TECHNISCHE UNIVERSITÄT DRESDEN

Faculty of Computer Science Institute of Theoretical Computer Science

Chair of Algebraic and Logical Foundations of Computer Science

Probabilistic Causality in Markovian Models

Robin Ziemek

Dissertation

to achieve the academic degree

Doktor der Ingenieurwissenschaften (Dr. ing.)

Referee

Prof. Dr. Christel Baier

TECHNISCHE UNIVERSITÄT DRESDEN

Faculty of Computer Science Institute of Theoretical Computer Science

Chair of Algebraic and Logical Foundations of Computer Science

Abstract

The complexity of modern computer and software systems still seems to grow exponentially, while the human user is widely left without explanations on how to understand these systems. One of the central tasks of current computer science therefore lies in the development of methods and tools to build such an understanding. A similar task is addressed by formal verification which gives various verifiable justifications for the functionality of a system. As these still only give knowledge that a system functions properly they only address a portion of the task to make systems easier to comprehend. It is widely believed that cause-effect reasoning plays an important role in forming human understanding of complex relations. Thus, there are already many accounts on causality in modern computer science. However, most of them are focusing on a form of backward looking actual causality. This variant of causality is concerned with actual events after their occurrence and tries to reason about causes mostly in a counterfactual manner.

In this thesis we address a probabilistic form of causality which is forward looking by nature. As such, we define and discuss novel notions of probabilistic causes in discrete time Markov chains and Markov decision processes. For this we rely on two central principles of probabilistic causality. On one hand, the *probability-raising principle* states that a cause should raise the probability of its effect. On the other hand, *temporal priority* requires that a cause must occur before its effect. We build the mathematical and algorithmic foundations of our so called *probability-raising causes*. For this we work in a state-based setting where causes and effects are reachability properties of sets of states. In order to measure the predictive power of states we define quality-measures for which we interpret causes as binary classifiers. With these tools we address the algorithmic questions of checking cause-effect relations if both a cause candidate and an effect are given and finding quality-optimal causes if only the effect is given. We discuss possible extensions of this basic state-based framework to more general formulations of causes and effects as ω-regular properties.

Acknowledgments

I sincerely want to thank my supervisor *Christel Baier* first and foremost. I especially want to thank you for believing in me when I started out with only a Master degree in Mathematics and little to no experience in computer science. Thanks to your guidance and experience I was able to find a place among the researchers for theoretical computer science.

I would also like to thank all my colleagues at the chair for Algebraic and Logic Foundations of Computer Science at TU Dresden for creating a healthy and lively atmosphere for research and working in general, even though my first years were heavily influenced by the global pandemic of Covid-19. In particular I want to thank my colleague and (kind of) supervisor *Jakob Piribauer* for our many discussions on Probability-Raising Causality in MDPs and for his many contributions to this topic. I also want to thank my colleagues *Sascha Klüppelholz* and *Max Korn* for their invaluable high-level view to many of my problems. Your comments were immensely helpful with starting and finishing the actual thesis.

I am indepted to my colleagues in *CPEC* for providing countless opportunities for scientific exchange by providing talks, seminars and regular trips to Saarbrücken. In the same breath, I am very grateful to the DFG for supporting CPEC and in turn my doctorate studies financially.

Of course, there are many people outside of my direct research which have also heavily influenced my development as a researcher. In particular, I want to thank Prof. *Ulrich Krähmer* from the Chair of Geometric Methods in Mathematics at TU Dresden who has patiently supervised me for both my Bachelor and Master studies. The energy and skill you brought to every lecture and meeting has inspired my to actually stay in research after my Masters degree and pursue a Doctoral degree.

It is also impossible for me not to mention my very good friend and (past) fellow student *Frederic Dörband* here. I would have never been able to start a doctoral study, let alone finish a Dissertation, without you having my back in research and in life. Funnily enough, you gave me the opportunity to meet Christel in the first place.

Moreover, there is a life outside of the university and research. I am extremely grateful for the support and love I received from all my friends and family. A special thanks i want to dedicate to my girlfriend *Alicia* for always listening to my loud thoughts on my research related topics (and even giving valuable comments ever so often).

Contents

Abstract 3

1 Introduction 7
1.1 Contributions and Outline . . . 15
1.2 Related Work . . . 17

2 Preliminaries 21
2.1 Markov Decision Processes . . . 22
2.1.1 Graph of an MDP . . . 22
2.1.2 Schedulers and Probability Measure . . . 23
2.1.3 Maximal and Minimal Probabilities . . . 25
2.1.4 Frequencies and Balance Equations . . . 26
2.1.5 MR Scheduler in MDPs without ECs . . . 26
2.1.6 MEC-Quotient . . . 27
2.2 Automata and ω-Regular Languages . . . 30
2.3 Probability-Raising Causality . . . 31

3 Probability-Raising Causality in Markov Decision Processes 33
3.1 Setting and Definition . . . 33
3.1.1 Related Work revisited . . . 38
3.2 Examples . . . 42
3.2.1 Conceptual Examples . . . 42
3.2.2 Racetrack MDP . . . 47
3.3 Quality Measures for Probability-Raising Causes . . . 52
3.3.1 PR Causes as Binary Classifiers in MDPs . . . 53
3.3.2 Quality Measures for a given PR Cause . . . 54

4 Algorithmic Considerations 59
4.1 Checking the Probability-Raising Conditions 60
4.1.1 Canonical MDP for given Cause Set 61
4.1.2 Checking the SPR Condition and the Existence of PR Causes 72
4.1.3 Checking the GPR Condition 77
4.2 Computing the Quality of a PR Cause 88
4.2.1 Max/Min Ratios for Disjoint Sets of Terminal States 92
4.3 Quality-Optimal Probability-Raising Causes 97
4.3.1 Quality-Optimal SPR Causes 98
4.3.2 Quality-Optimal GPR Causes 108

5 Variants of Probability-Raising Causality in MDPs 115
5.1 Probability-Raising Scheduler and Potential PR Causes 116
5.1.1 Checking for SPR Scheduler 119
5.1.2 Checking for GPR Scheduler 124
5.2 ω-Regular Effects - Reachability Causes 127
5.2.1 Checking Reachability Probability-Raising Causes 128
5.2.2 Quality and Optimality of Reachability PR Causes 133
5.3 ω-Regular Co-Safety Causes 137
5.3.1 Checking Co-Safety PR causes 139
5.3.2 Quality and Optimality of Co-Safety PR Causes 142

6 Conclusion 147

1 Introduction

Modern software and hardware systems are highly complex and therefore they are typically only understood by developers or domain experts. With the advancement of AI and automated decision making [Sco23] it is also evident that the development of such systems is so fast that it is very difficult even for a well-read user to comprehend the intricate behavior of such systems. On the other side, a lot of processes are now automated and even decisions over potential life-threatening behavior are sometimes in the hand of software systems [BCH23; New22]. For example, tools which enable autonomous driving for cars are already widely implemented in new generations of automobiles. Supporting systems like lane departure assistants or automated spacing to other cars have been around for over 20 years now. However, for a human driver it is not always clear why the car is autonomously slowing down even against the control of the driver. Naturally, the gap between complexity and comprehension of modern systems is bringing about a lot of challenges in areas like legislative and jurisdiction [MNH21] as well as trusting and working with automated systems [Mil17]. If, for example, a car causes an accident by some autonomous decision where people are harmed, it is important for the jurisdiction to know and understand *how* and *why* such a decision was made by the car in order to allocate blame. On the other side, people harmed by such autonomous behavior desire explanations for these decisions.

Addressing this gap is a huge undertaking and one of the major tasks of modern computer science, as can be seen by the many current research-projects addressing variations of this topic (for example consider the DFG Project CPEC [Cpe]). The area of formal verification already addresses some of the mentioned problems as it aims to prove the correctness of a system with respect to some specification. The results are thus often simple yes-no answers to the question whether the system behaves as specified or not. Even though these results give guarantees on the behavior of a system, this is not satisfactory for enabling understanding of the inner workings of a system. However, problems arising from this have been addressed and studied extensively by introducing counterexamples, witnesses and invariants

(see e.g. [MP95; CGP99; Nam01; FJB20]). More than just a guarantee, these certificates give the user a verifiable justification that a system does or does not behave according to the specification.

Though, such a verifiable certificate can help make the system more comprehensible for the user, it fails to explain the behavior of the system in an accessible manner. Formulated in epistemic terms, such results from formal verification provide *knowledge that* the system satisfies a specification (or not). This comes by giving an assertion (the yes/no output of the verification) and justification (the witness or counterexample) to increase the belief in the result. However, this process usually does not provide *understanding* of why the system behaves that way. This establishes the need for a more accessible understanding of *why* and *how* a system behaves the way it does.

Causality in Computer Science

A common ground found by many researchers is that the concept of cause and effect is one of the most important foundations for understanding complex relations. For example, when humans react to and work with problems, their decisions are strongly influenced by familiarity with the problem and their available causal information [Zhe+20]. Furthermore, the analysis of cause-effect relations have also been identified to be helpful in a variety of areas ranging from finance [Kle11] to medicine [KH11] and aeronautics [Ibr+20]. In a societal context, causality has a strong role in determining moral responsibility [CH04; CHK08; BH12; HIH23] or legal accountability [Fei+11], and ultimately fosters user acceptance by providing an increased level of transparency [Mil17]. In computer science, methods building on causality are already used to derive explanations [Gro+06; Bee+12; Coe+22a].

The topic of causality has been extensively researched for centuries going back to the first formulations by Aristotle [CR21]. Among the abundance of work is the very influential formulation of Hume [Hum39; Hum48] who defined causality based on counterfactual reasoning stating that causality means that

> "*if the first object had not been, the second never had existed*" [Hum48, Section VII].

Consider, for example, a setting where two cars driving in the same lane collide after the car in front suddenly stops. In such a scenario one would argue, if the front car would not have stopped, then both cars would not have collided.

Centuries later the ideas of Hume were picked up by Lewis [Lew73] resulting in a more structured approach to counterfactual reasoning as a basis for causality using a definition of "closest" possible worlds. If in the closest possible worlds, where the cause C does not happen, the effect E also does not happen, then C is indeed a cause of E. In the above example, in a world where the car in front does not rapidly slow down, the collision does not

happen. This idea proves very influential. Even in modern computer science current papers on causality in formal systems are still basing their notion of causality on it [PPB23; FS23].

However, arguably the most influential work bridging from causality to computer science is the work of Halpern and Pearl [HP01] which uses causal models based on structural equations [Pea09] to define actual causes using the idea of counterfactual reasoning. This definition was improved and modified over the years in multiple accounts [HP05a; HP05b; Hal15]. A lot of work in formal systems stems from this formalization ranging from responsibility, blame and harm [CH04; CHK08; BCH23] over fault localization [KLL11; LL13; Bee+15] to explanations [Bee+12; Coe+22a]. Moreover, foundational work also addresses the algorithmic complexity of checking various notions of actual causality [GM13; CGL18; IP20; Coe+22b].

What all of these accounts on causality have in common is that a form of *actual causality* (also called token causality [Kle12]) with a *backwards looking* perspective is used. Typically, they are given an actual execution of the system or a counterexample to a property and then analyze what caused the execution to bring about the erroneous behavior. Therefore, the behavior of the system as a whole is not in question, but rather more specifically the given actual execution. For example, after an autonomous car had an accident, actual causality considers the history of the car up until the time of the accident. This would find that a temporary power shortage has caused the spacing sensors to malfunction leading the car to abruptly stop such that a second car collides with it.

Another branch of causality therefore looks at general connections and correlations between events to find a notion of causality which is *forward looking*. Often, this means looking at the situation as a whole and then infer causality between possible events. This kind of causality is also called *type causality* [Hal16]. For the above example of the car collision, one would find that generally power shortage can lead to malfunctions of sensors which then may cause collisions. Such an analysis then enables the manufacturers to design the car accordingly such that safety risks like these are taken care of.

Example 1 (Actual Causality vs. Type Causality). For an in depth comparison of actual and type causality we turn to another example. Suppose, there are two persons, Billy and Suzy, throwing rocks at a bottle of glass. This example has been widely discussed for in literature of causality, especially in terms of *actual causality* [Hal04; FG11; PH13; Hal15]. We show our own variant of this in Figure 1.1 where we depict different states of the process by labeled circles and a change of states with an arrow. Readers familiar with MDPs may identify Figure 1.1 with an MDP.

Here, a process decides probabilistically whether Suzy or Billy are allowed to throw a rock at the bottle or whether the experiment ends. The respective probabilities are found on the labeled arrows in Figure 1.1. After this does not choose to end the experiment, the chosen person can now decide whether they take the chance and throw the rock (label $\mathfrak{t}$) or wait (label w) and thus restart the decision process. This is shown by the corresponding labeled

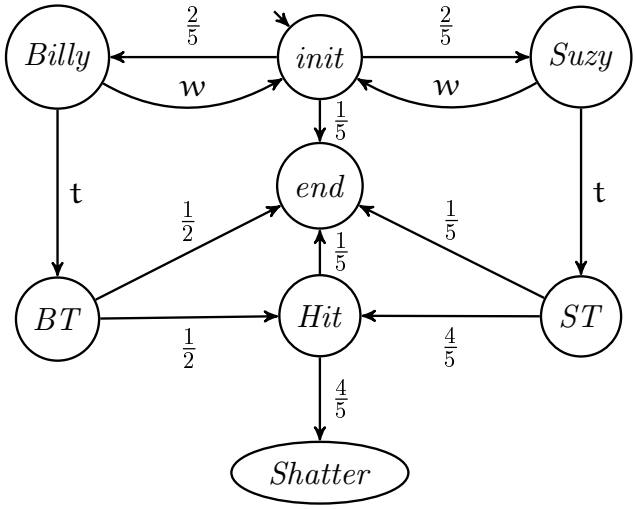

Figure 1.1: Modeling a variant of the rock throw example

arrows. If someone decides to throw their rock, the process changes its state (ST for "Suzy throws" and BT for "Billy throws"). However, Suzy is more accurate than Billy and therefore hits the bottle with a higher probability. This is denoted by the labeled probabilities on the arrows. If the bottle was hit it has a high chance to shatter. After a throw, the experiment will not be repeated independent on whether or not the bottle shatters.

Consider the case where both Billy and Suzy will choose to throw the rock whenever they are allowed to. (This corresponds to a scheduler $\mathfrak{S}$ always choosing $\mathfrak{t}$.) If the process chooses Billy then he decides to throw his rock. Let us assume the case where Billy actually hits the bottle and the bottle shatters. In this scenario it would be natural to say that Billy's decision to throw the rock actually caused the bottle to shatter.

However, a probabilistic analysis gives us that the probability of shattering under the condition that Billy throws is lower that the overall probability of shattering in the specified situation where both persons will take throw given the chance. Formally, this is reflected by

$$\Pr^{\mathfrak{S}}(\Diamond \mathit{Shatter} \mid \Diamond BT) = \frac{1}{2} \cdot \frac{4}{5} = \frac{50}{125} < \frac{52}{125} = \Pr^{\mathfrak{S}}(\Diamond \mathit{Shatter}),$$

where the left hand side is the conditional probability of the bottle shattering, given that Billy throws and the right hand side is the overall probability of the bottle shattering in the given situation. In other words, the decision of Billy to throw his rock lowered the probability of the bottle shattering since Suzy is a better thrower and would "probably" be chosen by the process afterwards. Therefore, in the modeled system the bottle has already a high chance to break anyway. Here, we see the difference between *actual causality* and *type causality*.

Actual causality argues about events which have actually passed. As we discussed above, a prominent approach for actual causality is counterfactual reasoning [HP01]. In our example one would argue that if Billy would not have thrown the rock, the bottle would note have been shattered.

