Formal Quality Measures for Predictors in Markov Decision Processes

Chistel Baier[1], **Sascha Klüppelholz**[1] **Jakob Piribauer**[1,2], **Robin Ziemek**[1] *

[1]Technische Universität Dresden
[2]Universität Leipzig
{christel.baier, sascha.klueppelholz, jakob.piribauer, robin.ziemek}@tu-dresden.de

Abstract

In adaptive systems, predictors are used to anticipate changes in the system's state or behavior that may require system adaption, e.g., changing its configuration or adjusting resource allocation. Therefore, the quality of predictors is crucial for the overall reliability and performance of the system under control. This paper studies predictors in systems exhibiting probabilistic and non-deterministic behavior modeled as Markov decision processes (MDPs). Main contributions are the introduction of quantitative notions that measure the effectiveness of predictors in terms of their average capability to predict the occurrence of failures or other undesired system behaviors. The average is taken over all memoryless policies. We study two classes of such notions. One class is inspired by concepts that have been introduced in statistical analysis to explain the impact of features on the decisions of binary classifiers (such as precision, recall, f-score). Second, we study a measure that borrows ideas from recent work on probability-raising causality in MDPs and determines the quality of a predictor by the fraction of memoryless policies under which (the set of states in) the predictor is a probability-raising cause for the considered failure scenario.

Extended version — https://arxiv.org/abs/2412.11754

1 Introduction

In modern days, AI systems grow ever more complex and harder to understand, e.g. code designed by artificial intelligence tends to be very abstruse and thus is not comprehensible in a simple way. Since a full understanding of such systems is difficult to establish, it is important to predict certain events within such systems. In particular, situations in which the system produces unwanted or even disastrous results need to be predicted early and precisely.

In the area of formal verification, counterexamples, invariants and related certificates are often used to provide a verifiable justification that a system does or does not behave according to a specification (see e.g., (Manna and Pnueli 1995; Clarke, Grumberg, and Peled 1999; Namjoshi 2001)). However, most AI systems can not be designed in a way that failure can be excluded and then certificates do not provide enough insights on the systems decisions to predict its behavior. In order to get an understanding *why* a system behaves the way it does, we introduce measures in how well certain events in a system serve as a predictor for undesired outcomes. Events which have a cause-effect relation to such outcomes constitute a special case of such predictors (Halpern and Pearl 2005; Pearl 2009).

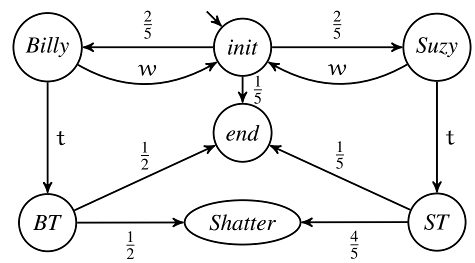

Figure 1: An experiment where we want to predict *Shatter*

In this paper, we consider binary predictors in Markov decision processes (MDPs) which are a stochastic operational model with non-deterministic choices. We interpret the nondeterminism as uncertainty about the future behavior and thus it may or may not be resolved adversarial to our goals.

For example, consider an experiment with two participating persons "Suzy" and "Billy" which are asked to throw a rock at a bottle of glass. This example has been widely discussed for in philosophic literature on *causality* (Hall 2004; Chockler and Halpern 2004; Halpern 2015). In our variant (Fig. 1) a randomized process decides, which person is allowed to take a throw or whether the experiment ends. However, the participants can choose to wait w or to throw t. If someone decides to throw their rock, the state is changed accordingly (*ST* for "Suzy throws" and *BT* for "Billy throws"). For predictions in this system, we do not make any assumptions on the decisions of the participants and consider both the outcomes and predictors described by sets of states. For example, reaching the state *ST* is intuitively a good predictor since the throw of Suzy has a high probability of hitting. However, if Suzy does not feel confident and thus has a decides to throw with a low probability, then state *ST* will only be reached with low probability and so a lot of scenarios in which the bottle shatters come from Billys throw. In order to distinguish between the quality of different predictors, we

*Authors are listed in alphabetical order

Copyright © 2025, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

use measures from statistical analysis (Powers 2011). However, the quality of a prediction may rely on the distribution over the decisions, as e.g. between w and t for *Suzy*. So, we consider an *average* case scenario with respect to the non-determinism for the quality of a prediction.

Furthermore, we also consider whether a chosen predictor has a cause-effect relation to the undesired outcome. Such a probabilistic cause-effect relation in MDPs is introduced in (Baier, Piribauer, and Ziemek 2024) where the probability-raising (PR) principle is invoked for each possible resolution of the non-determinism. Inspired by this, we introduce *probability-raising policies*, which witness a PR condition. A predictor can then also be rated by the relative amount of resolutions (the *causal volume*) in which it has a probabilistic cause-effect relation with the predicted event.

Contributions By considering binary predictors in MDPs we formally introduce an average case analysis for quality measures depending on the non-determinism in order to rate the quality of a predictor (Sec. 3). For this we use a uniform measure over the memoryless randomized policies. We also introduce the concept of probability-raising policies (Sec. 4), which we use to define causal volumes for predictors in MDPs as an additional way to get information about the quality of a predictor (Sec. 4.1). We then address the complexity of deciding the existence of PR policies (Sec. 4.2).

Related Work Considering the quality of a prediction has connections to responsibility (Chockler and Halpern 2004; Chockler, Halpern, and Kupferman 2008), blameworthiness (Halpern and Kleiman-Weiner 2018) and harm (Beckers, Chockler, and Halpern 2023). In (Mascle et al. 2021) and (Baier, Funke, and Majumdar 2021) forward responsibilities based on the Shapley value are allocated in Kripke and game structures. Recent work considers backwards responsibility in deterministic AI systems (Baier et al. 2024).

All these notions of responsibility are based on causality as a necessary condition (Braham and van Hees 2012). A causality based account of responsibility can be found in (Chockler, Halpern, and Kupferman 2008). In stochastic operational models, probabilistic causes are used as predictors in (Ziemek et al. 2022) and interpreted as binary classifiers in (Baier, Piribauer, and Ziemek 2024).

Predicting events in Markovian models also has connections to monitoring properties. The fact that randomization improves monitors for non-probabilistic systems has been examined in (Chadha, Sistla, and Viswanathan 2009). A current risk-value is estimated for states in partially observable MDPs in (Junges, Torfah, and Seshia 2021).

2 Preliminaries

In the context of this work a *Markov decision process (MDP)* is a 4-tuple $\mathcal{M} = (S, Act, P, init)$ where S is a finite set of states, Act a finite set of actions, $init \in S$ the initial state and $P : S \times Act \times S \to [0, 1]$ the probabilistic transition function such that $\sum_{t \in S} P(s, \alpha, t) \in \{0, 1\}$ for all states $s \in S$ and actions $\alpha \in Act$. An action α is *enabled* in state $s \in S$ if $\sum_{t \in S} P(s, \alpha, t) = 1$ and $Act(s)$ denotes the set of enabled actions in S. A state t is *terminal* if $Act(t) = \emptyset$. A *path* in an MDP $\mathcal{M}$ is a (finite or infinite) alternating sequence $\pi = s_0\, \alpha_0\, s_1\, \alpha_1\, s_2 \cdots \in (S \times Act)^* \cup (S \times Act)^\omega$ such that $P(s_i, \alpha_i, s_{i+1}) > 0$ for all indices i. A path is called maximal if it is infinite or finite and ends in a terminal state. An MDP can be seen as a Kripke structure in which transitions go from states to probability distributions over states.

