current system state, and denote by $\pi(s)(a)$ the probability with which alternative a is selected at state s. Given a generic policy π_g and a fair policy π_f, their convex combination $\pi[x]$ for $0 \leq x \leq 1$ is defined by

$$\pi[x](s)(a) = (1-x)\,\pi_g(s)(a) + x\,\pi_f(s)(a)$$

for all states s and all alternatives a. For $0 < x \leq 1$, policy $\pi[x]$ is fair, and for $x = 0$ it coincides with π_g. Consider a function h from policies to real numbers; the value $h(\pi)$ can represent for example a performance index of the system under policy π. To show that the value of the performance index corresponding to π_g can be approximated by fair policies, it suffices to prove that $\lim_{x \to 0} h(\pi[x]) = h(\pi[0]) = h(\pi_g)$. Often, this proof can be carried out using standard methods from calculus and linear algebra. With minor variations, this approach to the construction of approximating policies will be used to justify all the verification algorithms presented in the paper.

1.3 Paper outline

After providing a standard definition for probabilistic systems, we introduce three notions of fairness. The first one is probabilistic fairness; the second one is *unbounded fairness,* a weaker variant of probabilistic fairness that shares some of its properties, and the third one is *path fairness,* which is essentially the notion studied in [31,17]. We show that probabilistic and unbounded fairness, unlike path fairness, are invariant with respect to synchronous composition. We then compare the three notions of fairness with respect to three classes of properties:

Maximum acceptance probability. This class of properties concerns the maximum probability with which a path satisfies the Rabin acceptance condition of an ω-automaton, and it is related to the maximum probability of satisfying linear-time temporal logic formulas.

Minimum reachability cost. This class of properties concerns the minimum expected cost for reaching a subset of target states. The cost can represent various quantities of interest, such as the amount of time elapsed before the target is reached.

Maximum long-run average outcome. This class of properties is related to the long-run average outcome of system tasks, such as the request for a resource, or the sending of a message. Long-run average properties enable the specification of many classical performance and reliability indices [10].

We show that probabilistic fairness is equivalent to path fairness with respect to the maximum acceptance probability and the long-run average outcome classes of properties, and it is equivalent to unbounded fairness with respect to the minimum reachability cost class. Finally, for each of these notions of fairness and classes of properties we present model-checking algorithms that can be used to solve the verification problem on finite-state probabilistic systems.

2 Probabilistic Systems and Fairness

Our model for probabilistic systems is based on *Markov decision processes* (MDPs). An MDP is a generalization of a Markov chain in which a set of possible actions is associated with each state. To each state-action pair corresponds a probability distribution on the states, which is used to select the successor state [14]. Markov decision processes are closely related to the *probabilistic automata* of [25], the *concurrent Markov chains* of [31], and the *simple probabilistic automata* of [29,28].

Given a countable set C we denote by $\mathcal{D}(C)$ the set of probability distributions over C, i.e. the set of functions $f : C \mapsto [0,1]$ such that $\sum_{x \in C} f(x) = 1$. An MDP $\mathcal{P} = (S, \mathit{Acts}, A, p)$ consists of the following components:

(i) A set S of states.

(ii) A set Acts of actions.

(iii) A function $A : S \mapsto 2^{\mathit{Acts}}$, which associates with each $s \in S$ a finite set $A(s) \subseteq \mathit{Acts}$ of actions available at s.

(iv) A function $p : S \times \mathit{Acts} \mapsto \mathcal{D}(S)$, which associates with each $s, t \in S$ and $a \in A(s)$ the probability $p(s,a)(t)$ of a transition from s to t when action a is selected.

We will often associate with an MDP additional labelings to represent quantities of interest; the labelings will be simply added to the list of components.

A *path* of an MDP is an infinite sequence $\theta : s_0, a_0, s_1, a_1, \ldots$ of alternating states and actions, such that $s_i \in S$, $a_i \in A(s_i)$ and $p(s_i, a_i)(s_{i+1}) > 0$ for all $i \geq 0$. For $i \geq 0$, the sequence is constructed by iterating a two-phase selection process. First, an action $a_i \in A(s_i)$ is selected nondeterministically; second, the successor state s_{i+1} is chosen according to the probability distribution

$p(s_i, a)$. Given a path $\theta : s_0, a_0, s_1, a_1, \ldots$ and $k \geq 0$, we denote by $X_k(\theta)$, $Y_k(\theta)$ its k-th state s_k and its k-th action a_k, respectively.

For every state $s \in S$, we denote by Θ_s the set of (infinite) paths having s as initial state, and we denote by Σ_s the set of finite path prefixes having s as initial state. The set of all paths is $\Theta = \bigcup_{s \in S} \Theta_s$. Given two paths (or path prefixes) θ_1 and θ_2, we denote by $\theta_1 \preceq \theta_2$ the fact that θ_1 is a prefix of θ_2. Following the classical definition of [18], we let $\mathcal{B}_s \subseteq 2^{\Theta_s}$ be the σ-algebra of *measurable* subsets of Θ_s, defined as the smallest algebra that contains all the *cylinder sets* $\{\theta \in \Theta_s \mid \sigma \preceq \theta\}$, for σ that ranges over Σ_s, and that is closed under complementation and countable unions (and hence also countable intersections). The elements of $\mathcal{B}_s$ are called *events,* and they are the measurable sets of paths to which we will associate a probability.

2.1 Policies

To assign a probability to the events in $\mathcal{B}_s$, for all $s \in S$, we need to specify the criteria with which the actions are chosen. To this end, we use the concept of *policy* [14], closely related to the *schedulers* of [31] and to the *adversaries* of [29,28]. Denoting with S^+ the set of non-empty finite sequences of states, a policy π is a mapping $\pi : S^+ \mapsto \mathcal{D}(Acts)$, which associates with each sequence of states $s_0, s_1, \ldots, s_n \in S^+$ and each $a \in A(s_n)$ the probability $\pi(s_0, s_1, \ldots, s_n)(a)$ of choosing a after following the sequence of states $s_0, s_1, \ldots, s_n$. We require that $\pi(s_0, s_1, \ldots, s_n)(a) > 0$ implies $a \in A(s_n)$: a policy can choose only among the actions that are available at the state where the choice is made. We indicate with Π the set of all policies. According to this definition, policies are randomized, differently from the *schedulers* of [31,23], which are deterministic. The consideration of randomized policies is fundamental to the further developments of this paper. From these definitions, the probability of following a finite path prefix $s_0, a_0, s_1, a_1, \ldots, s_n$ under policy $\pi \in \Pi$ is given by

$$\prod_{i=0}^{n-1} p(s_i, a_i)(s_{i+1})\ \pi(s_0, \ldots, s_i)(a_i) .$$

These probabilities for prefixes give rise to a unique probability measure on $\mathcal{B}_s$. For $\mathcal{A} \in \bigcup_{s \in S} \mathcal{B}_s$, we write $\Pr_s^\pi(\mathcal{A})$ to denote the probability of event $\mathcal{A} \cap \mathcal{B}_s$ starting from the initial state $s \in S$ under policy π. For example, given a set $R \subseteq S$ of states, we denote by

$$(\Diamond R) = \{\theta \in \Theta \mid \exists k \geq 0 \,.\, X_k(\theta) \in R\}$$

the event of reaching R. The probability of reaching R starting from state s under policy π is then $\Pr_s^\pi(\Diamond R)$. Similarly, for all $s \in S$, if $f : \Theta_s \mapsto \mathbb{R}$ is a measurable function, we denote by $\mathrm{E}_s^\pi\{f\}$ the expectation of f from state s under policy π. For example, given a set $R \subseteq S$, for all paths $\theta \in \Theta$ we denote by

$$T_R(\theta) = \min\{k \mid X_k(\theta) \in R\}$$

the first-passage time of θ in R, with the convention that $\min \emptyset = \infty$. For all $s \in S$ the function $T_R : \Theta_s \mapsto \mathbb{R}$ is measurable, and the expected first-passage time in R from $s \in S$ under policy π is written as $\mathrm{E}_s^\pi\{T_R\}$. Note that we omitted the argument θ of the random function $T_R(\theta)$: for conciseness, here and in the following we omit the generic path θ that is the argument of random functions whenever we take expectations or probability measures.

2.2 Fairness

Given an MDP $\mathcal{P} = (S, Acts, A, p)$, a *fairness constraint* $\mathcal{F}$ for $\mathcal{P}$ is a mapping $\mathcal{F} : S \mapsto 2^{Acts}$ that associates with each $s \in S$ a subset $\mathcal{F}(s) \subseteq A(s)$ of *fair actions* at s. The intended meaning is that the choice at s among actions in $\mathcal{F}(s)$ should be "fair." The various notions of fairness differ in the way in which this "fairness" is defined. We denote by $SAPairs(\mathcal{P}) = \{(s, a) \mid s \in S \wedge a \in A(s)\}$ the set of *state-action pairs* of the MDP [14]. Given a path θ, we denote by

$$InfS(\theta) = \{s \in S \mid \overset{\infty}{\exists}\, k \,.\, X_k(\theta) = s\}$$
$$InfSA(\theta) = \big\{(s, a) \in SAPairs(\mathcal{P}) \mid \overset{\infty}{\exists}\, k \,.\, (X_k(\theta), Y_k(\theta)) = (s, a)\big\}$$

the sets of states and of state-action pairs that are repeated infinitely often along θ, where the notation $\overset{\infty}{\exists}\, k$ is an abbreviation for "there are infinitely many distinct values for k". For each policy and each initial state $s \in S$, the functions $InfS : \Theta_s \mapsto 2^S$ and $InfSA : \Theta_s \mapsto 2^{(S \times Acts)}$ are measurable.

Path fairness

Path fairness essentially coincides with the fairness considered in [31], and is called *weak fairness* in [17]. We say that a policy π is *path-fair* if, for all initial states, the paths that arise under π satisfy with probability 1 the following condition: *whenever a path visits infinitely often a state t, each action in $\mathcal{F}(t)$ is chosen infinitely often at t.* More precisely, π is path fair with respect to constraint $\mathcal{F}$ if, for all initial states $s \in S$ and all state-action pairs $(t, a) \in SAPairs(\mathcal{P})$ with $a \in \mathcal{F}(t)$,

$$\Pr_s^\pi\big(t \in InfS \ \ implies \ \ (t, a) \in InfSA\big) = 1 \ .$$

We call this notion of fairness *path fairness* because the fairness of a policy is established on the basis of the paths that arise under the policy. In contrast, our next notions of fairness refer directly to the policies.

Probabilistic fairness and unbounded fairness

Probabilistic fairness is a local notion of fairness that refers directly to the behavior of the policies a the various system states. Denote by S^* the set of finite (and possibly empty) sequences of states. A policy π is *probabilistically fair* with respect to the constraint $\mathcal{F}$ if there is an $\varepsilon > 0$ such that $\pi(\bar{s}, s)(a) >$

ε for all $\bar{s} \in S^*$, all $s \in S$ and all $a \in \mathcal{F}(s)$. In other words, a policy π is probabilistically fair with respect to $\mathcal{F}$ if there is a lower bound $\varepsilon > 0$ for the probability of choosing a fair action, throughout the system's behavior [8,11]. This requirement can also be written as:

$$\inf\{\pi(\bar{s}, s)(a) \mid \bar{s} \in S^* \wedge s \in S \wedge a \in \mathcal{F}(s)\} > 0 .$$

In the definition of probabilistic fairness, the bound ε can depend on the policy π, but it cannot depend on the past sequence $\bar{s}$ of states. If ε could depend on $\bar{s}$, then probabilistic fairness would reduce to a very weak notion of fairness, which we call *unbounded fairness.* A policy π is *unboundedly fair* with respect to the constraint $\mathcal{F}$ if we have

$$\pi(\bar{s}, s)(a) > 0$$

for all $\bar{s} \in S^*$, all $s \in S$, and all $a \in \mathcal{F}(s)$.