Type causality (sometimes called general causality [Hal16]) on the other hand argues about the whole system and tries to find causes, often in a forward looking manner. One would generally argue that Billy throwing a rock at the bottle does overall lower the chances of the bottle shattering. Thus, Billy's throw is not a general cause for the bottle to shatter since in the experiment, set up with the intend of Suzy and Billy to throw, the bottle already shatters with a higher probability. For a very in depth distinction between these two variants of causality we refer to the introduction of [Eel91, p. 8-20]. ◁

A very common approach to type causality comes from using the uncertainty of looking into the future as a probabilistic relation. Since the discovery of quantum mechanics in physics has led to the abandon of deterministic world views in philosophy, this has also strengthened the arguments for probabilistic relations between events [Sup70]. Therefore, probabilistic causality has formed as an alternative way to define cause-effect relations. At the heart of probabilistic causality lies the idea that causes raise the probability that the effect occurs. This *probability-raising* principle has famously been formalized by Reichenbach [Rei56] using conditional probabilities in the following way.

$$\textit{For a cause } \mathrm{C} \textit{ and the effect } \mathrm{E} \textit{ we have } \mathrm{P}(\mathrm{E} \mid \mathrm{C}) > \mathrm{P}(\mathrm{E} \mid \neg \mathrm{C}). \tag{C1}$$

Here, $\mathrm{P}(\mathrm{E} \mid \mathrm{C})$ is the probability that E occurs under the condition that C occurs and $\mathrm{P}(\mathrm{E} \mid \neg \mathrm{C})$ is the probability that E occurs under the condition that C does not occur. For example, if we say C is describing a power shortage in an autonomous car, then we would argue that this makes it more likely for a car to collide with other cars E. Thus, power shortage in an autonomous car *raises* the probability of a collision. However, this does not mean that every time there is a power shortage in a car it will have a collision. It may also happen that the car will just safely stop, for example if an experienced human driver interferes or if the safety mechanisms use a separate power circuit. Rather, there are more cases where a power shortage can lead to a collision then there are cases where the absence of a power shortage leads to a collision.

By nature, probabilistic causality is forward looking as it considers the probabilities that events *will* occur. Moreover, it can also be seen as a weaker variant of counterfactual reasoning. Instead of stating that E would not have happened if C would not have happened, it states:

If C *would not have happened then the probability that* E *happens would be lower.*

The probability-raising principle on its own is obviously not satisfactory for a characterization of what makes a cause. For instance, it is symmetric in E and C in the sense that the roles of E and C in (C1) can be switched by

$$\mathrm{P}(\mathrm{E} \mid \mathrm{C}) > \mathrm{P}(\mathrm{E} \mid \neg \mathrm{C}) \qquad \text{if and only if} \qquad \mathrm{P}(\mathrm{C} \mid \mathrm{E}) > \mathrm{P}(\mathrm{C} \mid \neg \mathrm{E}).$$

This is one of the reasons why research in probabilistic causality has also taken many different approaches. Reichenbach himself has already formulated the principle of the *common cause* which can deviate the probabilities of two events such that it seems that one caused the other according to probability-raising [Rei56]. In the above example where an autonomous car abruptly stops, leading to a collision, we evaluated that a temporary power shortage was a cause for the collision. However, the temporary power shortage also increases the probability that the display inside of the car will flicker. Just judging from the probabilities it might seem as if the flickering of the display also raises the probability of a collision.

Since a sound notion for probabilistic causality needs more than the probability-raising principle (C1), another central feature of causality agreed upon by many philosophers is that

Causes must happen before *their effects.* (C2)

This is called *temporal priority* of causes and is sufficient to resolve the symmetry between cause and effect in (C1). This temporal aspect of causality is also very common in everyday language. Following Reichenbachs account, Suppes defines *prima facie causes* which include probability-raising and temporal priority of causes [Sup70] . Furthermore, Suppes defines spurious causes which only change the probability of the effect in a small way, as well as direct and supplementary causes. However, in [Car79] it is noted that probability-raising can also be counter intuitive in sub-populations by using Simpson's paradox [Sim51]. We will discuss Simpsons paradox thoroughly in a later example (Ex. 23). These and more paradox examples have been at the center of discussions about probabilistic causality [Sky81], leading to some accounts dealing with such problems by considering background contexts to distinguish between spurious relations and genuine causes [ES83; Eel91]. For a more detailed overview about probabilistic causality we refer to [Hit21].

Where the counterfactual approach to causality has been taken up by theoretical computer science by multiple accounts, work on probabilistic causality in formal systems is scarce [BFM21]. There has been work on using probabilistic causality in Markov chains and temporal logic by Kleinberg et. al. which focuses on using probabilistic causes in time series data by computing causal significance [KM09; Kle11; HK15b]. In [KM10] the type-level probabilistic causes are used to infer token-level causes. It has been noted that probabilistic causality can be expressed in temporal hyper-logic for the probabilistic model of Markov chains [ÁB18] and for the probabilistic model with non-determinism of Markov decision processes [DFT20]. Here, the model checking problem of these logics is in PSPACE for HyperPCTL in [ÁB18] and generally undecidable for PHL in [DFT20]. This motivates the goal of our research leading to this thesis:

The goal of this thesis is to define a notion of probabilistic causality in Markovian models and to investigate the complexity of checking and inferring this causality.

Approach

As the formulated goal suggests this thesis mainly focuses on the foundational aspects of establishing a concept of probabilistic causality in Markov decision processes (MDPs), which are formal models for probabilistic systems that combine both non-deterministic choices and probabilistic choices. Typically, it is interpreted that the non-deterministic choices are made by a user or a controller as an input to the system, whereas the probabilistic choices come from the influence of the environment. For example an autonomous car could be depicted as an MDP where the decisions made by the system, e.g. accelerating or slowing down, are non-deterministic but the actual outcomes depend on probabilistic behaviors, like the stopping distance is influenced by the interplay of underground and wheels.

So, one goal is to formalize the concepts (C1) and (C2) for MDPs. In order to tackle this, we start in a basic setting where we focus on reachability properties as the considered events. Thus, the effect and the cause are represented by sets of states in the MDP. In order to take temporal priority of causes (C2) into account, we consider a setting where the effect states are terminal (also called absorbing) after which the system model stops. Intuitively, these events occur whenever the system model reaches a corresponding state. For example, the effect can be a faulty behavior like sensor failure which is described by reaching an undesired failure state. Later, we apply a natural extension of this framework when we consider the effect to be an ω-regular event while causes are still either state-based reachability properties or ω-regular co-safety path properties. This allows for more general formulations of causality.

Whenever we talk about some event being a probability-raising (PR) cause, we want that the PR condition (C1) always holds, no matter how the non-determinism is resolved. Intuitively, it should not make a difference to the cause-effect relations if we change the input to our system. For example, no matter what the driver or the autonomous controller of a car does, when there is a temporary power shortage, this is generally considered as a probabilistic cause for collision. This means the PR condition (C1) needs to hold for all scheduler of the MDP. Furthermore, we consider two natural ways to interpret condition (C1):

> On one hand, the probability-raising principle is required locally for each element of the cause. So, whenever an execution of the system reaches a state of the cause, the probability is strictly raised, which results in a strict property. Such causes are called *strict probability-raising (SPR) causes* in our framework. Such an interpretation is suitable when the causal analysis is in a preventive manner, e.g. when states shall be identified which are to be avoided in order to prevent a faulty behavior.
>
> On the other hand, we can treat the cause as a unit and thus consider the PR principle in a global manner leading to the notion of *global probability-raising (GPR) cause*. This way, the causal analysis can be done without considering individual elements or executions to be causal. Considering the cause as a whole is also more natural as it takes

the different ways in which the cause can occur into account by weighting them by the distributions the cause states are reached with. Such interpretation is suitable when causality is used to find faulty components of the system as opposed to individual system states. This comes from the synergy of GPR causes with further constraints on cause candidates. For example, the set of non-terminal states of the MDP may be partitioned into sets of states S_i which are under control of different agents. Also, GPR causes are more appropriate when causes are used for monitoring purposes where the system in question can only be observed partially as then the cause candidates can be considered as sets of indistinguishable states.

The differing concepts of strict and global probability-raising causes, however, do not help to distinguish between the predictive powers of different causes with respect to an effect. For example, in an MDP that models an autonomous car in a very detailed fashion, a state that describes a temporary power shortage due to a very specific and extreme reason (e.g. because the fuel of the car caught fire) is by itself an SPR cause for a crash. Even if this cause very precisely predicts a crash, the cause-effect relation is too specific to predict a crash in many cases. In Markov decision processes the predictive coverage coincides with how well the executions exhibiting the cause cover executions showing the effect. However, solely focusing on a broader coverage can also lead to very trivial causality relations. For example, starting the car increases the probability that the car will crash, but this cause-effect relation is too general to have much value in a causal analysis for the effect of crashing.

In order to take this trade off between accurate prediction and coverage into account, we take inspiration from measures for binary classifiers in statistical analysis (see e.g. [Pow11]). For this, we consider causes as binary classifiers that classify an execution as causal as soon as the cause is reached. This classification can be true or false depending on whether or not the effect occurs after the cause. So, we define different quality measures for causes, which allows us to compare different causes and also to look for quality-optimal causes. While addressing general quality measures, we also consider three quality notions in particular:

- The *precision*, which is the probability that the effect indeed occurs after the cause,
- the *recall*, which is the probability that the effect is indeed preceded by the cause,
- and the *f-score*, which is the harmonic mean between precision and recall. As precision and recall are opposite ends of the spectrum between accurate prediction and coverage, the f-score aims at finding a good trade off between these qualities.

For our resulting causality notions there are always two natural underlying questions:

1. Given two events C and E, does C cause E to happen?
2. Given an event E, which events C cause E to happen?

These central questions show up in different flavors in our algorithmic considerations. On one hand, we consider how to check the GPR and SPR causality when both events cause and effect are given. This already ties into the second question which would then be to check whether SPR or GPR causes exist for a given effect. On the other hand, we address the computation of the quality measures for a given cause and its effect. This basically concerns how strong of a cause $\mathbf{C}$ is for $\mathbf{E}$. Here the second question is more of a central one, as we address the computational problem of computing quality-optimal causes for a given effect. This means we look for "good" causes for $\mathbf{E}$.

1.1 Contributions and Outline

As mentioned before, the contributions of this thesis are naturally on a level of foundational research as we build mathematical and algorithmic foundations for a probabilistic notion of causality in Markov decision processes (MDPs). The most contributions therefore come by the algorithmic investigations on how to check causality and how to compute causes.

1. Inspired by influential work on probabilistic causality in the philosophical literature, we define two novel notions of probabilistic cause in MDPs based the probability-raising principle (C1) and temporal priority of causes (C2). As such we consider a basic setting where we represent events as reachability properties for sets of states which then represent both cause and effect and discuss the differences between both notions in Section 3.1. We thoroughly discuss our notions of strict and global PR cause using both examples inspired by philosophical accounts on probabilistic causality and examples used in theoretical computer science in Section 3.2. We also discuss a slight variation of our causality notion which only considers potential probability-raising causes which are dependent on the scheduler of the MDP in Section 5.1.

2. We consider PR causes as binary classifiers which aim to predict the effect leading to the notion of quality measures in Chapter 3.3. These can be used to compare different coverage and prediction properties of PR causes. In particular we consider worst-case scenarios for these measures which is in line with the fact that cause-effect relations always hold independently of which scheduler is chosen for the MDP.

3. We investigate the computational problem of checking cause-effect relations for given sets of states as effect and cause as well as checking the existence of PR causes for a given effect in Section 4.1. For this, we first describe a polynomial time model transformation of the MDP into a canonical causal form, which has the advantage that cause states can only be reached once and that the MDP has no non-trivial end-components. We then check the strict probability-raising condition directly in Section 4.1.2. This algorithm can then also be used to check the existence of PR causes for a given effect.

The global probability-raising condition is checked in Section 4.1.3 by formulating a quadratic system of inequalities using state-action frequencies of the canonical causal form of the MDP.

4. In Section 4.2 we give algorithms for computing the quality measures for given sets of states as effect and cause candidate. Here, we again make use of the canonical causal form of the MDP. In particular we reduce the computation of a quality-measure to the computation of either a conditional probability or a ratio of reachability probabilities for two disjoint sets of states.

5. We investigate the algorithmic complexity for computing quality-optimal causes for a given set of effect states in Section 4.3. It turns out that there are big differences between the strict- and the global PR causality condition as a basis. By exploiting the fact that SPR causes are stable under union we give algorithms to compute quality-optimal SPR causes. We show NP-hardness for the threshold problem of whether there is a GPR cause with a quality-measure above a given threshold by reductions from the Knapsack problem. Algorithms for computing quality-optimal GPR causes then rely on brute-force using the computation algorithms for quality-measures.

6. We introduce and discuss the notion of potential PR causes arising from a switch of quantification over scheduler in Section 5.1. For this, we split off the scheduler from the notion of PR cause to define probability-raising schedulers. These are schedulers for which either a strict or a global PR condition holds. When checking whether there is a scheduler for which the SPR condition holds in Section 5.1.1, we need different techniques as introduced in Chapter 4. For checking the existence of a witnessing GPR scheduler addressed in Section 5.1.2, we can mostly rely on established methods but need other techniques of proving correctness.

7. We extend the established state-based framework of PR causes in MDPs to ω-regular effect scenarios in Section 5.2. We algorithmically investigate the resulting notion of *reachability PR causes* by a reduction to the state-based case using a product MDP construction with a deterministic automaton representing the ω-regular effect. A further extension where the cause is considered as a ω-regular co-safety property is considered in Section 5.3. Here, a reduction via product MDPs is still possible but shows additional complications especially with respect to the temporal priority (C2).

Throughout the thesis we also discuss the special case of Markov chains, where no non-deterministic choices are present. We conclude the thesis in Chapter 6, where we also address a cost-oriented optimality criterion for causes in Markov chains as in [Zie+22].

Related Publications

The thesis in mainly based on the following peer-reviewed publications.

[Bai+22b] Christel Baier, Florian Funke, Jakob Piribauer and Robin Ziemek. "On probability-raising causality in Markov decision processes". In: *Foundations of Software Science and Computation Structures*. 2022, pp. 40–60. ISBN: 978-3-030-99253-8. DOI: 10.1007/978-3-030-99253-8_3.

[BPZ24] Christel Baier, Jakob Piribauer and Robin Ziemek. "Foundations of probability-raising causality in Markov decision processes". In: *Logical Methods in Computer Science* Volume 20, Issue 1 (Jan. 2024). DOI: 10.46298/lmcs-20(1:4)2024.

These publications are strongly influenced by our prior work on probabilistic causality in Markov chains.

[Bai+21b] Christel Baier, Clemens Dubslaff, Florian Funke, Simon Jantsch, Rupak Majumdar, Jakob Piribauer and Robin Ziemek. "From Verification to Causality-Based Explications (Invited Talk)". In: *48th International Colloquium on Automata, Languages, and Programming, (ICALP)*. Ed. by Nikhil Bansal, Emanuela Merelli and James Worrell. Vol. 198. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021, 1:1–1:20.

[Bai+21c] Christel Baier, Florian Funke, Simon Jantsch, Jakob Piribauer and Robin Ziemek. "Probabilistic Causes in Markov Chains". In: *Automated Technology for Verification and Analysis*. Ed. by Zhe Hou and Vijay Ganesh. Springer International Publishing, 2021, pp. 205–221. ISBN: 978-3-030-88885-5.

[Zie+22] Robin Ziemek, Jakob Piribauer, Florian Funke, Simon Jantsch and Christel Baier. "Probabilistic causes in Markov chains". In: *Innovations in Systems and Software Engineering* (2022). ISSN: 1614-5054.

1.2 Related Work

Before we dive into some related work to our account of probability-raising causality, we like to point to a very in depth summary of different approaches to causality in computer science given in [Bai+21b].

Actual Causality. As mentioned before, a lot of work in theoretical computer science has focused on actual causality, where most approaches draw more or less direct inspiration from the influential definitions of actual cause given by Halpern and Pearl [HP01; HP05a; Hal15].

In a series of work Chockler et al. have extended these definitions of actual causality to responsibility and blame [CH04; CHK08] and (more recently) harm [BCH23]. In [CHK08] the structural model approach to causality of [HP05a] is applied to coverage metrics in formal systems to introduce a notion of responsibility. Here, responsibility of a variable is basically the number of changes which have to be made to a variable in order to make it a counterfactual cause for a property. Similarly, a notion of harm based on actual causality [Hal15] is given in [BCH23] by using utility functions.

In [KLL11] the authors use fault trees to represent actual causes in the spirit of Halpern and Pearl. They introduce a method to derive fault trees from probabilistic counterexamples. This work is extended in [LL13] using Event Order Logic on transition systems to give a method for fault localization, which was further improved and extended in [Bee+15] and [CGL18].