A *(randomized) policy* $\mathfrak{S}$ is a function that maps each finite non-maximal path $s_0\alpha_0 \ldots \alpha_{n-1}s_n$ to a distribution over $Act(s_n)$. $\mathfrak{S}$ is called deterministic if $\mathfrak{S}(\pi)$ is a Dirac distribution for all finite non-maximal paths π. If the chosen action only depends on the last state of the path, $\mathfrak{S}$ is called *memoryless*. We write MR for the class of memoryless (randomized) and MD for the class of memoryless deterministic policies. *Finite-memory* policies are those that are representable by a finite-state automaton.

A policy $\mathfrak{S}$ of $\mathcal{M}$ induces a (possibly infinite) Markov chain. We write $\Pr^{\mathfrak{S}}_{\mathcal{M},s}$ for the standard probability measure on measurable sets of maximal paths in the Markov chain induced by $\mathfrak{S}$ with initial state s. We use the abbreviation $\Pr^{\mathfrak{S}}_{\mathcal{M}} = \Pr^{\mathfrak{S}}_{\mathcal{M},init}$. We use linear temporal logic (LTL) modalities such as $\Diamond$ (eventually) and U (until) to denote path properties. For $X, T \subseteq S$ the formula $X\,U\,T$ is satisfied by $\pi = s_0s_1 \ldots$ if there is $j \geqslant 0$ such that for all $i < j : s_i \in X$ and $s_j \in T$ and $\Diamond T = S\,U\,T$. It is well-known that $\Pr^{\min}_{\mathcal{M}}(X\,U\,T)$ and $\Pr^{\max}_{\mathcal{M}}(X\,U\,T)$ and corresponding optimal MD-policies are computable in polynomial time.

For $s \in S$ and $\alpha \in Act(s)$, (s, α) is a state-action pair of $\mathcal{M}$. We denote the set of state-action pairs of $\mathcal{M}$ by *StAct*. An *end component* (EC) of an MDP $\mathcal{M}$ is a strongly connected sub-MDP containing at least one state-action pair.

For a policy $\mathfrak{S}$ of $\mathcal{M}$, the expected *frequencies* of state-action pairs (s, α) are

$$freq_{\mathfrak{S}}(s, \alpha) = \mathrm{E}^{\mathfrak{S}}_{\mathcal{M}}(\text{no. of visits to } s \text{ in which } \alpha \text{ is taken})$$

In end-component free MDPs we can specify MR policies by their state-action frequencies (see e.g. (Kallenberg 2020, Theorem 4.7)), by considering a linear constraint system over variables $x_{s,\alpha}$ for each $(s, \alpha) \in StAct$:

$$x_{s,\alpha} \geqslant 0 \qquad \text{for all } (s, \alpha) \in StAct, \tag{S1}$$

$$x_{init} = 1 + \sum_{(t,\alpha) \in StAct} x_{t,\alpha} \cdot P(t, \alpha, init), \tag{S2}$$

$$x_s = \sum_{(t,\alpha) \in StAct} x_{t,\alpha} \cdot P(t, \alpha, s) \quad \text{for all } s \in S \setminus \{init\}, \tag{S3}$$

where we use the short form notation $x_s = \sum_{\alpha \in Act(s)} x_{s,\alpha}$. By (Kallenberg 2020, Theorem 4.7) a solution $x \in \mathbb{R}^{StAct}$ to (S1)-(S3) corresponds one-to-one to an MR policy $\mathfrak{S}$ for $\mathcal{M}$ such that $x_{s,\alpha} = freq_{\mathfrak{S}}(s, \alpha)$ for all $(s, \alpha) \in StAct$.

3 Measuring the Quality of a Predictor

The goal of this section is to measure how well reaching a set of states C predicts that a set of terminal states E will be reached in an MDP $\mathcal{M}$. For this, we consider well-known measures for binary classifiers from statistical analysis. To apply such measures, the non-determinism in $\mathcal{M}$ needs to be resolved. However, without any assumptions about the resolution of the non-determinism, we propose to consider an average over possible resolutions of the non-determinism.

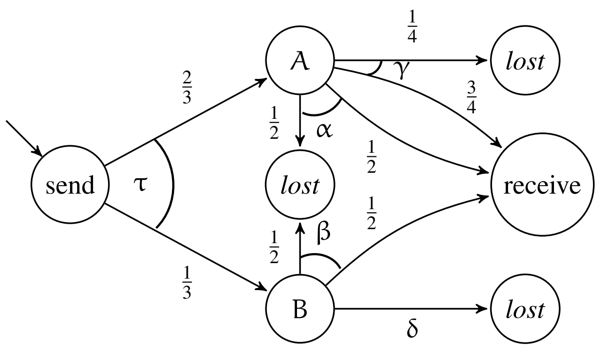

Figure 2: An MDP $\mathcal{M}$ depicting a communication network.

3.1 Averaging over Policies

To be able to compute this average, one has to choose which policies to consider and how to weigh them. In this paper, we use a uniform measure over MR policies. Since we consider reachability properties in particular, the choice of MR policies is justified, as any policy in an EC-free MDP for these properties can be approximated by MR policies (Kallenberg 2020). Furthermore, non-determinism can be used to model uncertainty in the transition probabilities as also represented in interval-Markov chains (see, e.g., (Kozine and Utkin 2002; Sen, Viswanathan, and Agha 2006)). In this case, actions can be used to model the extremal transition probabilities. Randomizing over such actions then allows for any possible concrete distribution over successor states.

Example 1. As an example, consider a communication network in which a message is sent by a sender to a receiver via various network nodes. Each node forwards the message to other nodes in a randomized fashion. We only know the successors of each network node as well as upper and lower bounds on the respective probability and on the probability that the message is lost between two nodes. Suppose, we have the suspicion that some sets of nodes are faulty and are the reason for many message losses. In order to check this claim, we want to measure how well reaching such a set of nodes serves as a predictor for a message loss by considering the average over possible resolutions of the non-determinism. A very simple example of such a network where actions model the known probability bounds is given in Figure 2. $\triangleleft$

Given an MDP $\mathcal{M} = (S, Act, P, init)$ with terminal states $T \subseteq S$, the set of all MR policies $\mathfrak{S}$ of $\mathcal{M}$ can be described by the set of vectors

$$\mathfrak{P} = \{x \in [0,1]^{StAct} \mid \sum_{\alpha \in Act(s)} x_{s,\alpha} = 1 \quad \text{for all } s \in S \setminus T\}.$$

The component $x_{s,\alpha}$ of $x \in \mathfrak{P}$ expresses the probability that the corresponding policy chooses α in s. $\mathfrak{P}$ is a product of $|S \setminus T|$-many regular simplices. For $s \in S \setminus T$ let $n(s) = |Act(s)| - 1$. Then $\mathfrak{P}$ can be identified with $\prod_{s \in S} \Delta^{n(s)}$, where $\Delta^m = \{y \in [0,1]^{m+1} \mid \sum_{i=1}^{m+1} y_i = 1\}$ is the regular m-simplex of dimension m in $\mathbb{R}^{m+1}$ for given $m \in \mathbb{N}$. In turn, the dimension of $\mathfrak{P}$ is $|StAct| - |S \setminus T|$.