3 Relations Among Fairness Notions

Given an MDP $\mathcal{P}$ and a fairness constraint $\mathcal{F}$ for $\mathcal{P}$, we denote by $\mathit{PathF}(\mathcal{P}, \mathcal{F})$, $\mathit{ProbF}(\mathcal{P}, \mathcal{F})$, and $\mathit{UnbF}(\mathcal{P}, \mathcal{F})$ the sets of policies that are fair according to path, probabilistic, and unbounded fairness, respectively. We also indicate with $\mathit{NoF}(\mathcal{P}) = \Pi$ the set of all policies, corresponding to the notion of no fairness. In the following, we omit the arguments $\mathcal{P}$ and $\mathcal{F}$ whenever they can be univocally understood from the context. The following preliminary proposition characterizes the hierarchy between these three fairness notions.

Proposition 1 *The following assertions hold:*

(i) *For all MDPs $\mathcal{P}$ and all fairness constraints $\mathcal{F}$, we have $\mathit{ProbF}(\mathcal{P}, \mathcal{F}) \subseteq \mathit{PathF}(\mathcal{P}, \mathcal{F})$, and $\mathit{ProbF}(\mathcal{P}, \mathcal{F}) \subseteq \mathit{UnbF}(\mathcal{P}, \mathcal{F})$.*

(ii) *Unbounded fairness and path fairness are incomparable:*
   (a) *There is an MDP $\mathcal{P}$ and a fairness constraint $\mathcal{F}$ such that $\mathit{PathF}(\mathcal{P}, \mathcal{F}) \not\subseteq \mathit{UnbF}(\mathcal{P}, \mathcal{F})$.*
   (b) *There is an MDP $\mathcal{P}$ and a fairness constraint $\mathcal{F}$ such that $\mathit{UnbF}(\mathcal{P}, \mathcal{F}) \not\subseteq \mathit{PathF}(\mathcal{P}, \mathcal{F})$.*

Proof. Assertion (i) follows immediately from the definitions of fairness.

The MDP $\mathcal{P}$ of Figure 1 with its fairness constraint $\mathcal{F}_{\mathcal{P}}$ is a witness for assertion (a). In fact, consider the policy π defined for all $k > 0$ by $\pi(s^k)(a) = 1$ if k is even, and $\pi(s^k)(a) = 0$ if k is odd (where s^k is the sequence consisting of k states s). Then $\pi \in \mathit{PathF}(\mathcal{P}, \mathcal{F}_{\mathcal{P}})$ and $\pi \notin \mathit{UnbF}(\mathcal{P}, \mathcal{F}_{\mathcal{P}})$.

The MDP $\mathcal{Q}$ of Figure 1 with its fairness constraint $\mathcal{F}_{\mathcal{Q}}$ is a witness for assertion (b). In fact, consider the policy π defined for all $k > 0$ by $\pi(s^k)(a) = 2^{-1/2^k}$. From this definition follows immediately that $\pi \in \mathit{UnbF}(\mathcal{Q}, \mathcal{F}_{\mathcal{Q}})$. To see that $\pi \notin \mathit{PathF}(\mathcal{Q}, \mathcal{F}_{\mathcal{Q}})$, it suffices to note that under policy π, a path that starts from s is confined to s (and takes only action a) with probability

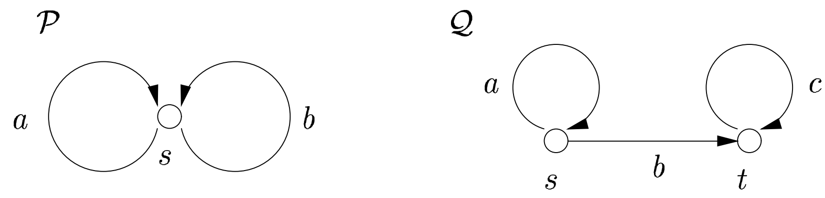

Fig. 1. Two MDPs $\mathcal{P}$ and $\mathcal{Q}$. The MDPs are deterministic, i.e. for each state and action, there is only one successor state, indicated in the diagram by a directed edge labeled with the action. The MDP $\mathcal{P} = (S, \mathit{Acts}, A, p)$ is defined by $S = \{s\}$, $\mathit{Acts} = \{a, b\}$, $A(s) = \{a, b\}$, and $p(s,a)(s) = p(s,b)(t) = 1$. The MDP $\mathcal{P}$ has an associated fairness constraint $\mathcal{F}_{\mathcal{P}}$ defined by $\mathcal{F}_{\mathcal{P}}(s) = \{a, b\}$. The MDP $\mathcal{Q}$ is similarly defined, has it an associated fairness constraint $\mathcal{F}_{\mathcal{Q}}$ defined by $\mathcal{F}_{\mathcal{Q}}(s) = \{b\}$, and $\mathcal{F}_{\mathcal{Q}}(t) = \emptyset$.

1/2. ∎ ∎

3.1 Fairness and synchronous composition

Path fairness does not posses the same invariance properties of probabilistic and unbounded fairness with respect to synchronous composition. In fact, it is possible that a policy that is path fair for an MDP when considered in isolation may not be path fair when the same MDP is considered composed synchronously with a non-interacting automaton. Since the MDP and the automaton do not interact, this means that the notion of path fairness is fragile, and the path fairness of a policy depends on the "environment" at large in which the system is studied. This undesirable characteristic is not shared by either probabilistic or unbounded fairness. The synchronous composition of an MDP and an automaton is important in verification, and the notion of α-fairness has been in part proposed to overcome this limitation of path fairness [24].

There are many definitions for synchronous composition, depending on the methods chosen for synchronizing the systems being composed. To emphasize that the phenomenon is independent of the particular definition adopted, we focus here on what is perhaps the simplest form of synchronous composition: the synchronous product between an MDP and a deterministic finite-state automaton with singleton input alphabet, where the MDP and the automaton are non-interacting. Even though this type of synchronous product is thoroughly trivial, it suffices to expose the different behavior of the various fairness notions.

Given an MDP $\mathcal{P} = (S, \mathit{Acts}, A, p)$ and an automaton $\mathcal{Q} = (T, \delta)$ with $\delta : T \mapsto T$, we define their synchronous product to be the MDP $\mathcal{P} \| \mathcal{Q} = (S \times T, \mathit{Acts}, B, q)$, where:

- for all $s \in S$ and $t \in T$, we have $B(s,t) = A(s)$.
- for all $s, s' \in S$, all $t, t' \in T$, and all $a \in A(s)$, the probability $p((s,t),a)(s',t')$ is equal to $p(s,a)(s')$ if $t' = \delta(t)$, and is equal to 0 otherwise.

Corresponding to a fairness constraint $\mathcal{F}_\mathcal{P}$ for $\mathcal{P}$, we define the fairness constraint $\mathcal{F}_{\mathcal{P}\|\mathcal{Q}}$ for $\mathcal{P}\|\mathcal{Q}$ by letting $\mathcal{F}_{\mathcal{P}\|\mathcal{Q}}(s,t) = \mathcal{F}_\mathcal{P}(s)$ for all $s \in S$ and $t \in T$. Corresponding to a policy $\pi_\mathcal{P}$ for $\mathcal{P}$, we define the policy $\pi_{\mathcal{P}\|\mathcal{Q}}$ for $\mathcal{P}\|\mathcal{Q}$ by letting

$$\pi_{\mathcal{P}\|\mathcal{Q}}\big((s_0,t_0),(s_1,t_1),\ldots(s_n,t_n)\big) = \pi_\mathcal{P}(s_0,s_1,\ldots,s_n)$$

for all $n > 0$, all $s_1, s_2, \ldots, s_n \in S^+$, and all $t_1, t_2, \ldots, t_n \in T^+$. With this notation, we can finally state the following theorem.

Theorem 1 *The following assertions hold:*

(i) *There is an MDP $\mathcal{P}$ with a fairness constraint $\mathcal{F}_\mathcal{P}$, there is a deterministic automaton $\mathcal{Q}$ with singleton alphabet, and there is a policy $\pi_\mathcal{P} \in \mathit{PathF}(\mathcal{P},\mathcal{F}_\mathcal{P})$ such that $\pi_{\mathcal{P}\|\mathcal{Q}} \notin \mathit{PathF}(\mathcal{P}\|\mathcal{Q},\mathcal{F}_{\mathcal{P}\|\mathcal{Q}})$.*

(ii) *Consider a fairness notion $\Phi \in \{\mathit{ProbF}, \mathit{UnbF}\}$. For all MDPs $\mathcal{P}$ with fairness constraint $\mathcal{F}_\mathcal{P}$, for all deterministic automata $\mathcal{Q}$ with singleton alphabet, and for all policies $\pi_\mathcal{P} \in \Phi(\mathcal{P},\mathcal{F}_\mathcal{P})$, we have $\pi_{\mathcal{P}\|\mathcal{Q}} \in \Phi(\mathcal{P}\|\mathcal{Q},\mathcal{F}_{\mathcal{P}\|\mathcal{Q}})$.*

Proof. For the first assertion, consider the MDP $\mathcal{P}$ and the automaton $\mathcal{Q}$ of Figure 2. The portion of the synchronous product $\mathcal{P}\|\mathcal{Q}$ that is reachable from the state (s_1, t_1) is also depicted in the figure. Consider the policy $\pi_\mathcal{P}$ defined for all $\bar{s} \in S^*$ by

$$\pi_\mathcal{P}(\bar{s}, s_1)(a) = \begin{cases} 1 & \text{if there are an even number of } s_1 \text{ in } \bar{s}; \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to check that $\pi_\mathcal{P} \in \mathit{PathF}(\mathcal{P},\mathcal{F}_\mathcal{P})$, while $\pi_{\mathcal{P}\|\mathcal{Q}} \notin \mathit{PathF}(\mathcal{P}\|\mathcal{Q},\mathcal{F}_{\mathcal{P}\|\mathcal{Q}})$.

The second assertion follows easily from the definition of probabilistic and unbounded fairness. ∎ ∎

3.2 Fairness and probabilistic properties

We analyze the relationship between the three fairness notions with respect to three classes of properties: *acceptance probability, reachability cost,* and *long-run average outcome.* In the following, we consider an MDP $\mathcal{P} = (S, \mathit{Acts}, A, p)$ together with a fairness constraint $\mathcal{F} : S \mapsto 2^{\mathit{Acts}}$, unless otherwise specified.