In [Bee+12] explanations for counterexample traces to temporal logic formulas are derived using the Halpern, Pearl notion of actual causality [HP05a] as interventions on traces.

By considering actual causality as a hyperproperty of a moore machine, which is a basic trace-based system model, Coenen et al. [Coe+22b] show that they can decide causality as a HyperQPTL model checking problem. Moreover, in [Coe+22a] they show that actual causality can be used to explain violations of hyperproperties formulated in HyperLTL.

Very recent work has taken up on Lewis' definition of causality using notions of closest possible worlds to define actual causes [Lew73]. Parreaux et al. [PPB23] use distance metrics to define closest possible worlds for traces of both transition systems and 2-player games and Finkbeiner and Sieber [FS23] extend the notion of total-similarity between worlds given by Lewis' to introduce QPTL with counterfactual conditionals.

Probabilistic Causality. In contrast to actual causes, probabilistic forms of causality in formal systems are not considered as broadly. In a series of work on probabilistic causality Kleinberg et al. define *prima facie causes* in Markov chains and give a measure for their causal significance [KM09; KM10; Kle11; KH11; HK15b; MK16; ZK17; ZCK18]. However, this work mostly focuses on deriving causes and their significance from time-series date with applications in law and medicine. We will compare our definition for probabilistic causality to their definition of prima facie cause [KM09] in Section 3.1.1.

It was also noted that probabilistic causality based on the probability-raising principle can be expressed as a hyperproperty. Abraham and Bonakdarpour define the probabilistic hyperlogic HyperPCTL working over Markov chains in [ÁB18], where in an example [ÁB18, Example 3.5] they show how to formulate probabilistic causality as a HyperPCTL formula. Similarly, Dimitrova et al. define another probabilisitic hyperlogic PHL working over Markov decision processes in [DFT20]. Here another example is given on how to define probabilistic causality based on actions [DFT20, Section 3.1]. Again, we will compare these examples to our definition in Section 3.1.1.

In [KP23] the authors define a temporal logic with counterfactual and interventional operators which then works over Markov decision processes. They use a structural causal model translation of the MDP by unrolling it up to a specific time point. Thus, they rely on counterfactual causality in a probabilistic model by allowing model manipulations.

The recent publications [Bai+21a; Bai+22a] use the principles of sufficiency and necessity are as a basis to define a notion of causality. We also encounter sufficient or necessary causes in our work as stated in Section 3.3.

2 Preliminaries

In this chapter, we collect notions and concepts which are known from literature and which we use when building our theory of probability-raising causality in Markovian models. For the most part, we only give definitions and do not go into detail about properties but rather give references for further reading. Throughout the thesis, we assume the reader is familiar with the concepts introduced in this chapter.

Notations

For a set S we denote the cardinality of S by $|S|$ and the power set of S by 2^S. We will denote the natural numbers by $\mathbb{N}$, the integers by $\mathbb{Z}$, the rational numbers by $\mathbb{Q}$ and the real numbers by $\mathbb{R}$. If we only consider non-negative numbers we add a subscript "$\geqslant 0$" and when considering positive numbers we add asubscript "> 0". Furthermore, for real numbers $a, b \in \mathbb{R}$ with $a \leqslant b$ we denote the closed interval by $[a, b]$ and the open interval by (a, b). Sometimes, we also use partially open intervals denoted by either $(a, b]$ or $[a, b)$. Whenever used, we denote vectors by small letters $x, y, z \in \mathbb{R}^n$ and matrices by capital letters $M \in \mathbb{R}^m \times \mathbb{R}^n$, where $n, m \in \mathbb{N}$.

For a function $f : S \to S'$ from a set S to a set S' we define the inverse function $f^{-1} : S' \to 2^S$ by $f^{-1}(s') = \{s \in S \mid f(s) = s'\}$. For a function $f : S \to X$ from a set S to a set of numbers X (e.g. $\mathbb{Z}, \mathbb{Q}$ or $\mathbb{R}$) we define the infimum $\inf_{s \in S'} f(s)$ for some subset $S' \subseteq S$ as the greatest lower bound of the set $f(S') = \{f(s) \mid s \in S'\}$. Analogously, we define the supremum $\sup_{s \in S'} f(s)$ as the lowest upper bound of the set $f(S')$. The minimum $\min_{s \in S'} f(s)$ is the smallest number in the set $f(S')$ if it exists, while the maximum $\max_{s \in S'}$ is the greatest number in the set $f(S')$ if it exists. If the infimum is an element of the set $f(S')$, then $\inf_{s \in S'} f(s) = \min_{s \in S'} f(s)$ and analogous for supremum and maximum.

For real valued functions $f : \mathbb{R} \to \mathbb{R}$ we denote the derivative of f by $\frac{df}{dx}$. We also consider partial derivatives of real valued functions $g : \mathbb{R}^n \to \mathbb{R}$. The partial derivative of g in the i-th coordinate is denoted by $\frac{\partial g}{\partial x_i}$.

We use the following complexity classes of decision problems (see e.g. [Sip13])

- P if it can be decided by a suitable Turing-machine in polynomial time,
- NP if it can be decided by a non-deterministic Turing-machine in polynomial time,
- coNP if the complement of the decision problem is in NP and
- PSPACE if it can be decided by a Turing-machine using a polynomial amount of space.

2.1 Markov Decision Processes

The main model we are concerned with are Markov decision processes, which we will introduce briefly here. For more details, we refer to [Put94; BK08; Kal20].

Definition 2 (MDP). In the context of this work a *Markov decision processes (MDP)* is a tuple $\mathcal{M} = (\mathrm{S}, Act, \mathrm{P}, init)$ where

- S is a finite set of *states*,
- Act is a finite set of *actions*,
- $init \in \mathrm{S}$ is a unique *initial state* and
- $\mathrm{P} : \mathrm{S} \times Act \times \mathrm{S} \to [0,1]$ is the *transition probability function* with $\sum_{t \in \mathrm{S}} \mathrm{P}(s, \alpha, t) \in \{0, 1\}$ for all states $s \in \mathrm{S}$ and actions $\alpha \in Act$. ◁

Note that we choose a setting where the MDP has a unique initial state. In the literature MDPs are alternatively defined as a tuple $\mathcal{M} = (\mathrm{S}, Act, \mathrm{P}, \iota)$ where ι is a distribution over S while all other components are as above. Given such an alternative MDP with initial distribution ι we can add a fresh state $init$ which is then considered as the unique initial state such that $\mathrm{P}(init, \cdot, \cdot) = \iota$. That is, from $init$ we have transitions according to ι.

For a state $s \in \mathrm{S}$ an action $\alpha \in Act$ and a set of states $\mathrm{T} \subseteq \mathrm{S}$ we use the abbreviation $\mathrm{P}(s, \alpha, \mathrm{T}) = \sum_{t \in \mathrm{T}} \mathrm{P}(s, \alpha, t)$. We say an action α is *enabled* in state $s \in \mathrm{S}$ if $\sum_{t \in \mathrm{S}} \mathrm{P}(s, \alpha, t) = 1$. We denote the set of enabled actions in state $s \in \mathrm{S}$ by $Act(s) = \{\alpha \in Act \mid \alpha \text{ is enabled in } s\}$. A *state-action pair* of $\mathcal{M}$ is a tuple (s, α) such that $s \in \mathrm{S}$ and $\alpha \in Act(s)$. The set of state-action pairs of $\mathcal{M}$ is denoted by $StAct = \{(s, \alpha) \in \mathrm{S} \times Act \mid \alpha \in Act(s)\}$. A state $t \in \mathrm{S}$ is *terminal* if $Act(t) = \emptyset$. A *Markov chain* (MC) is an MDP where Act is a singleton (we then write $\mathrm{P}(s, t)$ rather than $\mathrm{P}(s, \alpha, t)$).

2.1.1 Graph of an MDP

Given an MDP $\mathcal{M} = (\mathrm{S}, Act, \mathrm{P}, init)$ the *underlying graph* $\mathrm{G}(\mathcal{M}) = (\mathrm{S}, \mathrm{E})$ of $\mathcal{M}$ has the vertices S and the edges $\mathrm{E} = \{(s, t) \in \mathrm{S} \times \mathrm{S} \mid \text{there is } \alpha \in Act \text{ with } \mathrm{P}(s, \alpha, t) > 0\}$. The *size of an MDP* is the number of edges in the underlying graph. Therefore, the MDPs we are concerned with are of finite size as the state space S is finite.

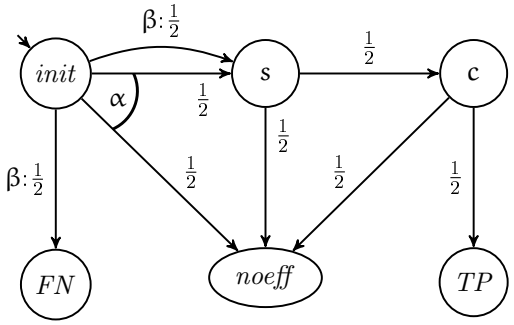

Figure 2.1: The graph of an MDP $\mathcal{M}$

Example 3 (Depicting an MDP). We represent MDPs as a graphs. For an example MDP $\mathcal{M}$ this is shown in Figure 2.1. This MDP has the state space $S = \{init, s, c, FN, noeff, TP\}$ and the set of actions $Act = \{\alpha, \beta\}$. The initial state $init$ is marked by an incoming edge which is not an outgoing edge of another state. For states $s, t \in S$ and action $\alpha \in Act$, if $P(s, \alpha, t) = p > 0$ then there is a directed edge from s to t labeled with p. Given a state-action pair $(s, \alpha) \in StAct$, we have two ways of connecting the corresponding transitions to $P(s, \alpha, \cdot)$. Either all corresponding transitions for (s, α) are connected by an arc, as seen for $(init, \alpha)$ in Figure 2.1, or the corresponding transitions have an additional label denoting the action, as for $(init, \beta)$. If there is only a single enabled action in a state, we omit the action label on outgoing transitions, as for s in Figure 2.1. Terminal states do not have outgoing transitions as they do not have enabled actions, like FN or TP in the example. ◁

A *sub-MDP* of an MDP $\mathcal{M} = (S, Act, P, init))$ is a tuple $\mathcal{E} = (T, \mathcal{A})$, where $T \subseteq S$ and $\mathcal{A} \subseteq Act$ such that for each $\alpha \in \mathcal{A}$ there is $t \in T$ with $\alpha \in Act(t)$ and

$$\{s \in S \mid P(t, \alpha, s) > 0\} \subseteq T \qquad \text{for all } t \in T, \alpha \in \mathcal{A}.$$

An *end-component* (EC) of $\mathcal{M}$ is a sub-MDP $\mathcal{E} = (T, \mathcal{A})$ of $\mathcal{M}$ for which the directed graph with vertices T and edges $\{(s, t) \in T \times T \mid \text{there is } \alpha \in \mathcal{A} \text{ with } P(s, \alpha, t) > 0\}$ is strongly connected An EC $\mathcal{E} = (T, \mathcal{A})$ can be identified with the set $\{(t, \alpha) \mid t \in T, \alpha \in Act(t) \cap \mathcal{A}\} \subseteq StAct$. An EC $\mathcal{E}$ is *maximal* (abbreviated MEC) if for each EC $\mathcal{E}' = (T', \mathcal{A}')$ with $T \subseteq T'$ and $\mathcal{A} \subseteq \mathcal{A}'$ we have $\mathcal{E} = \mathcal{E}'$.

2.1.2 Schedulers and Probability Measure

A *path* π in an MDP $\mathcal{M} = (S, Act, P, init)$ is a (finite or infinite) alternating sequence of states $s_i \in S$ and actions $\alpha_i \in Act$, e.g. $\pi = s_0\, \alpha_0\, s_1\, \alpha_1\, s_2 \cdots \in (S \times Act)^* \times S \cup (S \times Act)^\omega$, such that $P(s_i, \alpha_i, s_{i+1}) > 0$ for all indices i. We say a path is *maximal* if it is either infinite or it is finite and ends in a terminal state. Sometimes, we consider terminal states as states that only have one outgoing transition which is a self loop. In such a case all maximal paths are

infinite. The set of finite non-maximal paths in $\mathcal{M}$ is denoted by $Paths_{fin}(\mathcal{M})$ and the set of maximal paths in $\mathcal{M}$ by $Paths(\mathcal{M})$. For a finite path $\pi \in Paths_{fin}(\mathcal{M})$ we denote the last state of π by $last(\pi)$. That is, if $\pi = s_0\, \alpha_0 \ldots \alpha_{n-1}\, s_n$ then $last(\pi) = s_n$.

A *(randomized) scheduler* $\mathfrak{S}$ of an MDP $\mathcal{M}$ is a function that maps each finite non-maximal path $s_0\alpha_0 \ldots s_n$ to a distribution over $Act(s_n)$. Formally, $\mathfrak{S} : Paths_{fin}(\mathcal{M}) \times Act \to [0,1]$ with

$$\sum_{\alpha \in Act(last(\pi))} \mathfrak{S}(\pi, \alpha) = 1 \quad \text{for all } \pi \in Paths_{fin}(\mathcal{M})$$

A scheduler $\mathfrak{S}$ is called *memoryless* if for all paths π the chosen distribution of $\mathfrak{S}$ only depends on the last state of the path. We write MR for the class of memoryless (randomized) schedulers. Since an MR scheduler $\mathfrak{S}$ only depends on the last state s of a path we can also consider it as a function $\mathfrak{S} : S \times Act \to [0,1]$. Furthermore, a scheduler $\mathfrak{S}$ is called *deterministic* if $\mathfrak{S}(\pi, \cdot)$ is a Dirac distribution for all finite non-maximal paths π. We write MD for the class of memoryless deterministic schedulers. Such MD schedulers $\mathfrak{S}$ can be considered as functions $\mathfrak{S} : S \to Act$. Moreover, *finite-memory* schedulers are those which are representable by a finite-state automaton, which we introduce in the next Section 2.2.

A path π of $\mathcal{M}$ is said to be an $\mathfrak{S}$-path if $\mathfrak{S}(s_0\, \alpha_0 \ldots \alpha_{i-1}\, s_i, \alpha_i) > 0$ for each $i \in \{0, \ldots, n-1\}$. Given a path $\pi = s_0\, \alpha_0\, \ldots\, \alpha_{n-1}\, s_n$, the *residual scheduler* $res(\mathfrak{S}, \pi)$ of $\mathfrak{S}$ after π is defined by $res(\mathfrak{S}, \pi)(\phi, \cdot) = \mathfrak{S}(\pi \circ \phi, \cdot)$ for all finite paths ϕ starting in s_n. Here, $\pi \circ \phi$ denotes the concatenation of the paths π and ϕ. Intuitively speaking, $res(\mathfrak{S}, \pi)$ behaves like $\mathfrak{S}$ after π has already been seen.

A scheduler $\mathfrak{S}$ of $\mathcal{M}$ induces a (possibly infinite) Markov chain $\mathcal{M}_{\mathfrak{S}}$ by resolving the actions of $\mathcal{M}$ (also cf. [BK08, Definition 10.92]). On $\mathcal{M}_{\mathfrak{S}}$ the set of maximal paths $Paths(\mathcal{M}_{\mathfrak{S}})$ carries a probability measure. The associated σ-algebra to $\mathcal{M}_{\mathfrak{S}}$ is generated by the cylinder sets $Cyl(\tau) = \{\pi \in Paths(\mathcal{M}_{\mathfrak{S}}) \mid \tau \text{ is prefix of } \pi\}$, where τ is a finite path in $\mathcal{M}_{\mathfrak{S}}$ (also cf. [BK08, Definition 10.10]). For $\tau = init\ s_1\ s_2 \ldots s_n$ the probability of $Cyl(\tau)$ is given by

$$\Pr^{\mathfrak{S}}(Cyl(\tau)) = P(init, s_1) \cdot P(s_1, s_2) \cdot \ \ldots\ \cdot P(s_{n-1}, s_n).$$

We write $\Pr^{\mathfrak{S}}_{\mathcal{M},s}$ for this probability measure on measurable sets of maximal paths in the MC $\mathcal{M}_{\mathfrak{S}}$ with initial state $s \in S$. We omit the subscript s if the state is also the initial state $init$ of $\mathcal{M}$. Furthermore, whenever the MDP $\mathcal{M}$ for which we consider the probability measure is clear from the context we also omit the subscript $\mathcal{M}$. Then, $\Pr^{\mathfrak{S}}$ denotes the probability measure of $\mathcal{M}$ with initial state $init$.