	$\Diamond E$	$\neg\Diamond E$
$\Diamond C$	True positive $tp^{\mathfrak{S}}_{\mathcal{M}} = \Pr^{\mathfrak{S}}(\Diamond C \wedge \Diamond E)$	False positive $fp^{\mathfrak{S}}_{\mathcal{M}} = \Pr^{\mathfrak{S}}(\Diamond C \wedge \neg\Diamond E)$
$\neg\Diamond C$	False negative $fn^{\mathfrak{S}}_{\mathcal{M}} = \Pr^{\mathfrak{S}}(\neg\Diamond C \wedge \Diamond E)$	True negative $tn^{\mathfrak{S}}_{\mathcal{M}} = \Pr^{\mathfrak{S}}(\neg\Diamond C \wedge \neg\Diamond E)$

Figure 3: The confusion matrix for the prediction of $E \subseteq S$ by $C \subseteq S$ for a given policy $\mathfrak{S}$.

Example 2. For the MDP $\mathcal{M}$ depicted in Fig. 2 considering the state-action pairs $(init, \tau), (A, \alpha), (A, \gamma), (B, \beta), (B, \delta)$ we get $\mathfrak{P} = \{(1, p, 1-p, q, 1-q) \mid p, q \in [0,1]\}$. We can identify $\mathfrak{P}$ with the product of simplices $\{1\} \times [0,1] \times [0,1]$ or simply $[0,1]^2$. $\triangleleft$

Using the Lebesgue measure, we get a way to uniformly average over all MR policies. As mentioned above, we want to apply this to some quality measure for binary classifiers. For a quality measure $Q^x(C)$ for the predictor $C \subseteq S$ depending on a policy $x \in \mathfrak{P}$, the average of the quality measure is

$$\int_{\mathfrak{P}} Q^x(C)\, dx \;/\; \int_{\mathfrak{P}} 1\, dx.$$

The computation of $V(\mathfrak{P}) = \int_{\mathfrak{P}} 1\, dx$ can be done in polynomial time since $\mathfrak{P}$ is the product of standard simplices as described above. Then, by (Stein 1966) we have

$$\int_{\mathfrak{P}} 1\, dx = V(\prod_{s \in S} \Delta^{n(s)}) = \prod_{s \in S} V(\Delta^{n(s)}) = \prod_{s \in S} \frac{1}{n(s)!},$$

which is computable in polynomial time. In fact, integrating any polynomial of fixed degree over $\mathfrak{P}$ can also be done in polynomial time (Baldoni et al. 2011).

3.2 Quality Measures

Given a single policy $\mathfrak{S}$ for $\mathcal{M}$, we consider the so-called *confusion matrix* (Powers 2011) for binary classifiers as depicted in Figure 3. In statistical analysis, various measures for the quality of a classifier using the entries of the confusion matrix have been studied. Important measures are

$$precision^{\mathfrak{S}}(C) = \frac{tp^{\mathfrak{S}}}{tp^{\mathfrak{S}} + fp^{\mathfrak{S}}}, \; recall^{\mathfrak{S}}(C) = \frac{tp^{\mathfrak{S}}}{tp^{\mathfrak{S}} + fn^{\mathfrak{S}}},$$

$$fscore^{\mathfrak{S}}(C) = \frac{2}{\frac{1}{precision^{\mathfrak{S}}(C)} + \frac{1}{recall^{\mathfrak{S}}(C)}} = \frac{2tp^{\mathfrak{S}}}{2tp^{\mathfrak{S}} + fp^{\mathfrak{S}} + fn^{\mathfrak{S}}}.$$

Intuitively, the $precision^{\mathfrak{S}}(C)$ measures the probability that the prediction is indeed true after C is reached. The $recall^{\mathfrak{S}}(C)$, on the other hand, expresses the probability that reaching E was preceded by reaching C under the condition that E has indeed been reached. The $fscore^{\mathfrak{S}}(C)$ is the harmonic mean of precision and recall. A high f-score hence indicates that after reaching the predictor C, the probability to reach E is relatively high and at the same time relatively many executions leading to E pass through C before. An example for a more complex quality measures – similar in spirit to the f-score (Chicco and Jurman 2020) – is Matthews correlation coefficient given by $\mathrm{mcc}^{\mathfrak{S}}(C) =$

$$\frac{tp^{\mathfrak{S}} \cdot tn^{\mathfrak{S}} - fp^{\mathfrak{S}} \cdot fn^{\mathfrak{S}}}{\sqrt{(tp^{\mathfrak{S}} + fp^{\mathfrak{S}}) \cdot (tp^{\mathfrak{S}} + fn^{\mathfrak{S}}) \cdot (tn^{\mathfrak{S}} + fp^{\mathfrak{S}}) \cdot (tn^{\mathfrak{S}} + fn^{\mathfrak{S}})}}.$$

Computing average quality measures. In order to compute the average quality measures over the polytope $\mathfrak{P}$ of MR policies, we have to express the measures in terms of the policies. As the considered quality measures depend on the confusion matrix, the task hence is to express the entries of the confusion matrix in terms of the vectors $x \in \mathfrak{P}$

To express these values, we take a small detour via a model transformation such that we can distinguish whether C has occurred or not. We define the *two-copy MDP* $\mathcal{M}_C$ of $\mathcal{M}$ with respect to C in the following way:

- $\mathcal{M}_C$ consists of two copies of $\mathcal{M}$, namely $\mathcal{M}_0$ and $\mathcal{M}_1$,
- the initial state is $init_0$ of the first copy,
- whenever a state $c_0 \in C_0$ is reached in the first copy there is exactly one action which transitions to the corresponding state c_1 in the second copy with probability 1.

A policy $\mathfrak{S}$ of the original MDP $\mathcal{M}$ can be interpreted as a policy $\mathfrak{U}$ of the two-copy MDP $\mathcal{M}_C$ by mimicking the behavior in both copies ignoring the switch to the second copy. We denote the copies of the states in C and E in $\mathcal{M}_0$ and $\mathcal{M}_1$ by C_0, C_1, E_0, and E_1, respectively. Now, we can describe the entries of the confusion matrix for $\mathcal{M}$ directly by reachability probabilities in $\mathcal{M}_C$.