Acceptance probability

The first class of properties we consider concerns the maximum probability with which a path satisfies a *Rabin acceptance constraint.* This maximum probability is closely related to the the maximum probability of satisfying a linear-time temporal logic formula [9]. A *Rabin acceptance condition* is a set of pairs $\mathcal{A} = \{(Q_1^p, Q_1^r), \ldots, (Q_m^p, Q_m^r)\}$, where $Q_i^p, Q_i^r \subseteq S$ for all $1 \le i \le m$ [27,30]. A path θ of the MDP satisfies $\mathcal{A}$, written $\theta \models \mathcal{A}$, iff there is $1 \le i \le m$ such that

$$\mathit{InfS}(\theta) \subseteq Q_i^p \ ,$$

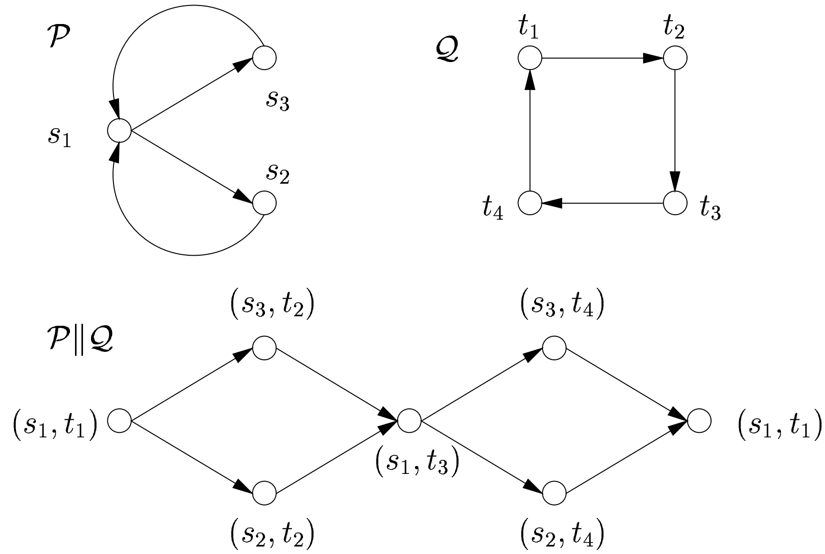

Fig. 2. An MDP $\mathcal{P}$, and an automaton $\mathcal{Q}$. The MDP is deterministic, and has an associated fairness constraint $\mathcal{F}$ defined by $\mathcal{F}(s_1) = \{a, b\}$. The automaton simply takes the only possible transition at every step. The portion of the synchronous product $\mathcal{P}\|\mathcal{Q}$ reachable from the state (s_1, t_1) is also depicted.

$$\mathit{InfS}(\theta) \cap Q_i^r \neq \emptyset \ .$$

Given a state $s \in S$, an acceptance condition $\mathcal{A}$, and a notion $\Phi \in \{\mathit{NoF}, \mathit{PathF}, \mathit{ProbF}, \mathit{UnbF}\}$ of fairness, the maximum acceptance probability $\Pr_s^+(\Phi, \mathcal{A})$ is defined as

$$\Pr_s^+(\Phi, \mathcal{A}) = \sup_{\pi \in \Phi} \Pr_s^\pi(\theta \models \mathcal{A}) \ . \tag{1}$$

Reachability cost

The second class of properties we consider concerns the expected cost of reaching a set of target states in the MDP. To define this quantity, let $c : S \times \mathit{Acts} \mapsto \mathbb{R}^+$ be a cost function that associates with each $s \in S$ and $a \in A(s)$ a cost $c(s, a) > 0$. For all initial states $s \in S$, target subsets $R \subseteq S$, and policies π, the expected cost of reaching R from s under policy π is given by

$$v_s^\pi(c, R) = \mathrm{E}_s^\pi \Big\{ \sum_{k=0}^{T_R - 1} c(X_k, Y_k) \Big\} \ , \tag{2}$$

where $T_R = \min\{k \mid X_k \in R\}$ is the first-passage time in R, with the convention that $\min \emptyset = \infty$. For a notion $\Phi \in \{\mathit{NoF}, \mathit{PathF}, \mathit{ProbF}, \mathit{UnbF}\}$ of fairness, the minimum expected reachability cost from s to R is then defined as

$$v_s^-(\Phi, c, R) = \inf_{\pi \in \Phi} v_s^\pi(c, R) \ . \tag{3}$$

Note that $v_s^-(\Phi, c, R)$ is infinite if R cannot be reached with probability 1 from s. If the cost $c(s, a)$ represents the time (or the expected time) spent at s when action $a \in A(s)$ is selected, then the quantity $v_s^-(\Phi, c, R)$ is equal to the minimum expected time from s to R. It is possible to consider also the more general case of non-negative costs, as done in [8], at the price of some mathematical complications.

Long-run average outcome

Long-run average properties are related to the average behavior of the system, measured over an interval of time whose length diverges to infinity [8,10]. The specification of these properties is based on the notion of *experiment*. An experiment is a finite portion of a path, which corresponds to a task of interest for the performance or reliability analysis of the system. An example of experiment consists in a request to access a shared resource, followed either by a grant or a rejection. With each experiment is associated a numerical value called the *outcome* of the experiment. The long-run average outcome of the experiment is simply the average value of such outcomes, measured over a period of time whose length diverges to infinity. In the previous example, if we associate outcome 0 with the experiments that end with a rejection, and outcome 1 with those that end with a grant, then the long-run average outcome of the experiment is equal to the long-run fraction of requests that are granted. The long-run average outcome is defined on the basis of two functions $R, W : S \times \mathit{Acts} \mapsto \mathbb{R}^+$ that associate with each $s \in S$ and $a \in A(s)$ the following quantities:

- the average outcome $R(s, a) \geq 0$ obtained when selecting action a at s;
- a completion rate $W(s, a) > 0$, equal to the probability of completing the experiment when selecting action a at s.

The restriction that W be non-zero is artificial, and in fact [8,10] considers the general case of non-negative W (and arbitrary R). We adopted this restriction because it leads to a considerably simpler mathematical treatment, while preserving the essence of the argument. Given $s \in S$, the functions R, W, and a policy π, the *expected long-run average outcome* $H_s^\pi(R, W)$ is defined as

$$H_s^\pi(R, W) = \mathrm{E}_s^\pi \left\{ \limsup_{n \to \infty} \frac{\sum_{k=0}^{n-1} R(X_k, Y_k)}{\sum_{k=0}^{n-1} W(X_k, Y_k)} \right\} . \tag{4}$$

For $n < \infty$, the numerator of (4) represents the total outcome obtained during the first n steps of the path, and the denominator represents the number of experiments performed. The limit for $n \to \infty$ of this ratio corresponds therefore to the average outcome per experiment along a path, and $H_s^\pi(R, W)$ is the expected value of this average outcome, computed considering all paths from s. Given $s \in S$, the functions R, W, and a notion of fairness $\Phi \in \{\mathit{NoF}, \mathit{PathF}, \mathit{ProbF}, \mathit{UnbF}\}$, we finally define the *maximum long-run average* *outcome* by:

$$H_s(\Phi, R, W) = \sup_{\pi \in \Phi} H_s^{\pi}(R, W) \ . \tag{5}$$

The quantity defined in (4) is related to the average reward of semi-Markov decision processes [26,3]. However, in the classical definition the limit and expectation are exchanged, and the expectation is distributed in two expectations, one above and one below the fraction line. The difference between the two definitions is discussed in [8].

3.2.1 Preview of the results

The behavior of the different notions of fairness with respect to the three above classes of properties are summarized by the following theorem.

Theorem 2 *For all states s, and for all $\mathcal{A}$, resp. all c, R, resp. all R, W, and for a general finite-state MDP with a fairness constraint, the following relations hold:*

(i) *Acceptance probability:*

$$\begin{aligned}\mathrm{Pr}_s^+(\mathit{NoF}, \mathcal{A}) &= \mathrm{Pr}_s^+(\mathit{UnbF}, \mathcal{A}) \\ &\geq \mathrm{Pr}_s^+(\mathit{PathF}, \mathcal{A}) = \mathrm{Pr}_s^+(\mathit{ProbF}, \mathcal{A})\end{aligned}$$

(ii) *Reachability cost:*

$$\begin{aligned}v_s^-(\mathit{NoF}, c, R) &= v_s^-(\mathit{PathF}, c, R) \\ &\leq v_s^-(\mathit{UnbF}, c, R) = v_s^-(\mathit{ProbF}, c, R)\end{aligned}$$

(iii) *Long-run average outcome:*

$$\begin{aligned}H_s(\mathit{NoF}, R, W) &= H_s(\mathit{UnbF}, R, W) \\ &\geq H_s(\mathit{PathF}, R, W) = H_s(\mathit{ProbF}, R, W)\end{aligned}$$

Moreover, the inequalities in the above relations cannot in general be replaced by equalities.

The above theorem tells us that probabilistic fairness sides with path fairness in finite-state systems, except for the case of reachability cost. This theorem also supports our claim that a probabilistic treatment of fairness is not any harder than a traditional one, except for the case of minimum expected reachability cost — and even in this case, we will show that working with probabilistic rather than path fairness entails only minor additional complications.

The simplicity of Theorem 2 is due in part to the fact that the quantities in (1), (3) and (5) have been defined using sup and inf, and we have not distinguished between the cases in which the suprema and infima can be achieved or not (i.e. whether sup and inf can be replaced with max and min).

This distinction would have blurred the insight provided by the theorem, and would have required the use of more complex model-checking algorithms. Algorithms that distinguish between these two cases for path fairness and Rabin acceptance conditions have been presented in [17].

In the remainder of the paper, we provide model-checking algorithms for all the combinations of the three notions of fairness and the three classes of properties. The equalities in Theorem 2 follow from the fact that the notions of fairness share the same model-checking algorithms. The fact that the inequalities cannot be in general replaced by equalities is shown by providing counterexamples.

4 Tools for Fairness

In this section, we present some results on MDPs that will be used in the construction and justification of the model-checking algorithms.

4.1 End components

Given an MDP $\mathcal{P} = (S, \mathit{Acts}, A, p)$, a *sub-MDP* is a pair (C, D), where $C \subseteq S$ is a subset of states and $D : S \mapsto 2^{\mathit{Acts}}$ is an *action assignment,* i.e. a function that associates to each $s \in S$ a subset $D(s) \subseteq A(s)$ of actions. The sub-MDP corresponds thus to a subset of states and actions of the original MDP. With each sub-MDP (C, D) we associate its set of state-action pairs

$$\mathit{SAPairs}(C, D) = \{(s, a) \in \mathit{SAPairs}(\mathcal{P}) \mid s \in C \wedge a \in D(s)\} .$$

Similarly, with each state-action set $\xi \subseteq \mathit{SAPairs}(\mathcal{P})$ we associate a sub-MDP $(C, D) = \mathit{SAPairs}^{-1}(\xi)$, defined by

$$C = \{s \in S \mid \exists a \in \mathit{Acts} \,.\, (s, a) \in \xi\}$$

and, for all $s \in S$, by

$$D(s) = \{a \in \mathit{Acts} \mid (s, a) \in \xi\} .$$

We say that a sub-MDP (C, D) is contained in a sub-MDP (C', D') if

$$\mathit{SAPairs}(C, D) \subseteq \mathit{SAPairs}(C', D') .$$

We say that a sub-MDP (C, D) is an *end component* (abbreviated by EC) if the following conditions hold:

- *Closure:* for all $s \in C$, all $a \in D(s)$, and all $t \in S$, if $p(s, a)(t) > 0$ then $t \in C$.
- *Connectivity:* Let $E = \big\{(s, t) \in C \times C \;\big|\; \exists a \in D(s) \,.\, p(s, a)(t) > 0\big\}$. The graph (C, E) is strongly connected.