Example 4 (Probability Measure on MDPs). For the MDP $\mathcal{M}$ from Figure 2.1 the function $\mathfrak{S}$ which maps $(init, \alpha)$ and $(init, \beta)$ to $\frac{1}{2}$ defines a memoryless randomized scheduler as in each other state a scheduler of $\mathcal{M}$ would only have a single choice. The sequence $\pi = init\ s\ c\ noeff$ is a maximal $\mathfrak{S}$-path with probability

$$\Pr^{\mathfrak{S}}(\{\pi\}) = \Pr^{\mathfrak{S}}(Cyl(\pi)) = 2 \cdot (\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}) = \frac{1}{8}.$$

2.1.3 Maximal and Minimal Probabilities

For a measurable set of maximal paths φ we define the *extremal probabilities* $\Pr^{\max}_{\mathcal{M},s}(\varphi)$ and $\Pr^{\min}_{\mathcal{M},s}(\varphi)$ as the supremum and respectively infimum of the probabilities for φ under all schedulers of $\mathcal{M}$, that is $\Pr^{\max}_{\mathcal{M},s}(\varphi) = \sup_{\mathfrak{S}} \Pr^{\mathfrak{S}}_{\mathcal{M},s}(\varphi)$ and analogous for $\Pr^{\min}_{\mathcal{M},s}$. The *expectation* of a random variable f in $\mathcal{M}$ with initial state $s \in S$ under scheduler $\mathfrak{S}$ is the expected value of f under the probability measure of $\mathcal{M}_{\mathfrak{S}}$ with initial state s. This expectation is then denoted by $\mathrm{E}^{\mathfrak{S}}_{\mathcal{M},s}(f)$. The supremum and infimum of the expectation of f over all schedulers $\mathfrak{S}$ are denoted by $\mathrm{E}^{\max}_{\mathcal{M},s}(f)$ and respectively $\mathrm{E}^{\min}_{\mathcal{M},s}(f)$. We omit the subscripts s if $s = \mathit{init}$ and $\mathcal{M}$ if the MDP is clear from context.

We use LTL-like temporal modalities such as $\Diamond$ (eventually) and U (until) to denote path properties. For two sets of states $X, T \subseteq S$ the LTL formula $X \,\mathrm{U}\, T$ is satisfied by those paths $\pi = s_0 s_1 \ldots$ for which there exists $j \geqslant 0$ such that for all $i < j : s_i \in X$ and $s_j \in T$. The *reachability property* of a set $T \subseteq S$ is defined as $\Diamond T = S \,\mathrm{U}\, T$. Given an MDP $\mathcal{M}$ we identify $\Diamond T$ and $X \,\mathrm{U}\, T$ with the set of maximal paths that satisfy the respective LTL formula, by abuse of notation. It is well-known that the extremal probabilities $\Pr^{\min}_{\mathcal{M}}(X \,\mathrm{U}\, T)$ and $\Pr^{\max}_{\mathcal{M}}(X \,\mathrm{U}\, T)$ are computable in polynomial time and corresponding optimal schedulers are memoryless deterministic and can also be derived in polynomial time. [Alf97; BK08]. We say an action $\alpha \in \mathit{Act}$ is *maximizing* for the property $X \,\mathrm{U}\, T$ if it is chosen by an MD scheduler $\mathfrak{S}$ for which $\Pr^{\mathfrak{S}}(X \,\mathrm{U}\, T) = \Pr^{\max}(X \,\mathrm{U}\, T)$. We define *minimizing* actions analogously.

We also use *conditional probabilities* in MDPs cf. [Bai+14; Mär20]. For two measurable sets of maximal paths φ and ψ we have

$$\Pr^{\max}_{\mathcal{M},s}(\varphi \mid \psi) = \max_{\mathfrak{S}} \Pr^{\mathfrak{S}}_{\mathcal{M},s}(\varphi \mid \psi) = \max_{\mathfrak{S}} \frac{\Pr^{\mathfrak{S}}_{\mathcal{M},s}(\varphi \cap \psi)}{\Pr^{\mathfrak{S}}_{\mathcal{M},s}(\psi)}$$

and $\Pr^{\min}_{\mathcal{M},s}(\varphi \mid \psi)$ likewise. If both φ and ψ are reachability properties then maximal and minimal conditional probabilities are computable in polynomial time [Bai+14].

Example 5 (Maximizing Action and Conditional Probability). In the MDP $\mathcal{M}$ from Figure 2.1 the maximal probability for reaching the state TP is $\Pr^{\max}(\Diamond \mathit{TP}) = \frac{1}{8}$, whereas the maximal probability for reaching state FN is $\Pr^{\max}(\Diamond \mathit{FN}) = \frac{1}{2}$. For $\Diamond \mathit{TP}$ a maximizing MD scheduler can choose both α or β in init, so both α and β are maximizing actions for $\Diamond \mathit{TP}$. However, for $\Diamond \mathit{FN}$ only β is a maximizing action. For a scheduler $\mathfrak{S}$ of $\mathcal{M}$ choosing α deterministically in init we have the conditional probability

$$\Pr^{\mathfrak{S}}(\Diamond \mathit{TP} \mid \Diamond c) = \frac{\Pr^{\mathfrak{S}}(\Diamond \mathit{TP} \cap \Diamond c)}{\Pr^{\mathfrak{S}}(\Diamond c)} = \frac{1}{8} / \frac{1}{4} = \frac{1}{2}.$$

2.1.4 Frequencies and Balance Equations

Given a scheduler $\mathfrak{S}$ for $\mathcal{M}$, the expected *frequencies* of state-action pairs (s, α), states $\mathrm{t} \in \mathrm{S}$ and set of states $\mathsf{T} \subseteq \mathrm{S}$ under $\mathfrak{S}$ are defined by:

$$\begin{aligned} \mathit{freq}_{\mathfrak{S}}(\mathrm{s}, \alpha) &= \mathrm{E}^{\mathfrak{S}}_{\mathcal{M}}(\text{number of visits to } \mathrm{s} \text{ in which } \alpha \text{ is taken}) \\ \mathit{freq}_{\mathfrak{S}}(\mathrm{t}) &= \sum_{\alpha \in \mathit{Act}(\mathrm{t})} \mathit{freq}_{\mathfrak{S}}(\mathrm{t}, \alpha), \qquad \mathit{freq}_{\mathfrak{S}}(\mathsf{T}) = \sum_{\mathrm{u} \in \mathsf{T}} \mathit{freq}_{\mathfrak{S}}(\mathrm{u}). \end{aligned}$$

In end-component free MDPs we can specify MR scheduler by their state-action frequencies (see e.g. [Kal20, Theorem 4.7]). For the set StAct of state-action pairs in $\mathcal{M}$ we consider a linear constraint system over the variables $\mathrm{x}_{\mathrm{s},\alpha}$ for each $(\mathrm{s}, \alpha) \in \mathit{StAct}$ where we use the short form notation $\mathrm{x}_{\mathrm{s}} = \sum_{\alpha \in \mathit{Act}(\mathrm{s})} \mathrm{x}_{\mathrm{s},\alpha}$.

$$\mathrm{x}_{\mathrm{s},\alpha} \geqslant 0 \qquad \text{for all } (\mathrm{s}, \alpha) \in \mathit{StAct} \tag{S1}$$

$$\mathrm{x}_{\mathit{init}} = 1 + \sum_{(\mathrm{t},\alpha) \in \mathit{StAct}} \mathrm{x}_{\mathrm{t},\alpha} \cdot \mathsf{P}(\mathrm{t}, \alpha, \mathit{init}) \tag{S2}$$

$$\mathrm{x}_{\mathrm{s}} = \sum_{(\mathrm{t},\alpha) \in \mathit{StAct}} \mathrm{x}_{\mathrm{t},\alpha} \cdot \mathsf{P}(\mathrm{t}, \alpha, \mathrm{s}) \qquad \text{for all } \mathrm{s} \in \mathrm{S} \setminus \{\mathit{init}\} \tag{S3}$$

By [Kal20, Theorem 4.7] there is a bijection between the set of memoryless schedulers of $\mathcal{M}$ (which are called "transient policies" in the terminology of [Kal20]) and feasible soultions to (S1) - (S3). Thus, a vector $\mathrm{x} \in \mathbb{R}^{\mathit{StAct}}$ which constitutes a solution to (S1) and the *balance equations* (S2) and (S3) corresponds one-to-one to an MR scheduler $\mathfrak{S}$ for $\mathcal{M}$ such that $\mathrm{x}_{\mathrm{s},\alpha} = \mathit{freq}_{\mathfrak{S}}(\mathrm{s}, \alpha)$ for all $(\mathrm{s}, \alpha) \in \mathit{StAct}$. Furthermore, this also characterizes possibly history dependent scheduler by [Kal20, Theorem 9.16]. Therefore, a solution x for (S1) - (S3) also corresponds one-to-one to a scheduler $\mathfrak{S}$ of $\mathcal{M}$ with $\mathrm{x}_{\mathrm{s},\alpha} = \mathit{freq}_{\mathfrak{S}}(\mathrm{s}, \alpha)$ for all $(\mathrm{s}, \alpha) \in \mathit{StAct}$.

2.1.5 MR Scheduler in MDPs without ECs

By [Kal20, Theorem 9.16] we can relate MR scheduler with general scheduler in MDPs without end-components.

Lemma 6 (From General Schedulers to MR Schedulers in MDPs without ECs). *Consider an MDP $\mathcal{M} = (\mathrm{S}, \mathit{Act}, \mathsf{P}, \mathit{init})$ without end-components. Then, for each scheduler $\mathfrak{S}$ for $\mathcal{M}$, there exists an MR scheduler $\mathfrak{T}$ such that:*

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathrm{t}) = \Pr^{\mathfrak{T}}_{\mathcal{M}}(\Diamond \mathrm{t}) \quad \textit{for each terminal state } \mathrm{t} \in \mathrm{S}.$$

This allows us to use convex combinations of scheduler in EC-free MDPs by the following Lemma 7.

Lemma 7 (Convex Combination of MR Schedulers). *For an end-component free MDP $\mathcal{M}$ and two schedulers $\mathfrak{S}$ and $\mathfrak{T}$ for $\mathcal{M}$ we have that for a real number $\lambda \in (0,1)$ there exists an MR scheduler $\mathfrak{U}$ of $\mathcal{M}$ such that:*

$$\Pr^{\mathfrak{U}}_{\mathcal{M}}(\Diamond t) \;=\; \lambda \cdot \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond t) + (1-\lambda) \cdot \Pr^{\mathfrak{T}}_{\mathcal{M}}(\Diamond t)$$

for each terminal state $t \in S$.

Proof. By Lemma 6 we suppose that $\mathfrak{S}$ and $\mathfrak{T}$ are MR-schedulers. Let

$$f(t,\alpha) = \lambda \cdot \mathit{freq}_{\mathfrak{S}}(t,\alpha) + (1-\lambda) \cdot \mathit{freq}_{\mathfrak{T}}(t,\alpha), \quad \text{and} \quad f(t) = \lambda \cdot \mathit{freq}_{\mathfrak{S}}(t) + (1-\lambda) \cdot \mathit{freq}_{\mathfrak{T}}(t).$$

We define the MR scheduler $\mathfrak{U}$ of $\mathcal{M}$ by setting $\mathfrak{U}(s,\alpha) = \frac{f(s,\alpha)}{f(s)}$ for each non-terminal state $s \in S$ with $f(s) > 0$ and for each action $\alpha \in \mathit{Act}(s)$. Moreover, if for some non-terminal state $s \in S$ we have $f(s) = 0$ then we let the MR scheduler $\mathfrak{U}$ select an arbitrary distribution over $\mathit{Act}(s)$.

By Lemma 6 we obtain $f(s) = \mathit{freq}_{\mathfrak{U}}(s)$ where s ranges over all states and state-action pairs in $\mathcal{M}$. This yields

$$\Pr^{\mathfrak{U}}_{\mathcal{M}}(\Diamond t) = f(t) = \lambda \cdot \mathit{freq}_{\mathfrak{S}}(t) + (1-\lambda) \cdot \mathit{freq}_{\mathfrak{T}}(t) = \lambda \cdot \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond t) + (1-\lambda) \cdot \Pr^{\mathfrak{T}}_{\mathcal{M}}(\Diamond t)$$

for each terminal state t. □

Let $\mathcal{M}, \mathfrak{S}, \mathfrak{T}, \lambda$ be as in Lemma 7. Then, the notation $\lambda\mathfrak{S} \oplus (1-\lambda)\mathfrak{T}$ will be used to denote an MR scheduler $\mathfrak{U}$ as in Lemma 7.

2.1.6 MEC-Quotient

Throughout the thesis there are multiple uses of the concept of the MEC-quotient, which is standard for the analysis of MDPs [Alf97]. Intuitively, the MEC-quotient of an MDP collapses all maximal end-components, ignoring the actions inside of the end-component while keeping outgoing transitions. More concretely, we use a modified version of the MEC-quotient with an additional trap state as in [Bai+18] which serves to mimic behaviors which stay inside an end-component of the original MDP.

Definition 8 (MEC-Quotient of an MDP). Let $\mathcal{M} = (S, \mathit{Act}, P, \mathit{init})$ be an MDP with end-components. Let $\mathcal{E}_1, \ldots, \mathcal{E}_k$ be the maximal end-components (MECs) of $\mathcal{M}$. We may suppose without loss of generality that the enabled actions of states are pairwise disjoint. That is, whenever $s_1, s_2 \in S$ with $s_1 \neq s_2$ then $\mathit{Act}_{\mathcal{M}}(s_1) \cap \mathit{Act}_{\mathcal{M}}(s_2) = \varnothing$. This way, we can consider any MEC $\mathcal{E}_i$ as a subset of Act. Let U_i denote the set of states that belong to $\mathcal{E}_i$ and let $U = U_1 \cup \ldots \cup U_k$. The *MEC-quotient of* $\mathcal{M}$ is the MDP $\mathcal{N} = \mathit{MEC}(\mathcal{M}) = (S', \mathit{Act}', P', \mathit{init}')$ and the function $\iota : S \to S'$ defined as follows:

- $S' = S \setminus U \cup \{s_{\mathcal{E}_1}, \ldots, s_{\mathcal{E}_k}, \bot\}$ where $s_{\mathcal{E}_1}, \ldots, s_{\mathcal{E}_k}, \bot$ are pairwise distinct fresh states,
- ι is given by $\iota(s) = s$ for $s \in S \setminus U$ and $\iota(u) = s_{\mathcal{E}_i}$ if $u \in U_i$,
- $init' = \iota(init)$,
- $Act' = Act \cup \{\tau\}$ where τ is a fresh action symbol and
- for $s \in S'$ the enabled actions $Act_{\mathcal{N}}(s)$ and transition probabilities are:
  - if $s \in S \cap S'$ then $Act_{\mathcal{N}}(s) = Act_{\mathcal{M}}(s)$ and $P'(s, \alpha, s') = P(s, \alpha, \iota^{-1}(s'))$ for all $s' \in S'$ and $\alpha \in Act_{\mathcal{M}}(s)$,
  - if $s = s_{\mathcal{E}_i}$ is representing the MEC $\mathcal{E}_i$ of $\mathcal{M}$ then (as we view $\mathcal{E}_i$ as a set of actions):

$$Act_{\mathcal{N}}(s_{\mathcal{E}_i}) = \bigcup_{u \in U_i} (Act_{\mathcal{M}}(u) \setminus \mathcal{E}_i) \cup \{\tau\},$$

    where for τ we have $P'(s_{\mathcal{E}_i}, \tau, \bot) = 1$ and for $u \in U_i$ and $\alpha \in Act_{\mathcal{M}}(u) \setminus \mathcal{E}_i$ we set $P'(s_{\mathcal{E}_i}, \alpha, s') = P(u, \alpha, \iota^{-1}(s'))$ for all $s' \in S'$.
  - $Act_{\mathcal{N}}(\bot) = \varnothing$. ◁

By construction, each terminal state of $\mathcal{M}$ is terminal in its MEC-quotient $MEC(\mathcal{M})$. Vice versa, every terminal state of $MEC(\mathcal{M})$ is either a terminal state of $\mathcal{M}$ or $\bot$. Moreover, $MEC(\mathcal{M})$ has no end-components, which implies that for every scheduler $\mathfrak{T}$ of $MEC(\mathcal{M})$, a terminal state will be reached with probability 1.