Lemma 3. *For an MR policy $\mathfrak{S}$ of $\mathcal{M}$ also viewed as a policy for $\mathcal{M}_C$ the following holds*

$$\begin{aligned} tp^{\mathfrak{S}}_{\mathcal{M}} &= \Pr^{\mathfrak{S}}_{\mathcal{M}_C}(\Diamond E_1), \quad fn^{\mathfrak{S}}_{\mathcal{M}} = \Pr^{\mathfrak{S}}_{\mathcal{M}_C}(\Diamond E_0), \\ fp^{\mathfrak{S}}_{\mathcal{M}} &= \sum_{c \in C_0} \Pr^{\mathfrak{S}}_{\mathcal{M}_C}(\Diamond c) \cdot (1 - \Pr^{\mathfrak{S}}_{\mathcal{M}_C}(\Diamond E_1)), \\ tn^{\mathfrak{S}}_{\mathcal{M}} &= 1 - tp^{\mathfrak{S}}_{\mathcal{M}} - fp^{\mathfrak{S}}_{\mathcal{M}} - fn^{\mathfrak{S}}_{\mathcal{M}}. \end{aligned}$$

Proof. Only the equation for $fp^{\mathfrak{S}}_{\mathcal{M}}$ must be proven since the other equations follow by construction. The probability that $c \in C_0$ is visited and afterwards $\neg\Diamond E_1$ holds is $\Pr^{\mathfrak{S}}_{\mathcal{M}_C}(\Diamond c) \cdot (1 - \Pr^{\mathfrak{S}}_{\mathcal{M}_C, c}(\Diamond E_1))$. As the set C_0 can only be reached once in $\mathcal{M}_C$, the equation for $fp^{\mathfrak{S}}_{\mathcal{M}}$ follows by adding the probabilities of these disjoint events. □

We denote the state space of $\mathcal{M}_C$ by S' and the set of terminal states by T'. Now, we describe reachability probabilities between states in $\mathcal{M}_C$ in terms of distributions given by MR policies. For $s, t \in S'$ and an MR policy $\mathfrak{S}$ of $\mathcal{M}$ let $a^{\mathfrak{S}}_{s,t} = \Pr^{\mathfrak{S}}_{s}(\Diamond t)$ be the probability to eventually reach t from s under $\mathfrak{S}$. So, $a^{\mathfrak{S}}_{s,s} = 1$ for all $s \in S'$. Using the representation of MR policies as vectors $x \in \mathfrak{P}$, we then have for non-terminal states $s \in S' \setminus T'$ and $t \neq s$,

$$a^x_{s,t} = \sum_{\alpha \in Act(s)} x_{s,\alpha} \cdot \sum_{u \in S} P(s, \alpha, u) \cdot a^x_{u,t}. \qquad \text{(reach)}$$

In fact, after determining the states s' from which t is not reachable in the Markov chain induced by x and $\mathcal{M}_C$ and setting the corresponding variables $a^x_{s',t}$ to 0, the equation system has a unique solution (see (Baier and Katoen 2008)). Given a rational-valued MR policy $x \in \mathfrak{P}$ the reachability probabilities can hence be derived in polynomial time by solving the linear equation system.

However, we want to compute the average of quality measures by taking an integral over the polytope $\mathfrak{P}$ of all MR policies. As we use the standard Lebesgue measure, the boundary of this polytope has measure 0. So, for our purpose, it is sufficient to express the quality measures as a function of the policy $x \in \mathfrak{P}$ on the interior of $\mathfrak{P}$.

Proposition 4. *The values $a^x_{s,t}$ for $s, t \in S'$ are rational functions in x on the interior of $\mathfrak{P}$. The degree of denominator and enumerator of these rational functions is $\leqslant 2|S|$.*

Proof. For a fixed $t \in S'$, the set A_0 of states s' from which t is not reachable in the Markov chain induced by x and $\mathcal{M}_C$ depends on which entries of x are 0. In the interior of $\mathfrak{P}$, no vector has a 0-entry. So, A_0 is independent of x and the corresponding variables can be set 0. As the resulting equation system with at most $|S'|$-many variables has a unique solution (see (Baier and Katoen 2008)), this solution can be expressed using fractions of determinants by Cramer's rule. The determinants are polynomials in the coefficients of the linear equations of degree at most $|S'| = 2|S|$ and the variables $x_{s,\alpha}$ appear only linearly in these coefficients. □

The polynomials in the resulting rational function contain exponentially many monomials in general. Nevertheless, we know that the values $a^x_{s,t}$ are rational functions in x on the interior of $\mathfrak{P}$ that only take values between 0 and 1.

Practical considerations for averages. If the transition relation is sparse, which is often the case for models with large state space, results on the efficient computation of symbolic determinants apply (Kaltofen and Villard 2005; Duriqi et al. 2024). Then, an exact representation of the rational functions $a^x_{s,t}$ can be obtained efficiently.

By Lemma 3, all entries of the confusion matrix under policy $x \in \mathfrak{P}$ can be expressed in terms of $a^x_{s,t}$ by simple arithmetic. Furthermore, the prominent measures for the quality of a predictor such as precision, recall, and f-score are linear rational functions in these entries. But also more complex measures such as the Matthews correlation coefficient are still relatively simple functions in the entries of the confusion matrix. So, all of these quality measures can be expressed as functions $Q(a^x)$ where a^x is the vector containing the values $a^x_{s,t}$. Unfortunately, an exact evaluation of the integral $\int_{\mathfrak{P}} Q(a^x)\, dx$ is nevertheless typically out of reach. But the quality measures take values in $[0, 1]$ or $[-1, 1]$ in the case of the MCC. Together with the fact, that the resulting function $Q(a^x)$ are smooth functions on the interior of $\mathfrak{P}$, standard approaches such as Monte Carlo integration can be used to approximate the integral (Press et al. 2007). Note, that we do not need an explicit representation of the functions $a^x_{s,t}$ when we sample policies $x \in \mathfrak{P}$ as the values $a^x_{s,t}$ can then be computed in polynomial time.

Example 5. Let us apply average quality measures to the network MDP $\mathcal{M}$ depicted in Figure 2. First, we consider A as the predictor. Here, we identify policies with pairs $(\mathsf{p},\mathsf{q}) \in [0,1]^2$ as in Example 2 where p is the probability to choose α in A and q the probability to choose β in B. For a policy $\mathsf{x} = (\mathsf{p},\mathsf{q})$, we get

$$tp^{\mathsf{x}}_{\mathcal{M}} = \frac{1}{6}(1+\mathsf{p}), \qquad fp^{\mathsf{x}}_{\mathcal{M}} = \frac{1}{6}(3-\mathsf{p}),$$
$$tn^{\mathsf{x}}_{\mathcal{M}} = \frac{1}{6}\mathsf{q}, \qquad fn^{\mathsf{x}}_{\mathcal{M}} = \frac{1}{6}(2-\mathsf{q}).$$

As the volume of $\mathfrak{P}$ is 1 in this case, we obtain the average f-score, for example, as

$$\int_{[0,1]^2} fscore^{\mathsf{x}}(\{\mathsf{A}\})\,\mathrm{d}\mathsf{x} = \int_{[0,1]^2} \frac{2tp^{\mathfrak{S}}}{2tp^{\mathfrak{S}} + fp^{\mathfrak{S}} + fn^{\mathfrak{S}}}\,\mathrm{d}\mathsf{x}$$
$$= \int_0^1\int_0^1 \frac{2+2\mathsf{p}}{7+\mathsf{p}-\mathsf{q}}\,\mathrm{d}\mathsf{p}\,\mathrm{d}\mathsf{q} \approx 0.43.$$

Analogously, we obtain

$$\int_{[0,1]^2} fscore^{\mathsf{x}}(\{\mathsf{B}\})\,\mathrm{d}\mathsf{x} = \int_0^1\int_0^1 \frac{4-2\mathsf{q}}{5+\mathsf{p}-\mathsf{q}}\,\mathrm{d}\mathsf{p}\,\mathrm{d}\mathsf{q} \approx 0.60.$$

So, according to the f-score, reaching B is on average a better predictor for a message loss than reaching A. $\lhd$

Remark 6. If a quality measure can be written as a linear rational function in terms of the entries of the confusion matrix, then its minimal or maximal value can be computed with the techniques presented in (Baier, Piribauer, and Ziemek 2024). This is e.g. the case for precision, recall and f-score. For this a standard model transformation is performed which collapses end components (de Alfaro 1997, 1999) while preserving relevant reachability probabilities (c.f. (Ciesinski et al. 2008, III.B) for a compact description). Afterwards, possible combinations of reachability probabilities can be expressed in terms of a linear constraint system for state-action pair frequencies. For linear rational quality measures, the resulting optimization problem can be solved in polynomial time. For more complex measures like Matthews correlation coefficient, more general optimization problems arise.