Given a subset $U \subseteq S$ of states, we say that an EC (C, D) is *maximal* in U if $C \subseteq U$, and if there is no other EC (C', D') with $C' \subseteq U$ that properly contains (C, D). We denote by $\mathit{Mec}(U)$ the set of maximal ECs in U; this set can be computed in time polynomial in the size of the MDP using simple graph algorithms. In a purely probabilistic system, fair end components correspond to the closed recurrent classes of the Markov chain underlying the system [18]. The significance of end components in the case of Markov decision processes is stated by the following theorem.

Theorem 3 [8] *For all $s \in S$ and all policies π, we have*
$$\Pr_s^\pi\big(SAPairs^{-1}(InfSA) \text{ is an } EC\big) = 1 \; .$$

Given a fairness constraint $\mathcal{F}$ for $\mathcal{P}$, we say that an end component (C, D) is a *fair end component* (FEC) if the following condition holds, in addition to closure and connectivity:

- *Fairness:* For all $s \in C$, we have $\mathcal{F}(s) \subseteq D(s)$.

We define containment and maximality for FECs as for ECs, and we denote by $MFec(U, \mathcal{F})$ the set of maximal FECs contained in $U \subseteq S$. Again, for each $U \subseteq S$ set $MFec(U, \mathcal{F})$ can be computed in time polynomial in the size of the MDP. The following theorem indicates that fair end components are the corresponding concept to end components in presence of fairness.

Theorem 4 *For all $s \in S$ and all $\pi \in ProbF \cup PathF$, we have*
$$\Pr_s^\pi\big(SAPairs^{-1}(InfSA) \text{ is a } FEC\big) = 1 \; .$$

This theorem was proved by [17] for path fairness, and by [8] for probabilistic fairness. The proof for probabilistic fairness is in fact immediate: one needs only examine the definition of probabilistic fairness to realize that Theorem 4 follows immediately from Theorem 3. Unbounded fairness behaves differently from path or probabilistic fairness with respect to end components, as shown by the following proposition.

Proposition 2 *For every EC (C, D) and $0 < q < 1$, we can construct a policy $\pi(q) \in UnbF$ such that for all $s \in C$,*
$$\Pr_s^{\pi(q)}\big(SAPairs^{-1}(InfSA) = (C, D)\big) \geq q \; .$$

Proof. Given q, we construct an infinite sequence $\{r_i(q)\}_{i \geq 0}$ of real numbers such that $0 < r_i(q) < 1$ for $i \geq 0$, and $\prod_{i=0}^\infty r_i(q) = q$, by letting $r_i(q) = q^{(1/2^{i+1})}$. Then, policy $\pi(q)$ can be constructed as follows: at step i of the path, if $X_i \notin C$, then π chooses uniformly at random an action from $A(X_i)$. If instead $X_i \in C$, then π chooses each action in $D(X_i)$ with probability
$$\frac{r_i(q)}{|D(X_i)|} + \frac{1 - r_i(q)}{|A(X_i)|} \; ,$$
and each action in $A(s) \setminus D(X_i)$ with probability $(1 - r_i(q))/|A(X_i)|$. It is easy to check that the policy $\pi(q)$ thus constructed has the required property. Note that policy $\pi(q)$ is *history-dependent,* i.e. its behavior at t depends on the prefix of path from the starting state s to t (in this case, the dependence

is through the length of the path prefix). ∎ ∎

4.2 *Parametric Markov chains*

To help with the construction of sets of approximating fair policies, we present some results on *parametric Markov chains.* In these chains the coefficients of the transition matrix are expressed as a function of a parameter. We present conditions that ensure that if the coefficients are continuous functions of the parameter, then also the steady-state distribution of the chain depends continuously on the parameter.

Given a memoryless policy π, we define a transition matrix $P = [p_{s,t}]_{s,t \in S}$ corresponding to π by taking, for all $s, t \in S$,

$$p_{s,t} = \sum_{a \in A(s)} \pi(s)(a)\, p(s,a)(t) \ .$$

Recall that a *sub-stochastic matrix* is a matrix $P = [p_{s,t}]_{s,t \in S}$ such that $0 \leq p_{s,t} \leq 1$ for $s, t \in S$, and $\sum_{t \in S} p_{s,t} \leq 1$ for all $s \in S$ [18]. The matrix corresponding to a memoryless policy is sub-stochastic (in fact, it is also *stochastic,* since $\sum_{t \in S} p_{s,t} = 1$ for all $s \in S$). Given a sub-stochastic matrix P, the steady-state (or limiting) matrix P^* of P is defined by $P^* = \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} P^n$. The following two propositions can be proved by linear algebra arguments [8], and they provide sufficient conditions under which the steady-state distribution of a Markov chain is a continuous function of a parameter. The first proposition covers the case in which the closed recurrent classes of the chain do not depend on the parameter.

Proposition 3 *For a fixed N, consider a family $P(x) = [p_{s,t}(x)]_{s,t \in S}$ of sub-stochastic matrices parameterized by a parameter $x \in I$, where $I \subseteq \mathbb{R}$ is an interval of real numbers. Assume that the Markov chain having P as transition matrix has the same set of closed recurrent classes for all $x \in I$. Then, if the coefficients of $P(x)$ depend continuously on x for $x \in I$, also the coefficients of the steady-state matrix $P^*(x)$ depend continuously on x for $x \in I$.*

A similar result holds for chains in which there is a single closed recurrent class (which may change as the parameter changes), and there is a fixed state that is always in that class, for all values of the parameter. To state the result, we say that a state is *surely recurrent* if the Markov chain has only one closed recurrent class, and the state belongs to that class. In this case, the steady-state matrix P^* can be written as $P^* = \mathbf{1}^t \boldsymbol{u}$, where $\mathbf{1}^t$ is the transpose of a vector consisting of $|S|$ 1's, and $\boldsymbol{u}$ is the vector of the steady-state (or limiting) distribution of the Markov chain.

Proposition 4 *For a fixed N, consider a family $P(x) = [p_{s,t}(x)]_{s,t \in S}$ of sub-stochastic matrices parameterized by a parameter $x \in I$, where $I \subseteq \mathbb{R}$ is an* *interval of real numbers. Assume that there is a state $1 \leq k_0 \leq N$ that is surely recurrent for all $x \in I$. Then, if the coefficients of $P(x)$ depend continuously on x for $x \in I$, also the coefficients of the steady-state distribution vector $\mathbf{u}(x)$ depend continuously on x for $x \in I$.*

4.3 Unconditionally fair policy

In the following arguments, it will be useful to have a fixed policy that is fair with respect to all notions of fairness discussed in this paper. Hence, we denote by π_f the memoryless policy that at each state $s \in S$ chooses uniformly at random an action $a \in A(s)$.

5 Acceptance Probability

In this section we prove Theorem 2, part (i), and we provide algorithms for computing the maximum acceptance probability under the different notions of fairness. The equalities in Theorem 2, part (i) are proved by showing that the algorithms for the relative notions of fairness coincide.

5.1 Probabilistic fairness

The algorithm for computing the maximum acceptance probability for probabilistic fairness is taken from [8]. By Theorem 4, with probability 1 the set of states repeated infinitely often along a path form a FEC. Given a Rabin acceptance condition $\mathcal{A} = \{(Q^p_1, Q^r_1), \ldots, (Q^p_m, Q^r_m)\}$ and a FEC (C, D), we say that the FEC *satisfies* $\mathcal{A}$ (written $(C, D) \models \mathcal{A}$) iff there is $1 \leq i \leq m$ such that $C \subseteq Q^p_i$ and $C \cap Q^r_i \neq \emptyset$. If (C, D) satisfies $\mathcal{A}$, and if a path starting from C chooses at each $s \in C$ an action in $D(s)$ uniformly at random, the path will satisfy $\mathcal{A}$ with probability 1. Hence, let

$$R_\mathcal{A} = \bigcup \{C \mid (C, D) \textit{ is a FEC and } (C, D) \models \mathcal{A}\}$$

be the union of the sets of states of all the FECs that satisfy $\mathcal{A}$. The set $R_\mathcal{A}$ can be computed more efficiently by

$$R_\mathcal{A} = \bigcup_{i=1}^{m} \bigcup \Big\{C \;\Big|\; (C, D) \in \mathit{MFec}(Q^p_i) \wedge C \cap Q^r_i \neq \emptyset \Big\} .$$

Once $R_\mathcal{A}$ is reached, it is easy to see that the acceptance condition can be satisfied with probability 1 under a probabilistically fair policy. In fact, there is a memoryless policy $\pi_r \in \mathit{ProbF}$ such that $\mathrm{Pr}^{\pi_r}_s(\theta \models \mathcal{A}) = 1$ for every $s \in R_\mathcal{A}$ (see [8, §8] for the details of the construction of π_r, inspired by [7]). The surprising fact is that it suffices to compute the maximum probability of reaching $R_\mathcal{A}$ *under any policy,* rather than under any probabilistically fair policy, as stated by the following proposition (as shown for path fairness in [17]).

Proposition 5 *For all $s \in S$, we have* $\Pr_s^+(\mathit{ProbF}, \mathcal{A}) = \max_{\pi\in\Pi} \Pr_s^\pi(\Diamond R_\mathcal{A})$.

In this proposition, $\Diamond R_\mathcal{A}$ denotes the event of reaching $R_\mathcal{A}$, as defined in Section 2.1. We write $\max_{\pi\in\Pi} \Pr_s^\pi(\Diamond R_\mathcal{A})$ instead of $\sup_{\pi\in\Pi} \Pr_s^\pi(\Diamond R_\mathcal{A})$, even though Π is an infinite set, to underline the fact that there is a policy $\pi_0 \in \Pi$ such that

$$\Pr_s^{\pi_0}(\Diamond R_\mathcal{A}) = \sup_{\pi\in\Pi} \Pr_s^\pi(\Diamond R_\mathcal{A}) \ .$$

A similar convention is used throughout the remainder of the paper. The interest of Proposition 5 lies in the fact that the quantity $\max_{\pi\in\Pi} \Pr_s^\pi(\Diamond R_\mathcal{A})$ can be computed using a well-known reduction to linear programming, which leads to a polynomial-time algorithm [6].