The $\mathcal{M}$ and $MEC(\mathcal{M})$ are connected by the following Lemma 9 (see also [Alf97; Alf99]).

Lemma 9 (Correspondence of an MDP and its MEC-Quotient). *Let $\mathcal{M}$ be an MDP and consider its MEC-quotient $MEC(\mathcal{M})$. Then, for each scheduler $\mathfrak{S}$ for $\mathcal{M}$ there is a scheduler $\mathfrak{T}$ for $MEC(\mathcal{M})$ such that*

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond t) = \Pr^{\mathfrak{T}}_{MEC(\mathcal{M})}(\Diamond t) \quad \textit{for each terminal state } t \textit{ of } \mathcal{M} \tag{2.1}$$

and vice versa. Moreover, if (2.1) *holds then $\Pr^{\mathfrak{T}}_{MEC(\mathcal{M})}(\Diamond \bot)$ equals the probability for $\mathfrak{S}$ to generate an infinite path in $\mathcal{M}$ that eventually enters and stays forever in an end-component.*

Proof. Given a scheduler $\mathfrak{T}$ for $\mathcal{N} = MEC(\mathcal{M})$, we pick an MD scheduler $\mathfrak{U}$ of $\mathcal{M}$ such that $\mathfrak{U}(u) \in \mathcal{E}_i$ for each $u \in U_i$. Then, the corresponding scheduler $\mathfrak{S}$ for $\mathcal{M}$ behaves as $\mathfrak{T}$ as long as $\mathfrak{T}$ does not choose the τ-transition to $\bot$. As soon as $\mathfrak{T}$ schedules τ then $\mathfrak{S}$ behaves as $\mathfrak{U}$ from this moment on.

Vice versa, given a scheduler $\mathfrak{S}$ for $\mathcal{M}$ then a corresponding scheduler $\mathfrak{T}$ for $\mathcal{N}$ mimics $\mathfrak{S}$ as long as $\mathfrak{S}$ has not visited a state belonging to an end-component $\mathcal{E}_i$ of $\mathcal{M}$. Scheduler $\mathfrak{T}$ ignores $\mathfrak{S}$'s transitions inside an MEC $\mathcal{E}_i$ and takes $\beta \in \bigcup_{u \in U_i}(Act_{\mathcal{M}}(u) \setminus \mathcal{E}_i)$ with the same probability as $\mathfrak{S}$ leaves $\mathcal{E}_i$. With the remaining probability mass, $\mathfrak{S}$ stays forever inside $\mathcal{E}_i$, which is mimicked by $\mathfrak{T}$ by taking the τ-transition to $\bot$.

For the formal definition of $\mathfrak{T}$, we use the following notation. For simplicity, let us assume that $init \notin U_1 \cup \ldots \cup U_k$. This yields $init = init'$. Given a finite path $\pi = s_0\, \alpha_0\, s_1\, \alpha_1 \ldots \alpha_{m-1}\, s_m$ in $\mathcal{M}$ with $s_0 = init$, let $\pi_{\mathcal{N}}$ be the path in $\mathcal{N}$ resulting from by replacing each maximal path fragment $s_h \alpha_h \ldots \alpha_{j-1} s_j$ consisting of actions inside an $\mathcal{E}_i$ with state $s_{\mathcal{E}_i}$. (Here, maximality means if $h > 0$ then $\alpha_{h-1} \notin \mathcal{E}_i$ and if $j < m$ then $\alpha_{j+1} \notin \mathcal{E}_i$.) Furthermore, let $p_\pi^{\mathfrak{S}}$ denote the probability for $\mathfrak{S}$ to generate the path π when starting in the first state of π.

Let ρ be a finite path in $\mathcal{N}$ with first state $init$ (recall that we suppose that $\mathcal{M}$'s initial state does not belong to an MEC, which yields $init = init'$) and $last(\rho) \neq \bot$. Then, Π_ρ denotes the set of finite paths $\pi = s_0\, \alpha_0\, s_1\, \alpha_1 \ldots \alpha_{m-1}\, s_m$ in $\mathcal{M}$ such that

- $\pi_{\mathcal{N}} = \rho$ and
- if $s_m \in U_i$ then $\alpha_{m-1} \notin \mathcal{E}_i$.

The formal definition of scheduler $\mathfrak{T}$ is as follows. Let ρ be a non-maximal finite path in $\mathcal{N}$ with last state s. If s is a state of $\mathcal{M}$ and does not belong to an MEC of $\mathcal{M}$ and $\beta \in Act_{\mathcal{M}}(s)$ then:

$$\mathfrak{T}(\rho, \beta) = \sum_{\pi \in \Pi_\rho} p_\pi^{\mathfrak{S}} \cdot \mathfrak{S}(\pi, \beta)$$

If $s = s_{\mathcal{E}_i}$ and $\beta \in Act_{\mathcal{N}}(s_{\mathcal{E}_i}) \setminus \{\tau\}$ then

$$\mathfrak{T}(\rho, \beta) = \sum_{\pi \in \Pi_\rho} p_\pi^{\mathfrak{S}} \cdot \Pr^{res(\mathfrak{S},\pi)}_{\mathcal{M}, last(\pi)}\big(\text{“leave } \mathcal{E}_i \text{ via action } \beta\text{”}\big)$$

where “leave $\mathcal{E}_i$ via action β” means the existence of a prefix whose action sequence consists of actions inside $\mathcal{E}_i$ followed by action β. The last state of this prefix, however, could be a state of U_i. Note $\beta \in Act_{\mathcal{N}}(s_{\mathcal{E}_i})$ means that β could have reached a state outside U_i, but there might be states inside U_i that are accessible via β. Similarly,

$$\mathfrak{T}(\rho, \tau) = \sum_{\pi \in \Pi_\rho} p_\pi^{\mathfrak{S}} \cdot \Pr^{res(\mathfrak{S},\pi)}_{\mathcal{M}, last(\pi)}\big(\text{“stay forever in } \mathcal{E}_i\text{”}\big)$$

where “stay forever in $\mathcal{E}_i$” means that only actions inside $\mathcal{E}_i$ are performed. By induction on the length of ρ we obtain

$$p_\rho^{\mathfrak{T}} = \sum_{\pi \in \Pi_\rho} p_\pi^{\mathfrak{S}},$$

which yields $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond t) = \Pr^{\mathfrak{T}}_{\mathcal{N}}(\Diamond t)$ for each terminal state t of $\mathcal{M}$. Moreover the probability under $\mathfrak{S}$ to eventually enter and stay forever in $\mathcal{E}_i$ equals the probability for $\mathfrak{T}$ to reach the terminal state $\bot$ via a path of the form $\rho\, \tau\, \bot$ where $last(\rho) = s_{\mathcal{E}_i}$. □

2.2 Automata and ω-Regular Languages

An *alphabet* is a (nonempty) finite set Σ. A *word* over Σ is a finite or infinite sequence $a_0 a_1 \dots$ of symbols $a_i \in \Sigma$. The empty word will be denoted by ε. A *prefix* of an infinite word $\sigma = a_0 a_1 \dots$ is a finite word $w = a_0 a_1 \dots a_n$. A prefix of a finite word $\sigma = a_0 a_1 \dots a_m$ with $len(\sigma) = n$ is a finite word $w = a_0 a_1 \dots a_i$ of length at most n. By Σ^* and Σ^ω we respectively denote the set of all finite and infinite words over Σ. A (formal) *language* over Σ is a set of finite words $L \subseteq \Sigma^\star$, while an ω-language is a set of infinite words $\mathcal{L} \subseteq \Sigma^\omega$.

We will define ω-*regular languages* as those languages accepted by a *deterministic Rabin automaton* (DRA). However, note that there are many equivalent ways to define ω-regular languages [Tho90].

Definition 10 (DRA and ω-Regular Languages). A DRA is a tuple $\mathcal{A} = (Q, \Sigma, q_0, \delta, \mathrm{Acc})$ where

- Q is a finite set of states,
- Σ is an alphabet,
- q_0 is the initial state,
- $\delta : Q \times \Sigma \to Q$ is the (partial) transition function and
- $\mathrm{Acc} \subseteq 2^Q \times 2^Q$ is the acceptance set.

Its size $|\mathcal{A}|$ is the number of states and transitions in $\mathcal{A}$. The *run* of a DRA $\mathcal{A}$ on a word $\sigma = \sigma_0 \sigma_1 \dots \in \Sigma^\omega$ is the sequence $\tau = q_0 q_1 \dots$ of states such that $\delta(q_i, \sigma_i) = q_{i+1}$ for all i. A run is *accepting* if there exists a pair $(L, K) \in \mathrm{Acc}$ such that L is only visited finitely often and K is visited infinitely often by τ. The language $\mathcal{L}(\mathcal{A})$ is the set of all words $\sigma \in \Sigma^\omega$ on which the run of $\mathcal{A}$ is accepting. We say an ω-language $\mathcal{L}$ is ω-*regular* if there is a DRA $\mathcal{A}$ such that $\mathcal{L} = \mathcal{L}(\mathcal{A})$.

Similar to the introduction of ω-regular languages, we will introduce *regular languages* as exactly those which are accepted by a *deterministic finite automaton* (DFA). Again, note that there are many equivalent ways to define regular languages [Sip13].

Definition 11 (DFA and Regular Languages). A *deterministic finite automaton (DFA)* is a tuple $\mathcal{A} = (Q, \Sigma, q_0, \delta, F)$ where

- Q is a finite set of states,
- Σ is an alphabet,
- q_0 is an initial state,
- $\delta : Q \times \Sigma \to Q$ is the (partial) transition function and
- $F \subseteq Q$ is the acceptance set.

Its size $|\mathcal{A}|$ is the number of states and transitions in $\mathcal{A}$. A *run* of the DFA $\mathcal{A}$ on a finite word $w = a_0 a_1 \dots a_n$ is a sequence $\tau = q_0 q_1 \dots q_n$ of states such that $\delta(q_i, a_i) = q_{i+1}$ for all i. It is *accepting* if $q_n \in F$. A word w is accepted by $\mathcal{A}$ if its corresponding run on $\mathcal{A}$ is accepting. The language $L(\mathcal{A})$ is the set of all words $w \in \Sigma^*$ accepted by $\mathcal{A}$. A language L is regular if there is a DFA $\mathcal{A}$ such that $L(\mathcal{A}) = L$.

A *good prefix* w for an ω-regular language $\mathcal{L}$ is a finite word such that all infinite extensions of w belong to $\mathcal{L}$. An ω-regular language $\mathcal{L}$ is called a *co-safety language* if all words in $\mathcal{L}$ have a good prefix. A co-safety language $\mathcal{L}$ is uniquely determined by the regular language of minimal good prefixes of words in $\mathcal{L}$. Therefore, a co-safety language can be represented by a DFA accepting exactly all minimal good prefixes of $\mathcal{L}$.

Given an MDP $\mathcal{M} = (S, Act, P, init)$ and a DFA $\mathcal{A} = (Q, \Sigma, q_0, \delta, Acc)$ with $\Sigma \subseteq S$ we define the *product MDP* $\mathcal{M} \otimes \mathcal{A} = (S \times Q, Act, P', init')$ by defining

- $P'(< s, q >, \alpha, < t, r >) = P(s, \alpha, t)$ if $r = \delta(q, s)$ and 0 otherwise,
- and $init' = \delta(q_0, init)$.

Intuitively, the product MDP $\mathcal{M} \otimes \mathcal{A}$ is transitioning according to the transition function given by $\mathcal{M}$ while keeping track of the corresponding run of $\mathcal{A}$ according to the transitions in $\mathcal{M}$. Therefore, the product MDP $\mathcal{M} \otimes \mathcal{A}$ is equipped with the acceptance condition of $\mathcal{A}$ in the second component. The same construction works for the product of an MDP with a deterministic Rabin automaton. The difference comes from the acceptance condition encoded in the second components of states of the product MDP.

2.3 Probability-Raising Causality

In this section we will introduce and discuss the concept of probability-raising causality for random variables. For an in depth discussion of probabilistic causality we refer to [Pea09] and for a more compact summary to [Hit21].

Definition 12 (Probabilistic Cause). Let P be a probability measure on a σ-algebra of events. We say an event Eff is *caused probabilistically* by an event Cause whenever Cause occurs before Eff and for their probabilities we have

$$P(\mathit{Eff} \mid \mathit{Cause}) > P(\mathit{Eff}). \tag{PR}$$

For an example consider a person which often consumes fast food with a high amount of fat (described by an event $\mathit{FastFood}$). Studies have shown that this behavior makes it more likely to get a heart disease [Yoo+21] (described by event $\mathit{HeartDisease}$). However, not every person that regularly consumes fast food will get a heart disease. Much rather, the

probability of getting a heart disease is raised by consuming fast food. Therefore, one has $\mathsf{P}(\mathit{HeartDisease} \mid \mathit{FastFood}) > \mathsf{P}(\mathit{HeartDisease})$.

Note that for two random variables C and E and probability measure P the following holds

$$\mathsf{P}(\mathsf{E} \mid \mathsf{C}) > \mathsf{P}(\mathsf{E}) \quad \text{if and only if} \quad \mathsf{P}(\mathsf{C} \mid \mathsf{E}) > \mathsf{P}(\mathsf{C}) \quad \text{if and only if} \quad \mathsf{P}(\mathsf{E} \mid \mathsf{C}) > \mathsf{P}(\mathsf{E} \mid \neg\mathsf{C}),$$

since we can use the definition of conditional probabilities to get

$$\begin{aligned} \mathsf{P}(\mathsf{E} \mid \mathsf{C}) > \mathsf{P}(\mathsf{E}) &\iff \frac{\mathsf{P}(\mathsf{E} \wedge \mathsf{C})}{\mathsf{P}(\mathsf{C})} \cdot \mathsf{P}(\mathsf{C}) > \mathsf{P}(\mathsf{E}) \cdot \mathsf{P}(\mathsf{C}) \\ &\iff \frac{\mathsf{P}(\mathsf{E} \wedge \mathsf{C})}{\mathsf{P}(\mathsf{E})} > \mathsf{P}(\mathsf{C}) \\ &\iff \mathsf{P}(\mathsf{C} \mid \mathsf{E}) > \mathsf{P}(\mathsf{C}). \end{aligned}$$

We can further use $(1 - \mathsf{P}(\mathsf{C})) = \mathsf{P}(\neg\mathsf{C})$ for

$$\begin{aligned} \mathsf{P}(\mathsf{E} \mid \mathsf{C}) \;>\; \mathsf{P}(\mathsf{E}) &\iff \mathsf{P}(\mathsf{E} \mid \mathsf{C}) \;>\; \mathsf{P}(\mathsf{C}) \cdot \mathsf{P}(\mathsf{E} \mid \mathsf{C}) + \mathsf{P}(\neg\mathsf{C}) \cdot \mathsf{P}(\mathsf{E} \mid \neg\mathsf{C}) \\ &\iff (1 - \mathsf{P}(\mathsf{C})) \cdot \mathsf{P}(\mathsf{E} \mid \mathsf{C}) \;>\; \mathsf{P}(\neg\mathsf{C}) \cdot \mathsf{P}(\mathsf{E} \mid \neg\mathsf{C}) \\ &\iff \mathsf{P}(\mathsf{E} \mid \mathsf{C}) \;>\; \mathsf{P}(\mathsf{E} \mid \neg\mathsf{C}). \end{aligned}$$

Therefore, condition (PR) is equivalent to both

$$\mathsf{P}(\mathit{Cause} \mid \mathit{Eff}) > \mathsf{P}(\mathit{Cause}) \text{ and } \mathsf{P}(\mathit{Eff} \mid \mathit{Cause}) > \mathsf{P}(\mathit{Eff} \mid \neg\mathit{Cause}).$$

Also note that Definition 12 specifically requires Cause to happen *before* the effect Eff. This is what we call *temporal priority* of causes, which is important since without this requirement causes and effects would be symmetrical by above observations. So, without temporal priority every probabilistic cause for an effect is also probabilistically caused by the effect, which is counterintuitive. In the prior example this would result in determining a heart disease as the cause for the consumption of fast food. Obviously, for the actual causal relation consuming fast food occurs before getting a heart disease.