4 Probability-Raising Policies and the Causal Volume of a Predictor

In this section we investigate, whether in an MDP $\mathcal{M}$ a given predictor C has a probabilistic cause-effect relation with the undesired event E. The idea that a cause for an event serves as a good predictor comes very natural and the usage of probabilistic causes for predictions has already been considered in (Ziemek et al. 2022) for Markov chains. For MDPs we take inspiration from the notion of probability-raising causes from (Baier, Piribauer, and Ziemek 2024) to define two variants of *probability-raising (PR) policies*, which are, in simple terms, witnessing a probability-raising condition. Since the focus of this paper is to investigate the quality of a predictor, we introduce a measure which considers the relative amount of possible MR policies that witness a PR condition. Afterwards we study the complexity of deciding the existence of such PR policy for a given predictor.

Definition 7 (Probability-raising policy). Given $\mathcal{M}$, E and C, a policy $\mathfrak{S}$ of $\mathcal{M}$ is a *global probability-raising policy* (GPR policy) for C and E in $\mathcal{M}$ if the following conditions **(R)** and **(G)** hold

(R) $\quad \mathrm{Pr}^{\mathfrak{S}}(\Diamond C) > 0,$

(G) $\quad \mathrm{Pr}^{\mathfrak{S}}(\Diamond E \mid \Diamond C) > \mathrm{Pr}^{\mathfrak{S}}(\Diamond E). \qquad$ (GPR)

$\mathfrak{S}$ is a *strict probability-raising policy* (SPR policy) for C and E in $\mathcal{M}$ if **(R)** and the following condition **(S)** hold

(S) For all $\mathsf{c} \in C$ with $\mathrm{Pr}^{\mathfrak{S}}((\neg C)\,\mathrm{U}\,\mathsf{c}) > 0$:

$$\mathrm{Pr}^{\mathfrak{S}}(\Diamond E \mid (\neg C)\,\mathrm{U}\,\mathsf{c}) > \mathrm{Pr}^{\mathfrak{S}}(\Diamond E). \qquad \text{(SPR)}$$

There is an implication from strict to global, since the conditional probability in (GPR) is a weighted sum over the conditional probabilities in (SPR) (also cf. (Baier, Piribauer, and Ziemek 2024)). Similarly, for singletons $\{\mathsf{c}\}$ the equivalence of $\Diamond\mathsf{c}$ and $(\neg\mathsf{c})\,\mathrm{U}\,\mathsf{c}$ means that in such cases the strict and global PR conditions coincide.

Under a PR policy $\mathfrak{S}$ there is a causal relationship between the predictor C and the predicted event E. Namely, whenever C is reached under a PR policy $\mathfrak{S}$, the probability of reaching E is raised.

Example 8. For an example we consider the network MDP $\mathcal{M}$ from Figure 2 as before. Here we consider $C = \{\mathsf{B}\}$ as a predictor for $E = \{lost\}$. For the MD policy $\mathfrak{S}$ choosing γ in A and β in B we have

$$\mathrm{Pr}^{\mathfrak{S}}(\Diamond lost \mid \Diamond\mathsf{B}) = \frac{1}{2} > \frac{1}{3}\cdot\frac{1}{2} + \frac{2}{3}\cdot\frac{1}{4} = \frac{1}{3} = \mathrm{Pr}^{\mathfrak{S}}(\Diamond lost)$$

and thus $\mathfrak{S}$ satisfies (GPR) (and (SPR)) and constitutes a PR policy. However, there are policies which do not satisfy (GPR), e.g., the MD policy $\mathfrak{U}$ choosing α in A and β in B. $\lhd$

4.1 Causal Volumes

With the definition of probability-raising policy in mind the question now arises, in how many cases such a causal relation holds between the predictor and the undesired outcome. For this we consider the relative portion of PR policies among all possible MR policies. Recall the polytope $\mathfrak{P}$ of distributions of MR policies in $\mathcal{M}$ from Section 3. Using the reachability matrix corresponding to a given policy $\mathsf{x} \in \mathfrak{P}$ from (reach) we express the conditions **(R)** , **(S)** and **(G)** from Def. 7 by

$$\sum_{\mathsf{c}\in C} \mathsf{a}^{\mathsf{x}}_{init,\mathsf{c}} > 0, \qquad \text{(r)}$$

$$\sum_{eff\in E} \mathsf{a}^{\mathsf{x}}_{\mathsf{c},eff} > \sum_{eff\in E} \mathsf{a}^{\mathsf{x}}_{init,eff} \quad \text{for all } \mathsf{c} \in C \text{ with } \mathsf{a}^{\mathsf{x}}_{init,\mathsf{c}} > 0, \qquad \text{(s)}$$

$$tp^{\mathsf{x}} \cdot tn^{\mathsf{x}} - fp^{\mathsf{x}} \cdot fn^{\mathsf{x}} > 0. \qquad \text{(g)}$$

We define the sets of MR policies which are SPR (or respectively GPR) policies by

$$\mathfrak{P}_{\mathsf{SPR}} = \{\mathsf{x} \in \mathfrak{P} \mid \mathsf{x} \text{ satisfies (reach), (r) and (s)}\},$$
$$\mathfrak{P}_{\mathsf{GPR}} = \{\mathsf{x} \in \mathfrak{P} \mid \mathsf{x} \text{ satisfies (reach), (r) and (g)}\}.$$

The three considered sets have the same dimension:

Theorem 9. *For an MDP $\mathcal{M}$, a set of terminal states E and a predictor set C for which a memoryless SPR (resp. GPR) policy exists, the sets $\mathfrak{P}$ and $\mathfrak{P}_{\mathsf{SPR}}$ (resp. $\mathfrak{P}_{\mathsf{GPR}}$) have the same dimension.*

Proof sketch. The claim follows from the fact that for any $x \in \mathfrak{P}_{\mathsf{SPR}}$ (resp. $\mathfrak{P}_{\mathsf{GPR}}$) there is an $\varepsilon > 0$ such that the ε-neighborhood of x is completely contained in $\mathfrak{P}$. □

Recall the set of all terminal states $\mathsf{T} \subset \mathsf{S}$ of $\mathcal{M}$. With this, we can now define the following *Volumes* with $|StAct| - |\mathsf{S} \setminus \mathsf{T}|$-dimensional Lebesgue integrals:

$$V(\mathfrak{P}) = \int_{\mathfrak{P}} 1 \, dx,$$

$$V(\mathfrak{P}_{\mathsf{SPR}}) = \int_{\mathfrak{P}_{\mathsf{SPR}}} 1 \, dx, \quad V(\mathfrak{P}_{\mathsf{GPR}}) = \int_{\mathfrak{P}_{\mathsf{GPR}}} 1 \, dx.$$