Proof. To prove Proposition 5, we prove that for all $s \in S$ the following equalities hold:

(6) $$\sup_{\pi\in \mathit{ProbF}} \Pr_s^\pi(\theta \models \mathcal{A}) = \sup_{\pi\in \mathit{ProbF}} \Pr_s^\pi(\Diamond R_\mathcal{A}) = \max_{\pi\in\Pi} \Pr_s^\pi(\Diamond R_\mathcal{A}) \ .$$

To prove (6), we first note that

$$\begin{aligned} \max_{\pi\in\Pi} \Pr_s^\pi(\Diamond R_\mathcal{A}) &\geq \sup_{\pi\in \mathit{ProbF}} \Pr_s^\pi(\Diamond R_\mathcal{A}) \\ &\geq \sup_{\pi\in \mathit{ProbF}} \Pr_s^\pi(\theta \models \mathcal{A}) \ . \end{aligned}$$

The first inequality is immediate; the second follows from the fact that a path from s follows with probability 1 a FEC, so that the probability of satisfying $\mathcal{A}$ without entering $R_\mathcal{A}$ is 0. In the reverse direction, a result on Markov decision processes establishes the existence of a memoryless deterministic policy π_d such that, for all $s \in S$,

$$\Pr_s^{\pi_d}(\Diamond R_\mathcal{A}) = \max_{\pi\in\Pi} \Pr_s^\pi(\Diamond R_\mathcal{A})$$

(see for example [6], and for a detailed proof, [8, §3]). Let also $B \subseteq S$ be the set of states that cannot reach $R_\mathcal{A}$. From π_d, we construct the policy π_e that coincides with π_d on $S \setminus (R_\mathcal{A} \cup B)$, with π_r on $R_\mathcal{A}$, and with π_f on B. Since π_e and π_d coincide on $C = S \setminus (R_\mathcal{A} \cup B)$, we have

(7) $$\Pr_s^{\pi_e}(\Diamond R_\mathcal{A}) = \Pr_s^{\pi_d}(\Diamond R_\mathcal{A}) = \max_{\pi\in\Pi} \Pr_s^\pi(\Diamond R_\mathcal{A})$$

for all $s \in S$. If $\pi_e \in \mathit{ProbF}$, then the argument is easily concluded. Otherwise, we construct a set of probabilistically fair policies that approximates π_e. For $0 \leq x \leq 1$, define the memoryless policy $\pi[x]$ by:

$$\pi[x](s)(a) = \begin{cases} \pi_r(s)(a) & \text{if } s \in R_\mathcal{A} \\ (1-x)\,\pi_e(s)(a) + x\,\pi_f(s)(a) & \text{otherwise.} \end{cases}$$

It is easy to check that $\pi[x] \in \mathit{ProbF}$ for $0 < x \leq 1$. Since for $0 < x \leq 1$ policy $\pi[x]$ is just one of many probabilistically fair ones that tries to satisfy $\mathcal{A}$, we have

(8) $$\sup_{\pi\in \mathit{ProbF}} \Pr_s^\pi(\theta \models \mathcal{A}) \geq \lim_{x\to 0} \Pr_s^{\pi[x]}(\Diamond R_\mathcal{A}) \ .$$

To complete the argument, from (7) it remains to show that

$$\lim_{x \to 0} \Pr_s^{\pi[x]}(\diamond R_{\mathcal{A}}) \geq \Pr_s^{\pi_e}(\diamond R_{\mathcal{A}}) \ . \tag{9}$$

To this end, denote by $P(x) = [p_{s,t}(x)]_{s,t \in S}$ the matrix corresponding to $\pi[x]$, for $0 \leq x \leq 1$. Note that $P(0)$ is equal to the matrix P_e corresponding to π_e. The closed recurrent classes of $P(x)$ are constant for $0 \leq x < 1$. In fact, for $0 \leq x < 1$ the closed recurrent classes of $\pi[x]$ are all contained in $B \cup R_{\mathcal{A}}$, and $\pi[x]$ does not depend on x in $B \cup R_{\mathcal{A}}$. Denoting by $P^*(x) = [p^*_{s,t}(x)]_{s,t \in S}$ the steady-state matrix corresponding to P, we can write the reachability probability of $R_{\mathcal{A}}$ for all $s \in S$ as

$$\Pr_s^{\pi[x]}(\diamond R_{\mathcal{A}}) = \sum_{t \in R_{\mathcal{A}}} p^*_{s,t}(x) \ .$$

From $\lim_{x \to 0} P(x) = P(0) = P_e$, by Proposition 3, we have $\lim_{x \to 0} P^*(x) = P^*_D$, from which we obtain (9), which together with (8) and (7) concludes the argument. ■ ■

5.2 Path fairness

Since Theorem 4 holds both for probabilistic and for path fairness, the first step in the computation of $\Pr_s^+(\mathit{PathF}, \mathcal{A})$ consists in computing the set $R_{\mathcal{A}} \subseteq S$, and it coincides with the first step of the computation of $\Pr_s^+(\mathit{PathF}, \mathcal{A})$. In fact, we want to prove that the algorithm for path fairness is the same as the one for probabilistic fairness, as stated by the following proposition.

Proposition 6 *For all $s \in S$, we have*

$$\Pr_s^+(\mathit{PathF}, \mathcal{A}) = \max_{\pi \in \Pi} \Pr_s^\pi(\diamond R_{\mathcal{A}}) = \Pr_s^+(\mathit{ProbF}, \mathcal{A}) \ .$$

Proof. To prove the proposition, we prove that the following equalities hold for all $s \in S$:

$$\sup_{\pi \in \mathit{PathF}} \Pr_s^\pi(\theta \models \mathcal{A}) = \sup_{\pi \in \mathit{PathF}} \Pr_s^\pi(\diamond R_{\mathcal{A}}) = \max_{\pi \in \Pi} \Pr_s^\pi(\diamond R_{\mathcal{A}}) \ . \tag{10}$$

Again, in one direction the inequalities follow easily:

$$\begin{aligned} \max_{\pi \in \Pi} \Pr_s^\pi(\diamond R_{\mathcal{A}}) &\geq \sup_{\pi \in \mathit{PathF}} \Pr_s^\pi(\diamond R_{\mathcal{A}}) \\ &\geq \sup_{\pi \in \mathit{PathF}} \Pr_s^\pi(\theta \models \mathcal{A}) \ . \end{aligned}$$

In the other direction, note that probabilistic fairness implies path fairness (Proposition 1). Thus, to prove that for all $s \in S$

$$\sup_{\pi \in \mathit{PathF}} \Pr_s^\pi(\theta \models \mathcal{A}) \geq \max_{\pi \in \Pi} \Pr_s^\pi(\diamond R_{\mathcal{A}}) \tag{11}$$

it suffices note that for $0 < x < 1$, the policy $\pi[x]$ used in the proof of (8) and (9) is also path fair. Hence, we can immediately duplicate the argument for (8) and (9) for path fairness, leading to (11) and finally (10). ■ ■

5.3 Unbounded fairness

For unbounded fairness, we define the set $R^\bullet_{\mathcal{A}}$ by

$$R^\bullet_{\mathcal{A}} = \bigcup \{C \mid (C,D) \text{ is an EC and } (C,D) \models \mathcal{A}\}$$
$$= \bigcup_{i=1}^{m} \bigcup \{C \mid (C,D) \in Mec(Q^p_i) \wedge C \cap Q^r_i \neq \emptyset\} \ .$$

Differently from $R_{\mathcal{A}}$, the set $R^\bullet_{\mathcal{A}}$ is computed disregarding the fairness constraints of the MDP. In fact, to compute the maximum acceptance probability for unbounded fairness, it turns out that it is not necessary to take fairness into account, as the following proposition states.

Proposition 7 *For all $s \in S$, we have*

$$\Pr^+_s(UnbF, \mathcal{A}) = \max_{\pi \in \Pi} \Pr^\pi_s(\Diamond R^\bullet_{\mathcal{A}}) = \Pr^+_s(NoF, \mathcal{A}) \ .$$

Proof. The rightmost equality simply encodes the algorithm for maximum acceptance probability without fairness [9]. Regarding the leftmost equality, again in one direction the inequalities follow easily: for all $s \in S$,

$$\max_{\pi \in \Pi} \Pr^\pi_s(\Diamond R^\bullet_{\mathcal{A}}) \geq \sup_{\pi \in UnbF} \Pr^\pi_s(\Diamond R^\bullet_{\mathcal{A}})$$
$$\geq \sup_{\pi \in UnbF} \Pr^\pi_s(\theta \models \mathcal{A})$$
$$= \Pr^+_s(UnbF, \mathcal{A}) \ .$$

In the other direction, in analogy with the proof of Proposition 2, for all $0 \leq \varepsilon \leq 1$ we can construct a policy $\pi[\varepsilon] \in UnbF$ such that for all $s \in S$ and all finite path prefixes σ ending in $R^\bullet_{\mathcal{A}}$, we have

$$Prb^{\pi[\varepsilon]}_s(\theta \models \mathcal{A} \mid \sigma \preceq \theta) \geq 1 - \varepsilon \ .$$

Let also π_d be a policy such that

$$\Pr^{\pi_d}_s(\Diamond R^\bullet_{\mathcal{A}}) = \max_{\pi \in \Pi} \Pr^\pi_s(\Diamond R^\bullet_{\mathcal{A}}) \ ,$$

and let $B \subseteq S$ be the set of states that cannot reach $R^\bullet_{\mathcal{A}}$. For all $s \in S$, all $a \in A(s)$, all $0 \leq x \leq 1$, all $0 < \varepsilon < 1$, and all $\bar{s} \in S^*$, we define policy $\pi[x, \varepsilon]$ by:

$$\pi[x,\varepsilon](\bar{s}, s)(a) = \begin{cases} (1-x)\, \pi_d(s)(a) + x\, \pi_f(s)(a) & \text{if } s \in S \setminus (B \cup R^\bullet_{\mathcal{A}}); \\ \pi_f(s)(a) & \text{if } s \in B; \\ \pi[\varepsilon](s)(a) & \text{if } s \in R^\bullet_{\mathcal{A}}. \end{cases}$$

For $0 < x, \varepsilon < 1$, we have $\pi[x, \varepsilon] \in UnbF$; the result then follows by noting that for all $s \in S$ we have

$$\lim_{x \to 0} \Pr^{\pi[x,\varepsilon]}_s(\Diamond R^\bullet_{\mathcal{A}}) = \Pr^{\pi_d}_s(\Diamond R^\bullet_{\mathcal{A}}) = \max_{\pi \in \Pi} \Pr^\pi_s(\Diamond R^\bullet_{\mathcal{A}}) \tag{12}$$

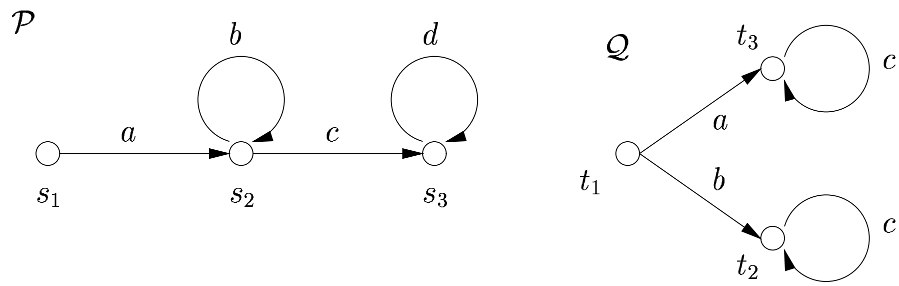

Fig. 3. Two MDPs $\mathcal{P}$ and $\mathcal{Q}$. The MDP $\mathcal{P}$ is deterministic, and has an associated fairness constraint $\mathcal{F}_\mathcal{P}$ defined by $\mathcal{F}_\mathcal{P}(s_2) = \{c\}$, and $\mathcal{F}_\mathcal{P}(s_1) = \mathcal{F}_\mathcal{P}(s_3) = \emptyset$. The MDP $\mathcal{Q}$ has an associated fairness constraint $\mathcal{F}_\mathcal{Q}$ defined by $\mathcal{F}_\mathcal{Q}(t_1) = \{a, b\}$, and $\mathcal{F}_\mathcal{Q}(t_2) = \mathcal{F}_\mathcal{Q}(t_3) = \emptyset$.

and hence

$$\begin{aligned}
\Pr_s^+(\mathit{UnbF}, \mathcal{A}) &= \sup_{\pi \in \mathit{UnbF}} \Pr_s^\pi(\theta \models \mathcal{A}) \\
&\geq \lim_{x\to 0} \lim_{\varepsilon \to 0} \Pr_s^{\pi[x,\varepsilon]}(\theta \models \mathcal{A}) \\
&= \lim_{x \to 0} \Pr_s^{\pi[x,\varepsilon]}(\Diamond R^\bullet_\mathcal{A}) \\
&= \Pr_s^{\pi_d}(\Diamond R^\bullet_\mathcal{A}) \ ,
\end{aligned} \tag{13}$$

as was to be shown. The proof of (12) and (13) follows the lines of the proofs of Propositions 5 and 6. ∎ ∎

Finally, the result of Theorem 2, part (i) follows by noting that $R_\mathcal{A} \subseteq R^\bullet_\mathcal{A}$, and by comparing Propositions 5, 6, and 7.