However, temporal priority requires the ability to reason about the temporal order of events. For linear temporal logic (LTL) temporal priority can easily be expressed as the formula $(\neg\mathit{Eff}) \,\mathrm{U}\, \mathit{Cause}$.

3 Probability-Raising Causality in Markov Decision Processes

In this chapter we start to define our two state-based notions of probability-raising (PR) causes and discuss the meaning of the resulting cause-effect relations. For this, we are incorporating the two main principles of probabilistic causes into our definition. Namely,

for a cause C *for the effect* E *we have* $\mathrm{P}(\mathrm{E} \mid \mathrm{C}) > \mathrm{P}(\mathrm{E})$ *and* (C1)

causes must happen before *their effects.* (C2)

Thus, (C1) describes the probability-raising principle and (C2) is the temporal priority of causes.

3.1 Setting and Definition

Before we state our definitions of probability-raising causes, we clarify the setting in which we are working. Throughout the thesis we suppose we are given an MDP $\mathcal{M} = (\mathrm{S}, Act, \mathrm{P}, init)$ such that all states are reachable, formally for all $\mathbf{s} \in \mathrm{S}$ we have $\Pr^{\max}_{\mathcal{M}}(\Diamond \mathbf{s}) > 0$. In this chapter we additionally consider a non-empty set of effect states $\mathit{Eff} \subset \mathrm{S}$ and assume all states $\boldsymbol{e} \in \mathit{Eff}$ to be terminal. These states represent an event for which consider a cause-effect analysis. It shall be noted here, that all effect states are chosen to be terminal to ensure temporal priority (C2) of our causes. We also suppose that $\Pr^{\min}_{\mathcal{M}}(\Diamond \mathit{Eff}) < 1$ as otherwise the effect will almost surely occur and thus we can not determine a probability-raising cause for the effect. In particular, for the unique initial state $init$ this implies that $init \notin \mathit{Eff}$.

Remark 13 (Handling more general Settings). The specified setting above may not be satisfied by more general MDPs. Thus, we briefly sketch how to simplify settings that differ from ours.

Suppose we are given a labeled MDP $\mathcal{M} = (\mathrm{S}, Act, \mathrm{P}, init, AP, \mathrm{L})$ and consider our probabilistic causality framework for an effect event described by some label $\mathfrak{a} \in \mathrm{L}$. We then consider the set of all states labeled $\mathfrak{a}$ and delete all outgoing transitions from these states, making them terminal. Formally, for the set $\mathit{Eff} = \{\mathfrak{s} \in \mathrm{S} \mid \mathfrak{a} \in \mathrm{L}(\mathfrak{s})\}$ we set the enabled actions to $Act(\mathfrak{s}) = \emptyset$ for all $\mathfrak{s} \in \mathit{Eff}$. Thus, we arrive in our state-based setting with a linear overhead in time-complexity for the deletion of transitions. The same can be done in the simpler case, for an (unlabeled) MDP where the states of Eff are not terminal.

Similarly, we can handle a setting in which the effect is a regular property on the atomic propositions specified by a deterministic-finite automaton (DFA) $\mathcal{A} = (\mathrm{Q}, 2^{AP}, \mathfrak{q}_0, \delta, \mathrm{F})$. So, the considered effect is the regular language $\mathcal{L}(\mathcal{A})$. For this, we consider the product of the MDP and the DFA $\mathcal{M} \otimes \mathcal{A} = (\mathrm{S} \times \mathrm{Q}, Act, \mathrm{P}', (init, \mathfrak{q}_0), AP, \mathrm{L}')$ (cf. Sec. 2.2). In this product MDP we consider all states $(\mathfrak{s}, \mathfrak{q})$ in which the second component is final in $\mathcal{A}$ and delete all outgoing edges making them terminal. Formally for $\mathit{Eff} = \{(\mathfrak{s}, \mathfrak{q}) \in \mathrm{S} \times \mathrm{Q} \mid \mathfrak{q} \in \mathrm{F}\}$ we set $Act(\mathfrak{s}, \mathfrak{q}) = \emptyset$ for all $(\mathfrak{s}, \mathfrak{q}) \in \mathit{Eff}$. Therefore, we establish the state-based setting with an polynomial overhead in both space complexity for the product MDP and time-complexity for the deletion of transitions. Obviously, this reduction of settings also works if the effect DFA works over the alphabet S.

Both of these transformations do not influence the idea of an event and thus faithfully describe the occurrence of the effect in the original MDP. ◁

We will now define the main notions of cause-effect relation using the probability-raising principle (C1), which we discussed in Section 2.3.

Definition 14 (Probability-Raising Causes). For an MDP $\mathcal{M} = (\mathrm{S}, Act, \mathrm{P}, init)$ and a terminal set of effect states $\mathit{Eff} \subset \mathrm{S}$ satisfying the above established assumptions we say a set of non-empty states $\mathit{Cause} \subseteq \mathrm{S} \setminus \mathit{Eff}$ satisfying $\Pr^{\max}_{\mathcal{M}}((\neg \mathit{Cause}) \,\mathrm{U}\, \mathfrak{c}) > 0$ for each $\mathfrak{c} \in \mathit{Cause}$ is a *global probability-raising cause* (short *GPR cause*) for Eff in $\mathcal{M}$ if the following condition **(G)** is satisfied.

(G) For each scheduler $\mathfrak{S}$ of $\mathcal{M}$ with $\Pr^{\mathfrak{S}}(\Diamond \mathit{Cause}) > 0$ we have

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff} \mid \Diamond \mathit{Cause}) > \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff}). \tag{GPR}$$

We say Cause is a *strict probability-raising cause* (short *SPR cause*) for Eff in $\mathcal{M}$ if the following condition **(S)** is satisfied.

(S) For each state $\mathfrak{c} \in \mathit{Cause}$ and each scheduler $\mathfrak{S}$ of $\mathcal{M}$ with $\Pr^{\mathfrak{S}}_{\mathcal{M}}((\neg \mathit{Cause}) \,\mathrm{U}\, \mathfrak{c}) > 0$ we have

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff} \mid (\neg \mathit{Cause}) \,\mathrm{U}\, \mathfrak{c}) > \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff}). \tag{SPR}$$

◁

We have discussed in Example 1 of the introduction that probability-raising causality is a form of type causality which takes a forward-looking perspective as opposed to actual causality which often takes a backward-looking perspective (also confer [Hit21]).

For SPR causes the intuition is that whenever the cause occurs the probability of the effect has been raised, captured by a condition for each individual state of Cause. This is a strict view of causality, since for SPR causes there can be no case in which the observation of the Cause does not result in a higher probability for observing Eff. However, in many cases causality is seen in a more global way, where in most cases the observation of a Cause is raising the probability of Eff. This is captured by condition (G) for GPR causes, where the conditional probability considers Cause as a whole. In particular, under all distributions in which a scheduler $\mathfrak{S}$ of $\mathcal{M}$ can enter Cause the observation of the cause raises the probability of reaching Eff. However, this does not restrict the conditional probability for individual elements of the Cause set. Since these two definitions of probability-raising (PR) cause vary in mathematical, algorithmic and philosophical meaning, for the remainder of this section we will discuss both notions and see some of their differences.

First, we see that the effect states being terminal now implies *temporal priority* (C2) of PR causes. For any path in $\mathcal{M}$ which visits both Cause and Eff, it is clear that the cause state has to be visited *before* the effect state is visited, as Cause does not contain effect states. However, with our notion we are not concerned with the number of transitions between Cause and Eff. For example, it might happen that a state $\mathbf{c} \in \mathit{Cause}$ is visited right after the initial state but Eff will only be visited after a very long time. This can be observed in the DTMC in Figure 3.1 for $\mathit{Cause} = \{\mathbf{c}\}$ and $\mathit{Eff} = \{\mathit{eff}\}$. Here, the dotted node can be an arbitrarily large DTMC without terminal states or BSCCs and with an outgoing transition to the effect state eff.

Second, the property that any PR cause satisfies $\Pr^{\max}_{\mathcal{M}}((\neg \mathit{Cause})\,\mathrm{U}\,\mathbf{c}) > 0$ can be seen as a minimality condition, as each state in $\mathbf{c} \in \mathit{Cause}$ can be reached without traversing Cause beforehand. This restricts the set Cause such that there is no state $\overline{\mathbf{c}}$ in Cause which has no impact on the probability-raising property for all scheduler, since any state $\overline{\mathbf{c}}$ for which $\Pr^{\max}_{\mathcal{M}}((\neg \mathit{Cause})\,\mathrm{U}\,\overline{\mathbf{c}}) = 0$ can only be visited after $\Diamond \mathit{Cause}$ has already been observed. We include the condition $\Pr^{\max}_{\mathcal{M}}((\neg \mathit{Cause})\,\mathrm{U}\,\mathbf{c}) > 0$ for all $\mathbf{c} \in \mathit{Cause}$ since the goal of our causality analysis is to give better explications of MDPs and as such, getting rid of redundant information is very useful. However, dropping the minimality condition does not affect most results of later chapters. This is because for fixed scheduler $\mathfrak{S}$ of $\mathcal{M}$ it may still happen that for some state $\mathbf{c} \in \mathit{Cause}$ we have $\Pr^{\mathfrak{S}}((\neg \mathit{Cause})\,\mathrm{U}\,\mathbf{c}) = 0$. So, for the scheduler $\mathfrak{S}$ the state $\mathbf{c}$ has no impact on $\Pr^{\mathfrak{S}}(\Diamond \mathit{Cause})$. However, by the minimality condition there must be some other scheduler $\mathfrak{U}$ for which $\mathbf{c}$ has an impact on this probability. A relation like this can be observed in the MDP $\mathcal{M}$ of Figure 3.2 with $\mathit{Cause} = \{\mathbf{c}_0, \mathbf{c}_1\}$ and $\mathit{Eff} = \{\mathit{eff}\}$. Here, the MD scheduler choosing α in $\mathbf{s}$ can not reach $\mathbf{c}_1$. However, $\mathbf{c}_1$ satisfies $\Pr^{\max}_{\mathcal{M}}((\neg \mathit{Cause})\,\mathrm{U}\,\mathbf{c}_1) > 0$ as e.g. the scheduler $\mathfrak{U}$ choosing β in $\mathbf{s}$ satisfies $\Pr^{\mathfrak{U}}((\neg \mathit{Cause})\,\mathrm{U}\,\mathbf{c}_1) = \frac{1}{6}$.

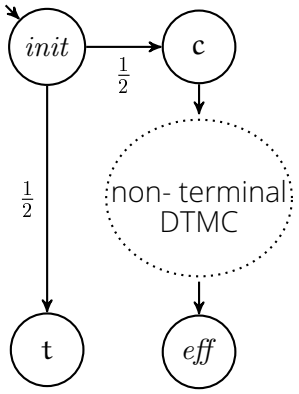

Figure 3.1: A DTMC where the cause is visited early and the effect "late"

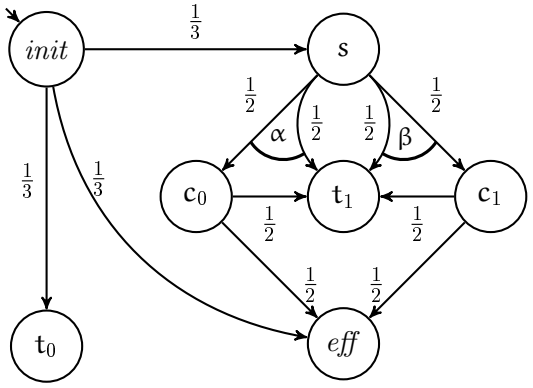

Figure 3.2: An MDP in which MD scheduler can only reach one cause state

With the definition of probability-raising causes we soundly capture the main idea of probabilistic causality. For SPR causes this is especially clear since whenever a system modeled by $\mathcal{M}$ transitions to a state $\mathbf{c} \in \mathit{Cause}$ for the first time, the probability of entering an effect state at some later point will be higher than at the starting point of the system. The probability-raising principle holds *strictly* for each individual state in the cause. Instead of a strict view that for each state the probability of the effect is higher, GPR causes impose that *globally* after visiting the set of states Cause the probability of the effect has been raised in comparison to the start of the system. In particular, GPR causes take the distribution with which Cause is reached by the system into account. Therefore, GPR causes take a global view on the system and only observe whether we have entered the set Cause without observing a specific state. Thus, GPR causes are more general as a cause-effect notion. This claim is supported by the next Proposition 15.

Proposition 15 (Strict PR implies Global PR). *An SPR cause Cause for Eff in $\mathcal{M}$ is also a GPR cause for Eff in $\mathcal{M}$.*

Proof. Assume that Cause is an SPR cause for Eff in $\mathcal{M}$ and let $\mathfrak{S}$ be a scheduler of $\mathcal{M}$ such that $\Pr^{\mathfrak{S}}(\Diamond \mathit{Cause}) > 0$. Let $\mathit{front}^{\mathfrak{S}}(\mathit{Cause}) = \{\mathbf{c} \in \mathit{Cause} \mid \Pr^{\mathfrak{S}}((\neg \mathit{Cause}) \,\mathrm{U}\, \mathbf{c}) > 0\}$ and let

$$\mathfrak{m} = \min_{\mathbf{c} \in \mathit{front}^{\mathfrak{S}}(\mathit{Cause})} \Pr^{\mathfrak{S}}(\Diamond \mathit{Eff} \mid (\neg \mathit{Cause}) \,\mathrm{U}\, \mathbf{c}).$$

As Cause is a SPR cause, we have $\mathfrak{m} > \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff})$. The set of $\mathfrak{S}$-paths satisfying $\Diamond \mathit{Cause}$ is the disjoint union of the sets of $\mathfrak{S}$-paths satisfying $((\neg \mathit{Cause}) \,\mathrm{U}\, \mathbf{c})$ for each $\mathbf{c} \in \mathit{front}^{\mathfrak{S}}(\mathit{Cause})$. Hence,

$$\Pr^{\mathfrak{S}}(\Diamond \mathit{Eff} \mid \Diamond \mathit{Cause}) = \frac{\sum_{\mathbf{c} \in \mathit{front}^{\mathfrak{S}}(\mathit{Cause})} \Pr^{\mathfrak{S}}(\Diamond \mathit{Eff} \mid (\neg \mathit{Cause}) \,\mathrm{U}\, \mathbf{c}) \cdot \Pr^{\mathfrak{S}}((\neg \mathit{Cause}) \,\mathrm{U}\, \mathbf{c})}{\sum_{\mathbf{c} \in \mathit{front}^{\mathfrak{S}}(\mathit{Cause})} \Pr^{\mathfrak{S}}((\neg \mathit{Cause}) \,\mathrm{U}\, \mathbf{c})} \geqslant \mathfrak{m}.$$

Thus, we have $\Pr^{\mathfrak{S}}(\Diamond \mathit{Eff} \mid \Diamond \mathit{Cause}) \geqslant \mathfrak{m} > \Pr^{\mathfrak{S}}(\Diamond \mathit{Eff})$ and (GPR) is satisfied under $\mathfrak{S}$. □

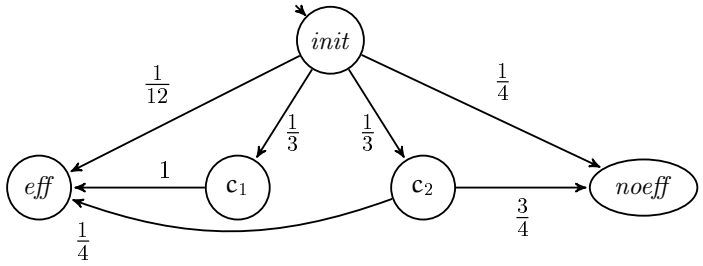

Figure 3.3: A MC with a non-strict GPR cause for eff

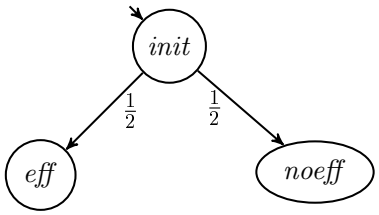

Figure 3.4: A MC with no PR cause for eff

Another connection between GPR and SPR causality is that for a singleton $\mathit{Cause} = \{\mathbf{c}\}$ the formula $((\neg \mathit{Cause})\,\mathrm{U}\,\mathbf{c})$ is equivalent to $\Diamond\mathbf{c}$ and therefore the conditions **(G)** and **(S)** are equivalent. In general the formula $((\neg \mathit{Cause})\,\mathrm{U}\,\mathbf{c})$ describes the event when $\mathbf{c}$ is (first) visited and no other state $\mathbf{c}' \in \mathit{Cause}$ was visited before. Thus, both the SPR condition **(S)** and the GPR condition **(G)** only consider the event when Cause is first visited. However, (SPR) considers only the first visited state in Cause while (GPR) considers the distribution with which Cause is reached.