Definition 10. Let $\mathcal{M}$ be an MDP, $E \subset \mathsf{S}$ a set of terminal states and $C \subseteq \mathsf{S} \setminus E$. We define the *strict causal volume* and resp. *global causal volume* of C for E as

$$\mathrm{sV}(C) = \frac{V(\mathfrak{P}_{\mathsf{SPR}})}{V(\mathfrak{P})} \quad \text{and} \quad \mathrm{gV}(C) = \frac{V(\mathfrak{P}_{\mathsf{GPR}})}{V(\mathfrak{P})}.$$

These causal volumes now express the fraction of MR policies which constitute SPR or GPR policies. Having an estimate of the causal volume of a predictor is an additional information about its quality. Consider, for example, a large network in which a message is sent from one node to another. The information that a component is a probabilistic cause for message to be lost in a lot of cases gives a good reason to predict the message loss whenever the component is part of the communication.

Example 11. For a smaller network example we again consider the MDP from Figure 2 where the undesired outcome $E = \{lost\}$ is to be predicted by $C = \{\mathsf{B}\}$. Here, we conclude from Example 8 that PR policies exist. Since C is a singleton the sets for SPR and GPR policies coincide and we have

$$\mathfrak{P}_{\mathsf{GPR}} = \{x \in [0,1]^{StAct} \mid x_{\mathsf{A},\alpha} < 1 \text{ or } x_{\mathsf{B},\beta} < 1\},$$

since only the MD policy choosing α and β does not constitute a PR policy. Thus, we also have $\mathrm{sV}(C) = \mathrm{gV}(C) = 1$ since almost all MR policies are both SPR and GPR policies. This means that from the perspective of probabilistic causality, the event $\Diamond\mathsf{B}$ is a good predictor for $\Diamond lost$. ◁

As mentioned in Section 3 the computation of $V(\mathfrak{P})$ can be done in polynomial time by representing $\mathfrak{P}$ as a product of simplices. However, in general the exact computation of the volume of a fully-dimensional polytope given in halfspace-representation such as $V(\mathfrak{P}_{\mathsf{SPR}})$ is #P-hard (Brightwell and Winkler 1991). Moreover, it has been shown that for any polynomial-time approximation there is a minimal gap between upper and lower bounds depending on the dimension of the problem (Bárány and Füredi 1987). With respect to these restrictions there are still efficient tools for the exact computation of polynomials over polytopes (De Loera et al. 2013). For $\mathfrak{P}_{\mathsf{GPR}}$ exact integration is even harder, since it is not a polytope by equation (g). However, since (g) is the only non-linear function restricting $\mathfrak{P}_{\mathsf{GPR}}$ this favors the usage of Monte Carlo integration algorithms (Press et al. 2007) to approximate $V(\mathfrak{P}_{\mathsf{GPR}})$.

4.2 Checking Probability-Raising Policies

While computing the causal volume for a given predictor and outcome is a difficult problem, we now want to address the existential query for probability-raising policies:

Given an MDP $\mathcal{M}$, a set of terminal states $E \subseteq \mathsf{S}$ and a predictor set $C \subseteq \mathsf{S} \setminus E$, is there a strict (resp. global) probability-raising policy for C and E in $\mathcal{M}$?

For this existence check we use a model transformation similar to the two-state MDP of $\mathcal{M}$ (Sec. 3). This then results in an end-component free MDP in which both C and E can only be visited once and there are exactly *four* terminal states corresponding to the entries of the confusion matrix. This will allow us to express these probabilities as reachability probabilities of single terminal states. We use the abbreviations from Figure 3.

Definition 12 (Canonical MDP). We transform the original MDP $\mathcal{M}$ to the *canonical MDP* $\mathcal{M}_{[C]}$ in the following way:

(i) All outgoing transitions from states $\mathsf{c} \in C$ are deleted and instead two fresh actions $\alpha_{\min}$ and $\alpha_{\max}$ are added. They transition to a new state TP with the minimal and resp. maximal probability to reach E from c. With the remaining probability they transition to a new state FP.

(ii) Collapse the maximal end-components (MECs) $\mathcal{E}$ of the resulting MDP into single states $\mathsf{s}_{\mathcal{E}}$ by taking the MEC-quotient, see e.g. (de Alfaro 1999).

(iii) Collapse the states of E to a fresh state FN and other terminal states to the fresh state TN. States $\mathsf{s}_{\mathcal{E}}$ representing MECs get the additional transition $P(\mathsf{s}_{\mathcal{E}}, \tau, TN) = 1$. The resulting MDP is $\mathcal{M}_{[C]}$.

In $\mathcal{M}_{[C]}$ we consider $E' = \{TP, FN\}$. ◁

Intuitively, the first step (i) of the transformation ensures, that each state $\mathsf{c} \in C$ is visited at most once, while still preserving all possible values for true positive and false positive predictions. In the second (ii) and third step (iii) the transformation gets rid of end-components by collapsing them into single states. Instead, the fresh action τ corresponds to the case that in the original MDP $\mathcal{M}$ a policy $\mathfrak{S}$ realizes a true negative by staying in an end-component $\mathcal{E}$ indefinitely. The soundness of the canonical MDP for a given set C with respect to **(G)** and **(S)** follows from the results of (Baier, Piribauer, and Ziemek 2024) and is given by the following Lemma 13 and Corollary 14.

Lemma 13 **($\mathcal{M}_{[C]}$ preserves Confusion Matrix).** *Given an MDP $\mathcal{M}$, a set of terminal states $E \subset \mathsf{S}$ and a set $C \subseteq \mathsf{S} \setminus E$, for each policy $\mathfrak{S}$ of $\mathcal{M}$ there is a policy $\mathfrak{U}$ of $\mathcal{M}_{[C]}$ and vice versa such that*

$$tp^{\mathfrak{S}}_{\mathcal{M}} = \Pr^{\mathfrak{U}}_{\mathcal{M}_{[C]}}(\Diamond TP), \qquad fp^{\mathfrak{S}}_{\mathcal{M}} = \Pr^{\mathfrak{U}}_{\mathcal{M}_{[C]}}(\Diamond FP),$$
$$fn^{\mathfrak{S}}_{\mathcal{M}} = \Pr^{\mathfrak{U}}_{\mathcal{M}_{[C]}}(\Diamond FN), \qquad tn^{\mathfrak{S}}_{\mathcal{M}} = \Pr^{\mathfrak{U}}_{\mathcal{M}_{[C]}}(\Diamond TN).$$

Corollary 14 ($\mathcal{M}_{[C]}$ **preserves Probability-Raising).** *Given an MDP $\mathcal{M}$, a set of terminal states $E \subset S$ and a set $C \subseteq S \setminus E$, a policy $\mathfrak{S}$ of $\mathcal{M}$ satisfies* **(S)** *(resp.* **(G)** *) for C and E in $\mathcal{M}$ iff the corresponding policy $\mathfrak{U}$ from Lemma 13 satisfies* **(S)** *(resp.* **(G)** *) for C and $\{TP, FN\}$ in $\mathcal{M}_{[C]}$.*

Checking for SPR Policies We now consider the following problem: Given an MDP $\mathcal{M}$ with a set of terminal states $E \subset S$ and a predictor set $C \subseteq S \setminus E$, is there a policy $\mathfrak{S}$ of $\mathcal{M}$ such that (SPR) holds for $\mathfrak{S}$? By the soundness of the canonical MDP w.r.t. the PR conditions (Cor. 14) and the confusion matrix (Cor. 14) we assume $\mathcal{M} = \mathcal{M}_{[C]}$.