5.4 A counterexample to equality

To see that the inequality in Theorem 2, part (i) cannot in general be replaced by equality, consider the MDP $\mathcal{P}$ of Figure 3, together with the acceptance condition $\mathcal{A} = \{(\{s_2\}, \{s_2\})\}$. We have $\Pr_{s_1}^+(\mathit{ProbF}, \mathcal{A}) = 0$ and $\Pr_{s_1}^+(\mathit{PathF}, \mathcal{A}) = 1$.

6 Reachability Cost

In this section, we study the algorithms for computing the minimum reachability cost under the various notions of fairness, and in the process we prove Theorem 2, part (ii).

Given a state $s \in S$ and an action $a \in A(s)$ for s, we denote by

$$\mathit{dest}(s, a) = \{t \in S \mid p(s, a)(t) > 0\}$$

the set of possible successors of s when a is selected.

Since the costs are strictly positive, the cost from a state $s \in S$ to the target set $R \subseteq S$ can be finite only if R can be reached from s with probability 1.

Hence, before presenting the algorithms for the various notions of fairness, we present an algorithm that computes the set of states from which R can be reached with probability 1, under a generic fairness constraint $\mathcal{G} : S \mapsto 2^{Acts}$ (not necessarily coinciding with the constraint $\mathcal{F}$ of the MDP). The algorithm is essentially the algorithm of [8], presented in an improved notation. To present the algorithm, we define the predicate $\mathit{FApre}(Y, X, \mathcal{G})$ over S, where $X, Y \subseteq S$ and $\mathcal{G} : S \mapsto 2^{Acts}$, by $s \models \mathit{FApre}(Y, X, \mathcal{G})$ iff:

$$\begin{aligned} &\forall a \in \mathcal{G}(s) \,.\, \mathit{dest}(s,a) \subseteq Y \\ \wedge\ &\exists a \in A(s) \,.\, \big(\mathit{dest}(s,a) \subseteq Y \wedge \mathit{dest}(s,a) \cap X \neq \emptyset\big) \,. \end{aligned}$$

For $R \subseteq S$ and $\mathcal{G} : S \mapsto 2^{Acts}$, we then define $\mathit{Reach}(R, \mathcal{G})$ by the μ-calculus formula:

$$\mathit{Reach}(R, \mathcal{G}) = \nu Y \,.\, \mu X \,.\, \big(\mathit{FApre}(Y, X, \mathcal{G}) \vee R\big) \,, \tag{14}$$

where we have used the slightly improper notation of using R as a predicate that holds exactly for the states in R. The following proposition can be proved by induction on the iterations used to compute the μ-calculus formula.

Proposition 8 *Given an absorbing target set $R \subseteq S$ and $\mathcal{G} : S \mapsto 2^{Acts}$, let U be the largest subset of states of S that satisfies the following two properties:*

- *For all $s \in U \setminus R$ and all $a \in \mathcal{G}(U)$, we have $\mathit{dest}(s,a) \subseteq U$.*
- *For all $s \in U$, there is a path from s to R in the graph (U, E), where*
$$E = \big\{(s,t) \in U \times U \;\big|\; \exists a \in A(s) \,.\, [\mathit{dest}(s,a) \subseteq U \wedge t \in \mathit{dest}(s,a)]\big\} \,.$$

Then, $U = \mathit{Reach}(R, \mathcal{G})$.

6.1 Probabilistic fairness

The following proposition establishes that $\mathit{Reach}(R, \mathcal{F})$ is the set of states from which the minimum cost to R converges.

Proposition 9 [8] *We have $v_s^-(\mathit{ProbF}, c, R) < \infty$ iff $s \in \mathit{Reach}(R, \mathcal{F})$.*

Proof. In one direction, Proposition 9 follows easily from Proposition 8. In fact, consider the policy that at each $t \in U$ chooses the action from $\{a \in A(t) \mid \mathit{dest}(t,a) \subseteq U\}$ uniformly at random. Under this policy, R is reached with probability 1 and within finite expected time from all $s \in U$, ensuring the convergence of the minimum cost. In the other direction, an inductive argument that follows the structure of (14) shows that if $s \notin U$, then $\Pr_s^\pi(\Diamond R) < 1$ for all $\pi \in \mathit{ProbF}$ (see [13] for related arguments), which leads to the result. ∎ ∎

For all $s \in U$, it is possible to compute the minimum cost to R under no fairness assumptions $v_s^-(\mathit{NoF}, c, R)$ by solving a stochastic shortest path problem [4]. The following result states that this cost is equal to the cost $v_s^-(\mathit{ProbF}, c, R)$ under probabilistic fairness. Together with Proposition 9, this yields an algorithm for the computation of $v_s^-(\mathit{ProbF}, c, R)$ for all $s \in S$.

Proposition 10 *For all $s \in \mathit{Reach}(R, \mathcal{F})$ we have*

$$v_s^-(\mathit{ProbF}, c, R) = v_s^-(\mathit{NoF}, c, R) \ .$$

Proof. To prove this result, we again use the idea of approximating the (possibly unfair) policy corresponding to the stochastic shortest path problem with a set of probabilistically fair policies. To this end, let $U = \mathit{Reach}(R, \mathcal{F})$, and let π_d be a memoryless policy such that for all $s \in U$ we have $v_s^{\pi_d}(c, R) = v_s^-(\mathit{NoF}, c, R)$ (for the existence of such a policy, see [4]). Let also π_u be any memoryless policy that at all $s \in U$ chooses an action from $\{a \in A(s) \mid \mathit{dest}(s, a) \subseteq U\}$ uniformly at random. For $0 \leq x \leq 1$, we define the memoryless policy $\pi[x]$ by, for all $s \in S$ and $a \in A(s)$,

(15) $$\pi[x](s)(a) = (1 - x)\, \pi_d(s)(a) + x\, \pi_u(s)(a) \ .$$

Note that for $0 < x \leq 1$ we have $\pi[x] \in \mathit{ProbF}$. We want to show that for all $s \in U$, we have

(16) $$\lim_{x \to 0} v_s^{\pi[x]}(c, R) = v_s^{\pi_d}(c, R) \ .$$

From this equation, Proposition 10 follows easily. To show (16), first observe that it suffices to focus on the set $V = U \setminus R$, since neither π_d nor π_u lead from U to outside U, and since the reachability cost from R is 0. Denote by $P(x) = [p_{s,t}(x)]_{s,t \in V}$ the probability transition matrix corresponding to the policy $\pi[x]$ restricted to set V, and note that $P(0)$ is the probability transition matrix corresponding to π_d. For $0 \leq x \leq 1$ define also the vector $\boldsymbol{z}(x) = [z_s(x)]_{s \in V}$ by

$$z_s(x) = \sum_{a \in A(s)} \pi[x](s)(a)\ c(s, a) \ .$$

With this notation, from (2) for $s \in V$ and $0 \leq x \leq 1$ we have

$$v_s^{\pi[x]}(c, R) = \sum_{k=0}^{\infty} P^k(x) \boldsymbol{z}(x) = (I - P(x))^{-1} \boldsymbol{z}(x) \ .$$

Since for $0 \leq x \leq 1$ the matrix $P(x)$ corresponds to a transient Markov chain, we have $\det(I - P(x)) \neq 0$ in this interval. Thus, for $0 \leq x \leq 1$ the coefficients of $(I - P(x))^{-1}$ are rational functions of x that have no poles in the interval $[0, 1]$. Since also $\boldsymbol{z}(x)$ is continuous in $[0, 1]$, we finally have

$$\begin{aligned} \lim_{x \to 0} v_s^{\pi[x]}(c, R) &= \lim_{x \to 0} (I - P(x))^{-1} \boldsymbol{z}(x) \\ &= (I - P(0))^{-1} \boldsymbol{z}(0) \\ &= v_s^{\pi_d}(c, R) \end{aligned}$$

thus proving (16). ∎ ∎

6.2 Unbounded fairness

The equivalent of Proposition 9 can be proved also for unbounded fairness.

Proposition 11 [8] *We have* $v_s^-(UnbF, c, R) < \infty$ *iff* $s \in Reach(R, \mathcal{F})$.

The rest of the analysis for the proof of Proposition 10 can then be carried through unchanged, observing that for all $0 < x \le 1$ the policy $\pi[x]$ defined by (15) is such that $\pi[x] \in UnbF$. Hence, we obtain the following result.

Proposition 12 *For all* $s \in S$, *we have* $v_s^-(UnbF, c, R) = v_s^-(ProbF, c, R)$.

6.3 Path fairness

With respect to reachability cost, path fairness behaves differently from the other two notions of fairness. The following proposition states that

$$v_s^-(PathF, c, R) = v_s^-(NoF, c, R)$$

for all $s \in S$.

Proposition 13 *Denote by* $(\lambda s.\emptyset) : S \mapsto 2^{Acts}$ *the empty fairness constraint. For all* $s \in S$, *the following assertions hold:*

(i) *If* $s \notin Reach(R, \lambda s.\emptyset)$, *then* $v_s^-(PathF, c, R) = v_s^-(NoF, c, R) = \infty$.

(ii) *If* $s \in Reach(R, \lambda s.\emptyset)$, *then* $v_s^-(PathF, c, R) = v_s^-(NoF, c, R)$.

Proof. Let $U^\bullet = Reach(R, \lambda s.\emptyset)$. The first assertion is shown by proving that if $s \notin U^\bullet$ then $\Pr_s^\pi(\Diamond R) < 1$ for all policies π, so that $v_s^\pi(c, R) = \infty$ for all policies π. This result is proved using an inductive argument on the iterations of (14).