Returning to the discussion of GPR and SPR causality, together with Proposition 15 we can show that **(G)** is a proper, more general condition than **(S)** . For this, we provide an Example 16 of a GPR cause which is not an SPR cause.

Example 16 (Non-Strict GPR Cause). Consider the Markov chain $\mathfrak{M}$ depicted in Figure 3.3 and let $\mathit{Eff} = \{\mathit{eff}\}$. Then the following holds

$$\begin{aligned}
\Pr(\Diamond \mathit{Eff}) &= \frac{1}{12} + \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{4} = \frac{1}{2},\\
\Pr(\Diamond \mathit{Eff} \mid \Diamond \mathbf{c}_1) &= \Pr_{\mathbf{c}_1}(\Diamond \mathit{eff}) = 1,\\
\Pr(\Diamond \mathit{Eff} \mid \Diamond \mathbf{c}_2) &= \Pr_{\mathbf{c}_2}(\Diamond \mathit{eff}) = \frac{1}{4}.
\end{aligned}$$

Thus, $\{\mathbf{c}_1\}$ is both an SPR and a GPR cause for Eff, while $\{\mathbf{c}_2\}$ is not. The set $\mathit{Cause} = \{\mathbf{c}_1, \mathbf{c}_2\}$, however, is a GPR cause for Eff but not an SPR cause since:

$$\begin{aligned}
\Pr(\Diamond \mathit{Eff} \mid \Diamond \mathit{Cause}) &= \frac{\Pr(\Diamond \mathit{Eff} \wedge \Diamond \mathit{Cause})}{\Pr(\Diamond \mathit{Cause})}\\
&= (\frac{1}{3} + \frac{1}{3} \cdot \frac{1}{4}) / (\frac{1}{3} + \frac{1}{3}) = \frac{5}{8}\\
&> \Pr(\Diamond \mathit{Eff}).
\end{aligned}$$

Cause is not an SPR cause as (SPR) does not hold for $\mathbf{c}_2$. ◁

Since we use a strict inequality in both (GPR) and (SPR) it can happen for a given effect that no probability-raising causes exist.

Remark 17 (PR Causes might not exist). Even in the special case of Markov chains probability-raising causes must not exist. Consider the simple Markov chain in Figure 3.4 with $\mathit{Eff} = \{\mathit{eff}\}$. Since noeff is terminal, the only cause candidate is the singleton $\{\mathit{init}\}$. However, $\{\mathit{init}\}$ cannot be a PR cause, since we require strict inequality in both (GPR) and (SPR). The same phenomenon occurs if all non-terminal states of an MDP reach the effect states with the same probability under all schedulers. In such cases, however, the non-existence of PR causes is well justified as the events $\Diamond \mathit{Eff}$ and $\Diamond \mathit{Cause}$ are stochastically independent for every set $\mathit{Cause} \subseteq S \setminus \mathit{Eff}$. ◁

Remark 17 makes the existence of PR causes a non-trivial property of a given MDP $\mathcal{M}$ and effect set Eff. Thus, a more detailed analysis for necessary and sufficient conditions on the MDP $\mathcal{M}$ and a set Cause to be a probability-raising cause for Eff is needed. Here, we list some necessary conditions for PR causality. However, we postpone the analysis of sufficient conditions and complete characterizations to Chapter 4 as a byproduct of algorithmic considerations for PR causality.

Remark 18 (Necessary Conditions for PR Causality). For a set $\mathit{Cause} \subseteq S \setminus \mathit{Eff}$ to be a GPR (resp. SPR) cause for Eff in $\mathcal{M}$ it needs to hold for each scheduler $\mathfrak{S}$ with $\Pr^{\mathfrak{S}}(\Diamond \mathit{Cause}) > 0$ that

1. $\Pr^{\mathfrak{S}}(\Diamond \mathit{Cause}) < 1$,
2. $\Pr^{\mathfrak{S}}(\Diamond \mathit{Eff}) < 1$ and
3. $\Pr^{\mathfrak{S}}(\Diamond \mathit{Eff}) > 0$.

If the first condition is not met by scheduler $\mathfrak{U}$ of $\mathcal{M}$ we have $\Pr^{\mathfrak{U}}(\Diamond \mathit{Cause}) = 1$ and thus $\Pr^{\mathfrak{U}}(\Diamond \mathit{Eff} \mid \Diamond \mathit{Cause}) = \Pr^{\mathfrak{U}}(\Diamond \mathit{Eff})$. On the other hand, if the second condition does not hold there is a scheduler $\mathfrak{T}$ for which $\Pr^{\mathfrak{T}}(\Diamond \mathit{Eff} \mid \Diamond \mathit{Cause}) \leqslant \Pr^{\mathfrak{T}}(\Diamond \mathit{Eff}) = 1$. Lastly, if the third condition violated by some scheduler $\mathfrak{V}$ of $\mathcal{M}$, we have $\Pr^{\mathfrak{V}}(\Diamond \mathit{Eff} \mid \Diamond \mathit{Cause}) = 0 = \Pr^{\mathfrak{V}}(\Diamond \mathit{Eff})$. In all of these cases the specified scheduler violates the (GPR) condition and thus Cause is not a GPR cause. By Proposition 15 this also means that Cause is not an SPR cause. ◁

3.1.1 Related Work revisited

Now that we have the precise definition of a probability-raising cause given by Definition 14 and collected some properties about this notion, we are able to consider some more specific comparisons to related notions of probabilistic causality in the realm of Markov chains and MDPs. Note that a more general overview over related work can be found in the introduction in Section 1.2.

Prima Facie Causes in DTMCs. Previous work in the direction of probabilistic causation in stochastic operational models has mainly concentrated on Markov chains. One of the earlier approaches to formalize causality based probability-raising comes from Kleinberg et al. (see e.g. [KM09; Kle10]) who introduce the notion of *prima facie causes* in finite Markov chains. There, both causes and effects are formalized as PCTL state formulae, and thus they can equivalently be interpreted as sets of states which satisfy the corresponding state formula.

We will now restate the conditions for prima facie causes from [KM09] in our setting to get the following:

Definition 19 (Prima Facie Cause, cf. [KM09]). Let $\mathcal{M} = (\mathrm{S}, \mathit{Act}, \mathrm{P}, \mathit{init})$ be a DTMC (so Act is a singleton) and let $\mathit{Eff} \subset \mathrm{S}$. Then a set $\mathit{Cause} \subseteq \mathrm{S}$ is a *prima facie cause* if there is $\mathrm{p} > 0$ such that

(1) Cause is reachable from init and $\mathit{Eff} \cap \mathit{Cause} = \emptyset$,

(2) for all $\mathbf{c} \in \mathit{Cause}$ we have $\Pr_{\mathbf{c}}(\Diamond \mathit{Eff}) \geqslant \mathrm{p}$ and

(3) $\Pr_{\mathit{init}}(\Diamond \mathit{Eff}) < \mathrm{p}$.

Note that in [KM09] the value p is implicitly existentially quantified. Since p is existentially quantified the conditions (2) and (3) combine to

$$\text{for all } \mathbf{c} \in \mathit{Cause}: \ \Pr_{\mathbf{c}}(\Diamond \mathit{Eff}) = \Pr_{\mathit{init}}(\Diamond \mathit{Eff} \mid (\neg \mathit{Cause}) \,\mathrm{U}\, \mathbf{c}) > \Pr_{\mathit{init}}(\Diamond \mathit{Cause}).$$

We see that this condition is identical to the strict probability-raising condition (SPR) since for a singleton $\mathit{Cause} = \{\mathbf{c}\}$ and each path $\pi \in \mathit{Paths}_{\mathit{fin}}(\mathcal{M})$ ending in $\mathbf{c}$ we have

$$\Pr(\Diamond \mathit{Eff} \mid \pi) = \Pr(\Diamond \mathit{Eff} \mid \neg\{\mathbf{c}\} \,\mathrm{U}\, \mathbf{c}) = \Pr_{\mathbf{c}}(\Diamond \mathit{Eff}).$$

Moreover, in [KM09] prima facie causes have temporal priority over the effect by explicitly requiring at least one time step to be in between the satisfaction of the PCTLstate formulas corresponding to cause and effect. For us, this is captured in condition (1) by requiring disjointedness of the sets Cause and Eff. We again see that SPR causes also entail these principles by having Cause and Eff disjoint.

Furthermore, in our setting the effect set Eff is terminal and thus the cause has temporal priority if it occurs. The reachability criteria for Cause is identical to ours. However, the formulation given by Kleinberg et al. does not entail the minimality condition for the cause as some states $\mathbf{c} \in \mathit{Cause}$ may not satisfy $\Pr((\neg \mathit{Cause}) \,\mathrm{U}\, \mathbf{c}) > 0$.

With this discussion we see that the notion of prima facie causes by Kleinberg et al. constitutes a special case of SPR causes in DTMCs if all states $\mathbf{c} \in \mathit{Cause}$ are reachable from init by path satisfying $(\neg \mathit{Cause}) \,\mathrm{U}\, \mathbf{c}$.

p-Causes in DTMCs. In [Zie+22] we define another notion of probabilistic causes inspired by the prima facie causes of Kleinberg et al. [KM09]. There, the considered Markov model $\mathfrak{M}$ is again a DTMC and the effect is an ω-regular property rEff. In [Zie+22] a so called $\mathfrak{p}$-*cause* for a given $\mathfrak{p} \in (0, 1]$ is a set of finite paths Π such that

1. almost all words in rEff have a prefix in Π, and
2. every finite path $\pi \in \Pi$ satisfies $\Pr(\mathit{rEff} \mid \pi) \geqslant \mathfrak{p}$.

For a value of $\mathfrak{p} > \Pr(\mathit{rEff})$ this is a strict probability-raising condition requiring probability-raising for each element of the cause. However, the probability for the satisfaction of an ω-regular property in a DTMCs is always either 0 or 1 after a finite number of steps, i.e. whenever an accepting or refuting bottom strongly connected component (BSCC) in the product of the DTMC and a deterministic automaton for the property is entered. Thus, the non-strict inequality has the consequence that $\mathfrak{p}$-causes always exist in a DTMC for which $\Pr(\mathit{rEff}) > 0$, which is not the case for PR causes.

However, we can compare the notions of $\mathfrak{p}$-cause to the probability-raising causes of this thesis in the following way: Consider $\mathit{rEff} = \Diamond \mathit{Eff}$ for a set of terminal states $\mathit{Eff} \subset \mathrm{S}$ and consider the set of finite paths Π which satisfy $\Diamond \mathit{Cause}$ for a set $\mathit{Cause} \subseteq \mathrm{S} \setminus \mathit{Eff}$ with $\Pr((\neg \mathit{Cause}) \,\mathrm{U}\, \mathfrak{c}) > 0$ for all $\mathfrak{c} \in \mathit{Cause}$. Then, the following are equivalent

- For some $\mathfrak{p} > \Pr(\Diamond \mathit{Eff})$ the set of finite paths Π is a $\mathfrak{p}$-cause for $\Diamond \mathit{Eff}$.
- The set Cause is an SPR cause for Eff and we have $\Pr(\Diamond \mathit{Cause} \mid \Diamond \mathit{Eff}) = 1$.

In the second statement, the conditional probability being 1 corresponds to the complete coverage of Π for $\Diamond \mathit{Eff}$. In other words, almost all paths reaching Eff must also reach Cause.

We will later extend our state-based notion of probability-raising causes to co-safety properties as cause and ω-regular properties as effects in Sections 5.2 and 5.3. There, we can observe that a minimal good prefix of a co-safety strict probability-raising cause in a Markov chain corresponds to a finite path of a $\mathfrak{p}$-cause.

Probabilistic Causality as a Hyperproperty. Ábrahám et al [ÁB18] introduce a probabilistic hyperlogic for Markov chains called HyperPCTL. This logic permits to formalize a very basic approach to probability-raising as a formula, which is consistent with Kleinberg's prima facie causes, and with strict probability-raising causes up to minimality. Thus, the above comparison holds for this related notion.

More generally probabilistic causality in MDPs was considered briefly in [DFT20] by Dimitrova et al. in a motivating example. While, the paper itself focuses on building a foundation for the introduced probabilistic hyperlogic PHL, in their first example [DFT20, Example 1] they illustrate how action causality can be formalized as a PHL formula.

Roughly, the presented formula expresses that taking a specific action α increases the probability for reaching effect states. Thus, it also relies on the probability-raising principle, but compares the "effect probabilities" under different schedulers which are either allowed to schedule α or can never schedule α rather than comparing probabilities under the same scheduler as in our PR condition. Even though our GPR and SPR causes are purely state-based with probability-raising conditions which compare probabilities under the same scheduler, we argue, however, that our notions of PR causes can reason about action causality as well.

Example 20 (Action Causality). Suppose, the task is to check whether taking action $\alpha \in Act$ in state $s \in S$ raises the effect probabilities compared to never scheduling α in state s as in Example 1 of [DFT20]. In order to argue about this causality, we make the strong assumption that there are no cycles in $\mathcal{M}$ containing s. Thus, choosing action α in s once is equivalent to choosing α in s at least once.

We now consider two copies of $\mathcal{M}$, namely $\mathcal{M}_0$ and $\mathcal{M}_1$, with the following modifications: In $\mathcal{M}_0$, the only enabled action of state s is α, while in $\mathcal{M}_1$ the enabled actions of state s are the elements of $Act_{\mathcal{M}}(s) \setminus \{\alpha\}$. We the consider the MDP $\mathcal{N}$ with a fresh initial state $init$ from which transitions with equal probabilities $1/2$ go to the copies of s in $\mathcal{M}_0$ and $\mathcal{M}_1$. The action α raises the effect probability in $\mathcal{M}$ if and only if for all schedulers $\mathfrak{S}$ for $\mathcal{N}$ the copy of s in $\mathcal{M}_0$ satisfies (SPR) for the effect set $Eff = Eff_0 \cup Eff_1$ in $\mathcal{N}$, where Eff_0 and Eff_1 are the copies of the effect set in $\mathcal{M}_0$ and $\mathcal{M}_1$ respectively.

We demonstrate this construction by considering the MDP $\mathcal{M}$ from Figure 3.5. We are interested whether taking α in s raises the probability to reach the effect state eff. The constructed MDP $\mathcal{N}$ with two adapted copies of $\mathcal{M}$ is depicted in Figure 3.6. For all schedulers $\mathfrak{S}$ of $\mathcal{N}$ the state s_0 satisfies (SPR) by

$$\Pr^{\mathfrak{S}}_{\mathcal{N}}(\Diamond\{eff_0, eff_1\} \mid \Diamond s_0) = \frac{1}{4} > \frac{1}{8} = \Pr^{\mathfrak{S}}_{\mathcal{N}}(\Diamond\{eff_0, eff_1\}),$$

which means that the action α does indeed raise the probability of reaching eff in $\mathcal{M}$.