For an SPR policy there needs to be a balancing between states $c \in C$ with a high value $p_{c,\max}$ and states $c' \in C$ with a low value $p_{c',\max}$. We will provide a characterization of SPR policies using this idea. For this, we consider the maximal probability to reach *TP* among all cause states:

$$p^\star = \max_{c \in C} p_{c,\max}$$

From $\mathcal{M}$ to $\mathcal{M}^\star$ we only change the actions in states $c \in C$. For each c the only enabled action in $\mathcal{M}^\star$ is δ with

$$P'(c, \delta, TP) = p^\star \text{ and } P'(c, \delta, FP) = 1 - p^\star$$

So, $\mathcal{M}^\star$ behaves as $\mathcal{M}$, but when a state $c \in C$ is reached the single enabled action leads to *TP* with $p^\star$. By construction, any policy $\mathfrak{U}$ of $\mathcal{M}^\star$ corresponds to a policy $\mathfrak{S}$ of $\mathcal{M}$.

Lemma 15 (Characterization of SPR Policies). *For an MDP $\mathcal{M}$ with set of terminal states $E \subset S$ and set of states $C \subseteq S \setminus E$ there is an SPR policy $\mathfrak{S}$ for C and E in $\mathcal{M}$ iff $\Pr^{\min}_{\mathcal{M}^\star}(\Diamond E) < p^\star$.*

Theorem 16 (Complexity for Checking SPR policy). *Given an MDP $\mathcal{M}$ with set of terminal states $E \subset S$ and set of states $C \subseteq S \setminus E$ the existence of an SPR policy for C and E in $\mathcal{M}$ can be decided in* P. *Moreover, a corresponding finite-memory (randomized) policy $\mathfrak{S}$ can be computed in polynomial time.*

Proof. We first note that we can transform $\mathcal{M}$ to $\mathcal{M}_{[C]}$ in polynomial time. We can now rely on the characterization of Lemma 15. So, by further transforming $\mathcal{M}_{[C]}$ to $\mathcal{M}^\star_{[C]}$ we can check whether $\Pr^{\min}_{\mathcal{M}^\star_{[C]}} < p^\star$ in polynomial time (Baier and Katoen 2008).

If this check is positive, an MD policy $\mathfrak{T}$ of $\mathcal{M}^\star_{[C]}$ with $\Pr^{\mathfrak{T}}_{\mathcal{M}^\star_{[C]}} = \Pr^{\min}_{\mathcal{M}^\star_{[C]}}$ can be derived. By (Baier, Piribauer, and Ziemek 2024)[Theorem 4.19] we can further derive the corresponding finite-memory (randomized) policy $\mathfrak{S}$ for $\mathcal{M}$ in polynomial time. $\square$

Remark 17 (Complexity for Singletons). The case where $C = \{c\}$ is a singleton can be decided more efficient. Note that there does not need to be a balancing between different states of C. Therefore, it is sufficient to only enable action $\alpha_{\max}$ in c. Let us call the resulting MDP $\mathcal{N}$. We then only have to check, whether $p_{\max,c} > \Pr^{\min}_{\mathcal{N}}(\Diamond E)$ as this corresponds to the best case SPR policy. $\triangleleft$

Checking for GPR Policies The next decision problem we consider is: Given an MDP $\mathcal{M}$, set of terminal states E and predictor set C is there a GPR policy $\mathfrak{S}$ for C and E in $\mathcal{M}$? By Lemma 13 we assume $\mathcal{M} = \mathcal{M}_{[C]}$ and use this to encode the inequality (GPR) in terms of state-action frequency variables. For this, we reformulate (GPR) in $\mathcal{M}_{[C]}$, using straight forward calculations, to

$$\Pr^{\mathfrak{S}}(\Diamond TP) \cdot \Pr^{\mathfrak{S}}(\Diamond TN) - \Pr^{\mathfrak{S}}(\Diamond FP) \cdot \Pr^{\mathfrak{S}}(\Diamond FN) > 0.$$

So, we consider this inequality in terms of frequencies together with the equations (S1)-(S3) (cf. Section 2):

$$x_{TP} \cdot x_{TN} - x_{FP} \cdot x_{FN} > 0. \qquad \text{(freq-GPR)}$$

Lemma 18 (Quadratic Program for GPR policy). *There is a GPR policy for C and E in $\mathcal{M}$ iff the system of inequalities given by (S1)-(S3) and (freq-GPR) has a solution.*

Proof. Let $\mathfrak{S}$ be a GPR policy for C and E in $\mathcal{M}$. By construction the corresponding frequencies of state-action pairs of $\mathfrak{S}$ are a solution to (S1)-(S3) and (freq-GPR).

Now, assume $x \in \mathbb{R}^{StAct}$ is a solution to (S1)-(S3) and (freq-GPR) and let $\mathfrak{S}$ be an MR policy corresponding to x. The inequality (freq-GPR) is equivalent to (GPR) in $\mathcal{M}_{[C]}$. As (freq-GPR) and (S1) hold, we have $x_{TP} \cdot x_{TN} > 0$ and thus $x_{TP} > 0$. Since $x_C = x_{TP} + x_{FP}$ we also get $\Pr^{\mathfrak{S}}(\Diamond C) > 0$. So, $\mathfrak{S}$ is a GPR policy for C and E in $\mathcal{M}$. $\square$

However, since (freq-GPR) is a strict inequality we can not directly use quadratic programming to decide the solvability (Vavasis 1990). We can still give an NP upper bound for deciding the existence of a GPR policy.

Theorem 19 (Complexity GPR Policies Check). *Deciding whether there is a GPR policy for C and E in $\mathcal{M}$ can be done in* NP.

Proof sketch. The proof uses distinct cases of (freq-GPR) within the set of possible state-action frequencies. It relies on the intermediate value theorem to show that, if there is a possible distribution of frequencies that forces (freq-GPR) to an equality, then there is an ε-neighborhood in which this becomes an inequality again. $\square$

By the result of (Baier, Piribauer, and Ziemek 2024)[Theorem 4.19] this existence check results in a finite-memory randomized GPR policy $\mathfrak{S}$ with exactly two memory cells.

5 Conclusion

In this paper, we proposed different measures to quantify the quality of predictors in adaptive systems. In the stochastic operational model of MDPs we considered predictors (and outcomes) as sets of states of the MDP. To this end, we proposed two approaches for measuring the effectiveness of the predictor. First, by averaging over the non-deterministic choices of the MDP we capture the effectiveness of the predictor to foresee undesired events in general. Second, we measure the fraction of policies that witness probability-raising causality of the predictor to the outcome, where we borrow ideas from recent work on probability-raising causality (Baier, Piribauer, and Ziemek 2024). For the proposed measures we provided insights on the complexity and discussed existing numerical methods for computing the respective measures.