For $s \in U^\bullet$, the second assertion can be proved as follows. Let π_d be the memoryless policy such that $v_s^{\pi_d}(c, R) = v_s^-(NoF, c, R)$. Define π_e to be the (history-dependent) policy that coincides with π_d until R is reached, and that chooses actions uniformly at random after R is reached. We have $\pi_e \in PathF$: in fact, under policy $\widetilde{\pi}_d$ any path that reaches R is fair, and the set of paths that never reach R has probability measure 0. It is then immediate to check that $v_s^{\pi_d}(c, R) = v_s^{\pi_e}(c, R)$, leading to the result. ∎ ∎

6.4 Fairness and reachability

Together, Propositions 9, 10, 12, and 13 prove Theorem 2, part (ii). Intuitively, Theorem 2, part (ii) can be interpreted as follows. Let

$$U = Reach(R, \mathcal{F})$$
$$U^\bullet = Reach(R, \lambda s.\emptyset) .$$

If $U = U^\bullet$, then under all three notions of fairness we can achieve a cost to R that is arbitrarily close to that achieved by the optimal (not necessarily fair) policy. If $U \subset U^\bullet$, on the other hand, the inequality in Theorem 2, part (ii) is strict for some $s \in U^\bullet \setminus U$. In this latter case, the difference between the behavior of probabilistic and unbounded fairness on one side, and path fairness on the other, is essentially due to the following phenomenon. Suppose that from a state s, in order to reach R, a path must visit a state t, with $A(t) = \{a, b\}$. From t, action a leads to R, and action b leads to a set of states that cannot reach R. Probabilistic and unbounded fairness require that a policy be fair at all steps. Hence, under a probabilistically or unboundedly fair policy, action b must be selected with non-zero probability, and the expected cost to R will be infinite. On the other hand, path fairness does not impose requirements on all steps of the paths. As long as a policy visits t only finitely often (which is the case here), the policy can deterministically select a at t, and the expected cost to R will converge.

6.5 *A counterexample to equality*

To see that the inequality in Theorem 2, part (ii) cannot in general be replaced by equality, consider the MDP $\mathcal{Q}$ of Figure 3. Let c be the cost function that associates 1 with all state-action pairs of the MDP, and let $R = \{t_2\}$. We have $v^-_{t_1}(\mathit{NoF}, c, R) = 1$ and $v^-_{t_1}(\mathit{ProbF}, c, R) = \infty$.

7 Long-Run Average Outcome

Before dealing with the case of general MDPs, we prove that the three notions of fairness lead to the same maximum long-run average outcome, provided the MDP is *strongly connected.* We say that the MDP $\mathcal{P} = (S, \mathit{Acts}, A, p)$ is strongly connected if the graph (S, E) is strongly connected, where $E = \{(s,t) \in S \times S \mid \exists a \in A(s) \,.\, p(s,a)(t) > 0\}$. The following proposition summarizes several results for strongly connected MDPs.

Proposition 14 [8, §5] *Consider a strongly connected MDP $\mathcal{P}$ with state space S, together with outcome and cost functions R, W. The following assertions hold.*

- *The value of $H_s(\mathit{NoF}, R, W)$ does not depend on $s \in S$. The common value $H(\mathit{NoF}, R, W)$ can be computed in time polynomial in the size of $\mathcal{P}$ by solving a linear programming problem.*
- *There is a memoryless policy π_d such that $H_s(\mathit{NoF}, R, W) = H^{\pi_d}_s(R, W)$ for all $s \in S$. Moreover, the transition probability matrix P_d induced by π_d corresponds to a Markov chain having a single closed recurrent class.*

Using this proposition, we can show that the maximum long-run average outcome coincide for our three notions of fairness on strongly connected MDPs.

Proposition 15 *On a strongly connected MDP, for all $s \in S$ and $\Phi \in \{ProbF, PathF, UnbF\}$, we have*

$$H_s(\Phi, R, W) = H(NoF, R, W) .$$

Proof. The proof of this proposition is once more based on approximating the optimal policy in the absence of fairness with a set of fair policies. Let π_d be as in Proposition 14. For $0 \leq x \leq 1$, all $s \in S$ and all $a \in A(s)$, we define the memoryless policy $\pi[x]$ by

$$\pi[x](s)(a) = (1-x)\,\pi_d(s)(a) + x\,\pi_f(s)(a) .$$

For $0 < x \leq 1$, we have $\pi[x] \in ProbF$. For $0 \leq x \leq 1$, denote by $P(x)$ the transition probability matrix arising from $\pi[x]$, and define the vectors $\boldsymbol{r}(x) = [r_s(x)]_{s \in S}$ and $\boldsymbol{w} = [w_s(x)]_{s \in S}$ by

$$r_s(x) = \sum_{a \in A(s)} R(s,a)\,\pi[x](s)(a)$$

$$w_s(x) = \sum_{a \in A(s)} W(s,a)\,\pi[x](s)(a) .$$

Denote by $P^*(x) = [p^*_{s,t}(x)]_{s,t \in S}$ the steady-state probability distribution matrix corresponding to $P(x)$. By our choice of π_d, the Markov chain corresponding to $P(0)$ has a single closed recurrent class $C \subseteq S$. Since the MDP is strongly connected, by definition of $\pi[x]$ all states of C are surely recurrent for $0 \leq x < 1$. Hence, as a consequence of standard facts on Markov chains we have

$$H_s^{\pi[x]}(R,W) = \frac{\sum_{t \in S} p^*_{s,t}(x)\; r_t(x)}{\sum_{t \in S} p^*_{s,t}(x)\; w_t(x)} .$$

Moreover, Proposition 4 ensures that $\lim_{x \to 0} P^*(x) = P(0)$. Since for all $t \in S$ quantities $r_t(x)$ and $w_t(x)$ are continuous for $x \to 0$, we have

$$\lim_{x \to 0} H_s^{\pi[x]}(R,W) = H(NoF, R, W) .$$

Hence, for all $s \in S$ we have $H_s(ProbF, R, W) \geq H(NoF, R, W)$. Since the reverse inequality is immediate, we conclude

$$H_s(ProbF, R, W) = H(NoF, R, W)$$

as was to be shown. The equivalent results for *PathF* and *UnbF* follow then immediately by observing that $ProbF \subseteq PathF$ and $ProbF \subseteq UnbF$. ∎ ∎

7.1 Probabilistic fairness

If the MDP $\mathcal{P}$ is strongly connected, Proposition 14 and 15 provide a method for the computation of $H_s(ProbF, R, W)$ at all $s \in S$. In the general case, from (4) we see that the expected long-run average outcome depends only on the states and actions that are repeated infinitely often. Theorem 4 states that these states and actions form a FEC with probability 1: hence, we can concentrate our attention on the maximal FECs. Let

$$MFec(S, \mathcal{F}) = \mathcal{L} = \{(C_1, D_1), \ldots, (C_n, D_n)\} ,$$

and note that for $1 \le i \le n$, the FEC (C_i, D_i) is a strongly connected sub-MDPs of the global MDP. Hence, for $1 \le i \le n$ we can associate with (C_i, D_i) the maximum long-run average outcome $H^i(NoF, R, W)$ that can be obtained when staying forever in (C_i, D_i), computed using Proposition 14 and 15.

Once the maximum long-run average outcomes for the maximal FECs have been computed, we can compute $H_s(ProbF, R, W)$ at all $s \in S$ using an idea that originates from [6]. For all $1 \le i \le n$, we add to the MDP a special state t_i, which signals the intention to stay in (C_i, D_i) forever. For $1 \le i \le n$, we let $A(t_i) = \{b_i\}$, where b_i is an action that leads back t_i, i.e. $dest(t_i, b_i) = \{t_i\}$. The set of states $\{t_1, \ldots, t_n\}$ is thus absorbing. For all $1 \le i \le n$ and all $s \in C_i$, we also add to $A(s)$ a new action a_i that leads deterministically to t_i: the choice of a_i represents the decision of staying in (C_i, D_i) from that point on. Finally, we associate with each state $s \in S \cup \{t_1, \ldots, t_n\}$ and $a \in A(s)$ of the new MDP a final reward $h(s)$ defined by

$$h(s, a) = \begin{cases} H^i(NoF, R, W) & \text{if } s \in C_i \text{ and } a = a_i, \text{ for } 1 \le i \le n; \\ 0 & \text{otherwise.} \end{cases}$$

For $1 \le i \le n$, the reward associated with a transition from C_i to t_i is thus equal to the maximum long-run average reward that can be obtained by staying in (C_i, D_i) forever; the reward h is 0 on all other transitions.

Denote by $\mathcal{P}[(C_1, D_1), \ldots, (C_n, D_n)]$ the MDP obtained from $\mathcal{P}$ in this fashion. The following proposition states that the maximum long-run average outcome $H_s(ProbF, R, W)$ at all $s \in S$ can be computed by solving a maximum expected total reward problem on $\mathcal{P}[(C_1, D_1), \ldots, (C_n, D_n)]$, using h as the reward.

Proposition 16 *Let the MDP $\mathcal{P}[(C_1, D_1), \ldots, (C_n, D_n)]$ and the reward h be as described above. Then, for all $s \in S$ we have:*

$$H_s(ProbF, R, W) = \max_{\pi \in \Pi} \mathrm{E}_s^{\pi} \left\{ \sum_{k=0}^{\infty} h(X_k, Y_k) \right\} , \qquad (17)$$

where the max *in (17) exists.*

The maximum expected total cost mentioned in the proposition can be solved in several ways: see for example [3] or, for more efficient algorithms tailored to this type of problem, [8, §7][12].

Proof. On the one hand, consider a memoryless policy π_e for the MDP $\mathcal{P}[(C_1, D_1), \ldots, (C_n, D_n)]$ that realizes the maximum in the total cost problem (17).

For $1 \le i \le n$, we can assume that if π_e chooses with positive probability action a_i at some state $s \in C_i$, then it chooses a_i deterministically at all states of C_i. In fact, assume towards the contradiction that at $t \in C_i$ there is a strictly better choice from the point of view of total cost. Since (C_i, D_i) is strongly connected, then a strictly better policy would be obtained by choosing all actions in D_i uniformly at random at all states of $C_i \setminus \{t\}$, until t is reached, and choosing the better choice at t, contradicting the hypothesis that π_e is optimal.

For $0 \leq x \leq 1$, from π_e we construct a memoryless policy $\pi[x]$ for $\mathcal{P}$ as follows. Policy $\pi[x]$ coincides with π_e on all $S \setminus \bigcup_{i=1}^n C_i$. For $1 \leq i \leq n$, if π_e does not choose a_i at C_i, then $\pi[x]$ coincides with π_e also on C_i. If π_e chooses a_i in C_i, for $1 \leq i \leq n$, then we take $\pi[x]$ to coincide with the probabilistically fair x-optimal policy for (C_i, D_i), constructed as in the proof of Proposition 15.

On the basis of $\pi[x]$, for $0 \leq \varepsilon \leq 1$ and $0 \leq x \leq 1$ we construct a memoryless policy $\pi[\varepsilon, x]$ by, for all $s \in S$ and $a \in A(s)$,

$$\pi[\varepsilon, x](s)(a) = \begin{cases} \pi[x](s)(a) & \text{if } s \in \bigcup_{i=1}^n C_i \\ (1-\varepsilon)\,\pi_e(s)(a) + \varepsilon\,\pi_f(s)(a) & \text{otherwise.} \end{cases}$$

Using arguments similar to those for Propositions 10 and 15, it is not difficult to prove that for all $s \in S$, we have

$$\lim_{\varepsilon \to 0} \lim_{x \to 0} H_s^{\pi[\varepsilon, x]}(R, W) = \mathrm{E}_s^{\pi_e}\left\{\sum_{k=0}^{\infty} h(X_k, Y_k)\right\} ,$$

which leads to the result.