Acknowledgments

The authors are supported by the DFG through the DFG grant 389792660 as part of TRR 248 and the Cluster of Excellence EXC 2050/1 (CeTI, project ID 390696704, as part of Germany's Excellence Strategy) and by BMBF (Federal Ministry of Education and Research) in DAAD project 57616814 (SECAI, School of Embedded and Composite AI) as part of the program Konrad Zuse Schools of Excellence in Artificial Intelligence.

References

Baier, C.; Funke, F.; and Majumdar, R. 2021. A Game-Theoretic Account of Responsibility Allocation. In Zhou, Z., ed., *30th International Joint Conference on Artificial Intelligence (IJCAI)*, 1773–1779. ijcai.org.

Baier, C.; and Katoen, J.-P. 2008. *Principles of Model Checking (Representation and Mind Series)*. The MIT Press, Cambridge, MA. ISBN 026202649X, 9780262026499.

Baier, C.; Piribauer, J.; and Ziemek, R. 2024. Foundations of probability-raising causality in Markov decision processes. *Logical Methods in Computer Science*, Volume 20, Issue 1.

Baier, C.; van den Bossche, R.; Klüppelholz, S.; Lehmann, J.; and Piribauer, J. 2024. Backward Responsibility in Transition Systems Using General Power Indices. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(18): 20320–20327.

Baldoni, V.; Berline, N.; de Loera, J. A.; Köppe, M.; and Vergne, M. 2011. How to Integrate a Polynomial over a Simplex. *Mathematics of Computation*, 80(273): 297–325.

Beckers, S.; Chockler, H.; and Halpern, J. Y. 2023. A Causal Analysis of Harm. arXiv:2210.05327.

Braham, M.; and van Hees, M. 2012. An Anatomy of Moral Responsibility. *Mind*, 121 (483): 601–634.

Brightwell, G.; and Winkler, P. 1991. Counting Linear Extensions. *Order*, 8: 225–242.

Bárány, I.; and Füredi, Z. 1987. Computing the Volume is Difficult. *Discrete & Computational Geometry*, 2: 319–326.

Chadha, R.; Sistla, A. P.; and Viswanathan, M. 2009. On the Expressiveness and Complexity of Randomization in Finite State Monitors. *J. ACM*, 56(5).

Chicco, D.; and Jurman, G. 2020. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC genomics*, 21(1): 1–13.

Chockler, H.; and Halpern, J. Y. 2004. Responsibility and Blame: A Structural-Model Approach. *J. Artif. Int. Res.*, 22(1): 93–115.

Chockler, H.; Halpern, J. Y.; and Kupferman, O. 2008. What causes a system to satisfy a specification? *ACM Transactions on Computational Logic*, 9(3): 20:1–20:26.

Ciesinski, F.; Baier, C.; Größer, M.; and Klein, J. 2008. Reduction Techniques for Model Checking Markov Decision Processes. In *2008 Fifth International Conference on Quantitative Evaluation of Systems*, 45–54.

Clarke, E. M.; Grumberg, O.; and Peled, D. 1999. *Model Checking*. MIT Press.

de Alfaro, L. 1997. *Formal Verification of Probabilistic Systems*. Phd thesis, Stanford University, Stanford, USA.

de Alfaro, L. 1999. Computing Minimum and Maximum Reachability Times in Probabilistic Systems. In Baeten, J. C. M.; and Mauw, S., eds., *10th International Conference on Concurrency Theory (CONCUR)*, volume 1664 of *Lecture Notes in Computer Science*, 66–81. Springer.

De Loera, J.; Dutra, B.; Köppe, M.; Moreinis, S.; Pinto, G.; and Wu, J. 2013. Software for exact integration of polynomials over polyhedra. *Computational Geometry*, 46(3): 232–252.

Duriqi, B.; Snopçe, H.; Salihu, A.; and Luma, A. 2024. An overview of parallel processing of rectangular determinant calculation. In *13th Mediterranean Conference on Embedded Computing, MECO 2024, Budva, Montenegro, June 11-14, 2024*, 1–7. IEEE.

Hall, N. 2004. Two Concepts of Causation. In Collins, J.; Hall, N.; and Paul, L., eds., *Causation and Counterfactuals*, 225–276. MIT Press.

Halpern, J. Y. 2015. A Modification of the Halpern-Pearl Definition of Causality. In *Proceedings of IJCAI'15*, 3022–3033. AAAI Press. ISBN 9781577357384.

Halpern, J. Y.; and Kleiman-Weiner, M. 2018. Towards Formal Definitions of Blameworthiness, Intention, and Moral Responsibility. In *AAAI Conference on Artificial Intelligence*.

Halpern, J. Y.; and Pearl, J. 2005. Causes and Explanations: A Structural-Model Approach. Part I: Causes. *The British Journal for the Philosophy of Science*, 56(4): 843–887.

Junges, S.; Torfah, H.; and Seshia, S. A. 2021. Runtime Monitors for Markov Decision Processes. In Silva, A.; and Leino, K. R. M., eds., *Computer Aided Verification*, 553–576. Cham: Springer International Publishing. ISBN 978-3-030-81688-9.

Kallenberg, L. 2020. *Lecture Notes Markov Decision Problems - version 2020*.

Kaltofen, E. L.; and Villard, G. 2005. On the complexity of computing determinants. *Comput. Complex.*, 13(3-4): 91–130.

Kozine, I.; and Utkin, L. V. 2002. Interval-Valued Finite Markov Chains. *Reliab. Comput.*, 8(2): 97–113.

Manna, Z.; and Pnueli, A. 1995. *The Temporal Logic of Reactive and Concurrent Systems: Safety*. Springer-Verlag.

Mascle, C.; Baier, C.; Funke, F.; Jantsch, S.; and Kiefer, S. 2021. Responsibility and verification: Importance value in temporal logics. In *2021 36th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, 1–14. Los Alamitos, CA, USA: IEEE Computer Society.

Namjoshi, K. S. 2001. Certifying Model Checkers. In *13th International Conference on Computer Aided Verification (CAV)*, volume 2102 of *Lecture Notes in Computer Science*, 2–13. Springer.

Pearl, J. 2009. *Causality*. Cambridge University Press, 2nd edition.

Powers, D. 2011. Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation. *Journal of Machine Learning Technologies*, 2(1): 37–63.

Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; and Flannery, B. P. 2007. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. USA: Cambridge University Press, 3 edition. ISBN 0521880688.

Sen, K.; Viswanathan, M.; and Agha, G. 2006. Model-Checking Markov Chains in the Presence of Uncertainties. In Hermanns, H.; and Palsberg, J., eds., *Tools and Algorithms for the Construction and Analysis of Systems*, 394–410. Berlin, Heidelberg: Springer Berlin Heidelberg. ISBN 978-3-540-33057-8.

Stein, P. 1966. A Note on the Volume of a Simplex. *The American Mathematical Monthly*, 73(3): 299–301.

Vavasis, S. A. 1990. Quadratic programming is in NP. *Information Processing Letters*, 36(2): 73–77.

Ziemek, R.; Piribauer, J.; Funke, F.; Jantsch, S.; and Baier, C. 2022. Probabilistic causes in Markov chains. *Innovations in Systems and Software Engineering*.