In the other direction, consider an arbitrary probabilistically fair policy π. Under policy π, the paths are with probability 1 eventually confined to some (C_i, D_i) with $1 \leq i \leq n$. Once confined in (C_i, D_i), it is possible to prove that π cannot do better than $H^i(\mathit{NoF}, R, W)$ (see [8] for a detailed argument). Hence, for all $s \in S$ we have

$$H_s^\pi(R, W) \leq \sum_{i=1}^{n} H^i(\mathit{NoF}, R, W)\, \mathrm{Pr}_s^\pi\big(\mathit{InfSA} = \mathit{SAPairs}(C_i, D_i)\big) ,$$

and from this follows easily the result. ∎ ∎

7.2 Path and unbounded fairness

Similarly to probabilistic fairness, also under path fairness the set of state-action pairs that are repeated infinitely often along a path forms a FEC with probability 1. Hence, we can repeat for path fairness the same reasoning done in the previous subsection for probabilistic fairness. From the equality of the algorithms for the computation of the maximum long-run average outcome for these two notions of fairness, we obtain that for all $s \in S$,

$$H_s(\mathit{ProbF}, R, W) = H_s(\mathit{PathF}, R, W) , \tag{18}$$

which is one part of Theorem 2, part (iii).

For unbounded fairness, Proposition 2 tells us that a path that enters an EC can stay forever in the EC with probability arbitrarily close to 1, even if the EC is not fair. This suggests that for dealing with unbounded fairness, the only modification needed to the algorithm of the previous section is to take $\mathcal{L} = \mathit{Mec}(S)$ instead of $\mathcal{L} = \mathit{MFec}(S)$, thus considering all ECs, including the unfair ones. This intuition is confirmed by the following proposition.

Proposition 17 *We have $H_s(\mathit{UnbF}, R, W) = H_s(\mathit{NoF}, R, W)$ for all $s \in S$.*

Proof. The inequality $H_s(\mathit{UnbF}, R, W) \geq H_s(\mathit{NoF}, R, W)$ holds trivially for all $s \in S$. To show the converse inequality, the key step is to show that, given an EC (C, D), we have a set of policies $\pi[x]$ such that, for all $t \in C$,

$$\lim_{x \to 0} H_t^{\pi[x]}(R, W) = H_t(\mathit{NoF}, R, W)[(C, D)] \tag{19}$$

$$\lim_{x \to 0} \Pr_t^{\pi[x]}\big(\forall k \geq 0 \,.\, (X_k \in C \wedge Y_k \in D(X_k))\big) = 1 \,, \tag{20}$$

where $H_t(\mathit{NoF}, R, W)[(C, D)]$ refers to the maximum long-run average outcome that can be obtained *on the EC (C, D)*, rather than on the whole MDP. To this end, let π_d be a memoryless policy such that $H_t^{\pi_d}(R, W) = H_t(\mathit{NoF}, R, W)[(C, D)]$ for all $t \in C$. By definition, we have that

$$\Pr_t^{\pi_d}(\forall k \geq 0 \,.\, (X_k \in C \wedge Y_k \in D(X_k)) = 1 \,.$$

For $0 \leq x \leq 1$, construct the policy $\pi[x]$ by

$$\pi[x](s_1, \ldots, s_k) = (1 - x)^{1/2^k} \pi_d(s_k) + \big(1 - (1 - x)^{1/2^k}\big) \pi_f(s_k)$$

for all $k \geq 1$ and all $s_1, s_2, \ldots, s_k \in C$, and by $\pi[x](s_1, \ldots, s_k) = \pi_f$ for $k \geq 1$ and $s_k \notin C$. A straightforward calculation shows that

$$\Pr_t^{\pi[x]}\big(\forall k \geq 0 \,.\, (X_k \in C \wedge Y_k \in D(X_k))\big) = 1 - x \,,$$

which shows (20). In addition, notice that policy $\pi[x]$ is a linear combination of π_d and π_f that is always at least as close to π_d as $(1 - x)\pi_d + x\pi_f$. Hence, (19) follows from the same arguments used to prove Proposition 15.

Once (19) and (20) have been proved, the results follows from considering the MDP $\mathcal{P}[(C_1, D_1), \ldots, (C_n, D_n)]$ obtained as for Proposition 16, except that $(C_1, D_1), \ldots, (C_n, D_n)$ are the ECs (instead of the FECs) of the original MDP, and that the final rewards are defined by $h(s, a) = H_s(\mathit{NoF}, R, W)[(C, D)]$ for all ECs (C, D) and all $s \in C$, and $h(s, a) = 0$ otherwise. The result can be obtained by reasoning as in the proof of Proposition 16. ∎ ∎

7.3 A counterexample to equality

To see that the inequality in Theorem 2, part (iii) cannot in general be replaced by equality, consider the MDP $\mathcal{P}$ of Figure 3. We consider two functions R and W, such that W is equal to 1 for all state-action pairs, and R is defined by $R(s_1, a) = R(s_2, c) = R(s_3, d) = 0$ and $R(s_2, b) = 1$. Then, it is easy to check that $H_{s_1}(\mathit{NoF}, R, W) = 1$ and $H_{s_1}(\mathit{ProbF}, R, W) = 0$.

Acknowledgments.

We thank Tom Henzinger, Marta Kwiatkowska, and Amir Pnueli for inspiring discussions on the notion of fairness in probabilistic and non-probabilistic systems. We also thank the anonymous referees for many helpful suggestions.

References

[1] R. Alur and T.A. Henzinger. Finitary fairness. In *Proc. 9th IEEE Symp. Logic in Comp. Sci.*, pages 52–61, 1994.

[2] A. Bianco and L. de Alfaro. Model checking of probabilistic and nondeterministic systems. In *Found. of Software Tech. and Theor. Comp. Sci.*, volume 1026 of *Lect. Notes in Comp. Sci.*, pages 499–513. Springer-Verlag, 1995.

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

[4] D.P. Bertsekas and J.N. Tsitsiklis. An analysis of stochastic shortest path problems. *Math. of Op. Res.*, 16(3):580–595, 1991.

[5] C. Courcoubetis and M. Yannakakis. Verifying temporal properties of finite-state probabilistic programs. In *Proc. 29th IEEE Symp. Found. of Comp. Sci.*, pages 338–354, 1988.

[6] C. Courcoubetis and M. Yannakakis. Markov decision processes and regular events. In *Proc. 17th Int. Colloq. Aut. Lang. Prog.*, volume 443 of *Lect. Notes in Comp. Sci.*, pages 336–349. Springer-Verlag, 1990.

[7] C. Courcoubetis and M. Yannakakis. The complexity of probabilistic verification. *J. ACM*, 42(4):857–907, 1995.

[8] L. de Alfaro. *Formal Verification of Probabilistic Systems.* PhD thesis, Stanford University, 1997. Technical Report STAN-CS-TR-98-1601.

[9] L. de Alfaro. Temporal logics for the specification of performance and reliability. In *Proc. of Symp. on Theor. Asp. of Comp. Sci.*, volume 1200 of *Lect. Notes in Comp. Sci.*, pages 165–176. Springer-Verlag, 1997.

[10] L. de Alfaro. How to specify and verify the long-run average behavior of probabilistic systems. In *Proc. 13th IEEE Symp. Logic in Comp. Sci.*, pages 454–465, 1998.

[11] L. de Alfaro. Stochastic transition systems. In *CONCUR'98: Concurrency Theory. 9th Int. Conf.*, Lect. Notes in Comp. Sci. Springer-Verlag, 1998.

[12] L. de Alfaro. Computing minimum and maximum reachability times in probabilistic systems. To appear in *CONCUR'99: Concurrency Theory. 10th Int. Conf.*, Lect. Notes in Comp. Sci., Springer-Verlag, 1999.

[13] L. de Alfaro, T.A. Henzinger, and O. Kupferman. Concurrent reachability games. In *Proc. 39th IEEE Symp. Found. of Comp. Sci.*, 1998.

[14] C. Derman. *Finite State Markovian Decision Processes.* Acedemic Press, 1970.

[15] N. Francez. *Fairness.* Springer-Verlag, 1986.

[16] S. Hart, M. Sharir, and A. Pnueli. Termination of probabilistic concurrent programs. *ACM Trans. Prog. Lang. Sys.*, 5(3):356–380, 1983.

[17] M. Kwiatkowska and C. Baier. Model checking for a probabilistic branching time logic with fairness. Technical Report CSR-96-12, University of Birmingham, 1996.

[18] J.G. Kemeny, J.L. Snell, and A.W. Knapp. *Denumerable Markov Chains.* D. Van Nostrand Company, 1966.

[19] M. Kwiatkowska. Survey of fairness notions. *Information and Software Technology*, 31(7):371–386, 1989.

[20] D. Lehmann, A. Pnueli, and J. Stavi. Impartiality, justice and fairness: the ethics of concurrent termination. In *Proc. 8th Int. Colloq. Aut. Lang. Prog.*, volume 115 of *Lect. Notes in Comp. Sci.*, pages 264–277. Springer-Verlag, 1981.

[21] Z. Manna and A. Pnueli. *The Temporal Logic of Reactive and Concurrent Systems: Specification.* Springer-Verlag, New York, 1991.

[22] A. Pnueli. On the extremely fair treatment of probabilistic algorithms. In *Proc. 15th ACM Symp. Theory of Comp.*, pages 278–290, 1983.

[23] A. Pnueli and L. Zuck. Probabilistic verification by tableaux. In *Proc. First IEEE Symp. Logic in Comp. Sci.*, pages 322–331, 1986.

[24] A. Pnueli and L.D. Zuck. Probabilistic verification. *Information and Computation*, 103:1–29, 1993.

[25] M.O. Rabin. Probabilistic automata. *Information and Computation*, 6:230–245, 1963.

[26] S.M. Ross. Average cost semi-Markov decision processes. *J. Appl. Prob.,* 7:649–656, 1970.

[27] S. Safra. On the complexity of ω-automata. In *Proc. 29th IEEE Symp. Found. of Comp. Sci.*, 1988.

[28] R. Segala. *Modeling and Verification of Randomized Distributed Real-Time Systems.* PhD thesis, MIT, 1995. Technical Report MIT/LCS/TR-676.

[29] R. Segala and N.A. Lynch. Probabilistic simulations for probabilistic processes. In *CONCUR'94: Concurrency Theory. 5th Int. Conf.*, volume 836 of *Lect. Notes in Comp. Sci.*, pages 481–496. Springer-Verlag, 1994.

[30] W. Thomas. Automata on infinite objects. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume B, chapter 4, pages 135–191. Elsevier Science Publishers (North-Holland), Amsterdam, 1990.

[31] M.Y. Vardi. Automatic verification of probabilistic concurrent finite-state systems. In *Proc. 26th IEEE Symp. Found. of Comp. Sci.*, pages 327–338, 1985.

[32] M. Vardi. Infinite games against nature, 1995.