◁

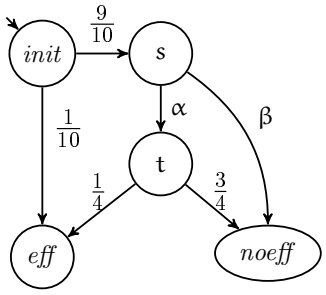

Figure 3.5: MDP $\mathfrak{M}$ where we consider action causality

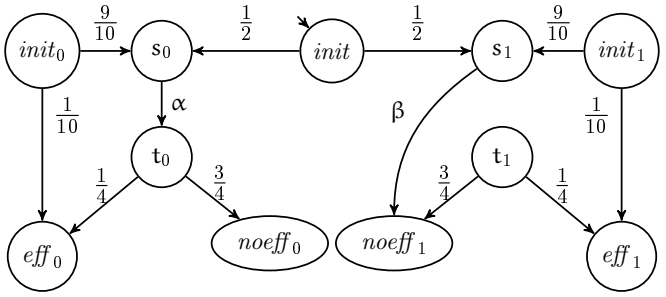

Figure 3.6: $\mathfrak{N}$ with the two copies $\mathfrak{M}_0$ and $\mathfrak{M}_1$

Let us note that the idea presented by this example can be generalized to check whether scheduler classes satisfying a regular constraint have higher effect probability compared to all other schedulers. In this case, we deal with an MDP $\mathfrak{N}$ as above, where $\mathfrak{M}_0$ and $\mathfrak{M}_1$ are defined as the synchronous product of deterministic finite automata and $\mathfrak{M}$.

In conclusion, our definition of probability-raising cause in MDPs (Def 14) aligns with other existing approaches to probabilistic causality in Markovian models.

3.2 Examples

In this section we will demonstrate applications of the framework of probability-raising causality in example MDPs. On one hand, there are a lot of examples in the philosophical literature on probabilistic causality which exemplify both the up- and downsides of cause-effect notions described using probability-raising as in (PR) paired with temporal priority of causes [Car79; Eel91; Hit21]. By nature, these examples are artificial but reflect real world examples, as their core purpose is to show properties of probabilistic causality. We replicate some properties by modeling philosophical examples as MDPs and considering GPR and SPR causes.

On the other hand, we will consider the Racetrack MDP [BBS95] which originated in AI autonomous decision making. In particular we will concentrate on specific instances with suitable side constraints. The Racetrack MDPs bridge challenges in autonomous driving and learning by neural networks to the world of probabilistic modeling via Markov decision processes [Gro+20]. With this example we see an application of probability-raising causality in MDPs on an example, which is also used in the literature of formal systems.

3.2.1 Conceptual Examples

We start with some abstract examples inspired by literature on probabilistic causality, e.g. [Car79; Eel91; Kle12]. With these examples we showcase properties of probability-raising and demonstrate why it makes sense to consider PR causality in MDPs. In other words we provide “proof-of-concept” examples. Thus, the probabilities given here do not resemble actual studies.

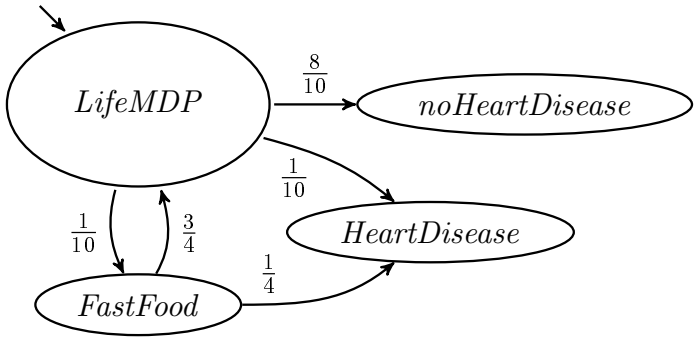

Figure 3.7: DTMC showing a causal relation between fast food and heart diseases

Example 21 (Temporal Priority). For the running example we take inspiration from an example given by [Car79] for causes of heart diseases. We assume an MDP which abstractly describes the probabilistic dependencies between choices in the life of an individual. As there are only a few parts of this MDP which we are interested in we collapse the remaining MDP into a single state called $\mathit{LifeMDP}$. For simplicity we assume there is a probability of $\frac{1}{10}$ to start a habit of eating fast food, represented by a state $\mathit{FastFood}$. For a person with such a habit we assume a $\frac{1}{4}$ chance to get a heart disease and a $\frac{3}{4}$ chance of stopping this habit. We also assume a $\frac{1}{10}$ chance of getting a heart disease while not having a habit of consuming fast food and a $\frac{1}{4}$ chance with the habit. This results in the DTMC represented in Figure 3.7.

So for $\mathit{Eff} = \{\mathit{HeartDisease}\}$ and $\mathit{Cause} = \{\mathit{FastFood}\}$ we derive the equations

$$\Pr(\Diamond \mathit{Eff}) = \frac{1}{10} + \frac{1}{10} \cdot \left(\frac{1}{4} + \frac{3}{4} \cdot \Pr(\Diamond \mathit{Eff})\right),$$
$$\Pr(\Diamond \mathit{Eff} \mid \Diamond \mathit{Cause}) = \frac{1}{4} + \frac{3}{4} \cdot \Pr(\Diamond \mathit{Eff}),$$

which imply

$$\Pr(\Diamond \mathit{Eff} \mid \Diamond \mathit{Cause}) = \frac{13}{37} > \frac{5}{37} = \Pr(\Diamond \mathit{Eff}).$$

Thus, $\mathit{FastFood}$ is a strict probability-raising cause for $\mathit{HeartDisease}$. Note, that Eff being terminal is important for such examples. Given an additional transition from $\mathit{HeartDisease}$ to $\mathit{LifeMDP}$ with a sufficiently small probability, while the remaining probability mass goes to a fresh terminal state $\mathfrak{t}$, we might deduce that $\mathit{HeartDisease}$ raises the probability of $\mathit{FastFood}$. Thus, the effect Eff consisting of terminal states forces *temporal priority* of probability-raising causes. This removes the symmetry between Eff and Cause in the definition of PR cause. However, the symmetry of the (PR) formula itself is still present. For example, it is still the case that $\Pr(\Diamond \mathit{HeartDisease} \mid \Diamond \mathit{FastFood}) > \Pr(\Diamond \mathit{HeartDisease})$. Furthermore, if we would add a further transition from $\mathit{noHeartDisease}$ to $\mathit{LifeMDP}$ then we would eventually reach $\mathit{Eff} = \{\mathit{HeartDisease}\}$ with probability 1 and there could be no probability-raising causality by Remark 18. Therefore, the existence of terminal non-effect states or end-components not containing effect states is also crucial for PR causality. ◁

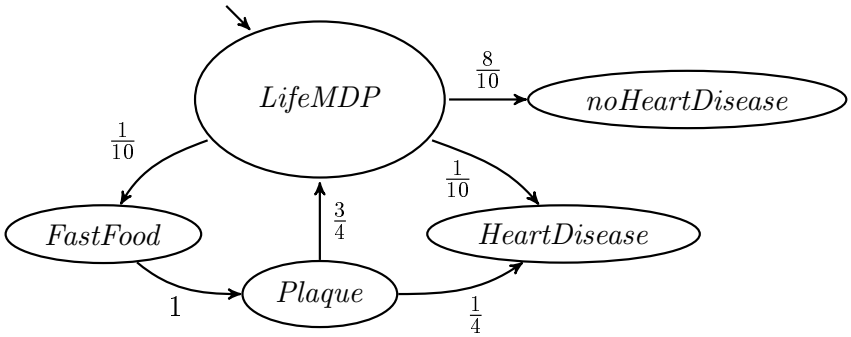

Figure 3.8: DTMC where plaque is screening of fast food from heard disease

Example 22 (Screening off). We now extend Example 21 by adding a detail to the relation between $\mathit{FastFood}$ and $\mathit{HeartDisease}$. In particular, we introduce a new state Plaque, which describes an individual having a build-up of plaque in their blood vessels. Instead of transitioning directly from a habit of consuming fast food to getting a heart disease we transition to this intermediate state with compatible probabilities. The resulting DTMC is depicted in Figure 3.8.

Here, we still have $\Pr(\Diamond \mathit{HeartDisease} \mid \Diamond \mathit{FastFood}) = \frac{13}{37} > \frac{5}{37} = \Pr(\Diamond \mathit{HeartDisease})$. However, we now have another strict probability-raising cause in $\{\mathit{Plaque}\}$, since

$$\Pr(\Diamond \mathit{HeartDisease} \mid \Diamond \mathit{Plaque}) = \frac{13}{37} > \frac{5}{37} = \Pr(\Diamond \mathit{HeartDisease}).$$

Moreover, in this example $\Diamond \mathit{FastFood}$ and $\Diamond \mathit{HeartDisease}$ are probabilistically independent conditional upon $\Diamond \mathit{Plaque}$. To see this, we use the definition of conditional probability

$$\Pr(\Diamond \mathit{HeartDisease} \wedge \Diamond \mathit{FastFood} \mid \Diamond \mathit{Plaque}) = \frac{\Pr(\Diamond \mathit{HeartDisease} \wedge \Diamond \mathit{FastFood} \wedge \Diamond \mathit{Plaque})}{\Pr(\Diamond \mathit{Plaque})}.$$

Abbreviating $\varphi = \Diamond \mathit{HeartDisease} \wedge \Diamond \mathit{FastFood} \wedge \Diamond \mathit{Plaque}$ we have

$$\begin{aligned} \Pr(\varphi) &= \frac{1}{10} \cdot \frac{3}{4} \cdot \Pr(\varphi) + \frac{1}{10} \cdot \left(\frac{3}{4} \cdot \frac{1}{10} + \frac{1}{4} \right) \\ \implies \Pr(\varphi) &= \frac{13}{370} \end{aligned}$$

and thus

$$\begin{aligned} \Pr(\Diamond \mathit{HeartDisease} \wedge \Diamond \mathit{FastFood} \mid \Diamond \mathit{Plaque}) &= \frac{13}{370} / \frac{1}{10} = \frac{13}{37}, \text{ and} \\ \Pr(\Diamond \mathit{HeartDisease} \mid \Diamond \mathit{Plaque}) \cdot \Pr(\Diamond \mathit{FastFood} \mid \Diamond \mathit{Plaque}) &= \frac{13}{37} \end{aligned}$$

which by definition means that $\Diamond \mathit{FastFood}$ and $\Diamond \mathit{HeartDisease}$ are probabilistically independent conditional upon $\Diamond \mathit{Plaque}$, as claimed.

In probabilistic causality [Rei56] [Hit21] this can be formulated as

$$\mathit{Plaque} \text{ screening off } \mathit{FastFood} \text{ from } \mathit{HeartDisease}.$$

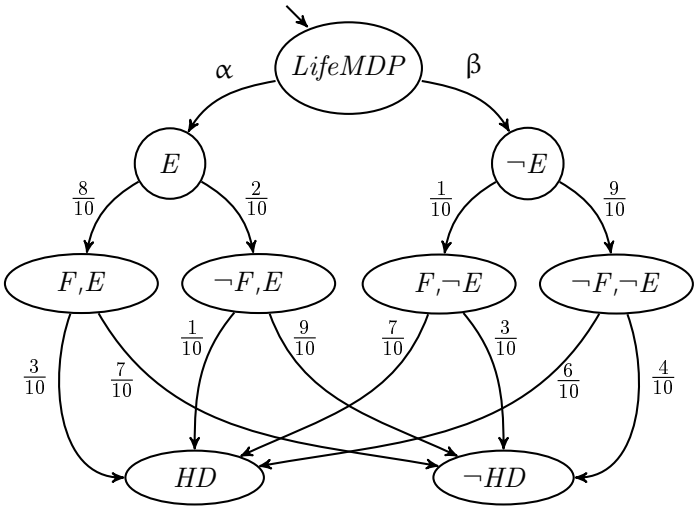

Figure 3.9: MDP exemplifying Simpson's paradox in sub-populations

In terminology of Reichenbach [Rei56] $Plaque$ is *causally between* $FastFood$ and $HeartDisease$. Given this example, we see that such phenomena can also occur for probability-raising causes by Definition 14. Whenever there is PR cause for Eff which is screening of another one, it still makes sense to consider both as causes for Eff. However, this indicates that these causes differ in how well they predict the effect. In order to distinguish between these, we use quality measures for PR causes introduced in Chapter 3.3. Moreover, we will later discuss an approach in which an SPR cause is described which does not screen off any other SPR cause (cf. Remark 71). ◁

Example 23 (Simpson's Paradox). We will now consider the same causal relation between consuming fast food and getting a heart disease with an additional twist, as we introduce the habit of exercising into the MDP. So, we consider states in which the individual is exercising regularly, labeled E, and where it does not, labeled $\neg\mathsf{E}$. In order to reach those states we have non-deterministic transitions with actions α and β respectively implying that it is purely non-probabilistic whether an individual is exercising. For the habit of consuming fast food, labeled F, we assume we are given statistical data on the distributions of such behavior given either E or $\neg\mathsf{E}$. Furthermore, we also have probabilities for each sub-population on how likely they get a heart disease, labeled HD. This MDP is represented by Figure 3.9.

We consider $Cause = \{(\mathsf{F},\mathsf{E}),(\mathsf{F},\neg\mathsf{E})\}$ and $Eff = \{\mathsf{HD}\}$. For the two deterministic scheduler $\mathfrak{S}_\alpha$ and $\mathfrak{S}_\beta$ choosing α or β in $LifeMDP$ respectively we note that

$$\Pr^{\mathfrak{S}_\alpha}(\Diamond Eff \mid \Diamond Cause) = \frac{3}{10} > \frac{8}{10}\cdot\frac{3}{10} + \frac{2}{10}\cdot\frac{1}{10} = \frac{26}{100} = \Pr^{\mathfrak{S}_\alpha}(\Diamond Eff),$$
$$\Pr^{\mathfrak{S}_\beta}(\Diamond Eff \mid \Diamond Cause) = \frac{7}{10} > \frac{1}{10}\cdot\frac{7}{10} + \frac{9}{10}\cdot\frac{6}{10} = \frac{61}{100} = \Pr^{\mathfrak{S}_\beta}(\Diamond Eff).$$

This shows that in the two sub-populations of exercisers and non-exercisers consuming fast food indeed raises the probability of getting a heart disease.

But, if we consider a randomized scheduler $\mathfrak{S}$ which chooses α and β with equal probability of $\frac{1}{2}$, we have

$$\Pr^{\mathfrak{S}}(\Diamond \mathit{Eff} \mid \Diamond \mathit{Cause}) = \frac{31}{90} = 0.3\overline{4} < 0.435 = \frac{87}{200} = \Pr^{\mathfrak{S}}(\Diamond \mathit{Eff}).$$

So, for $\mathfrak{S}$ consuming fast food in fact lowers the probability of getting a heart disease in a population where there are an equal amount of individuals are exercising and not exercising. Such a reversal of probabilistic inequalities through additional data in sub-populations is also called *Simpson's paradox* or *Simpson's reversal* [Sim51; Hit21]. We note, that in this example Cause is not a GPR cause as the randomized scheduler $\mathfrak{S}$ is refuting (GPR) and thus it is also not an SPR cause by Proposition 15. ◁

Remark 24 (Simpson's Paradox and Statistical Data.). Even though Simpson's paradox creates problems for theories of probabilistic causality, as noted by Cartwright [Car79] and Skyrms [Sky81], such problems are much more prevalent in statistical data. By nature, MDPs are inherently inappropriate for representing statistical data and dependencies. This can already be seen in the small Example 23. There, we have to encode the "history" of the data into the state space resulting in 4 states for 2 distinct features. In most cases, this is not a satisfactory method as one of the key properties of Markovian models is the Markov property that transitions are *not* history dependent. One approach to tackle such problems is to consider the additional data which brings about the reversal of conditional probabilities as a feature. There are also accounts on feature causality which can also be applied to MDPs [Dub+22; Dub21]. We note that in our theory of probability-raising causality we can also encounter Simpson's paradox. However, when the Markovian model in question does not encode the history of the execution the probabilities for (GPR) and (SPR) are not dependent on the history of the execution. In such cases, we do not run into such reversals of conditional probabilities. ◁

With the conceptual examples considered here, we show that Definition 14 reflects the ideas of probabilistic causality formulated in philosophical literature. Moreover, we demonstrate that phenomena of probabilistic causality can also occur setting of Markov decision processes in Examples 22 and 23. For Simpson's Paradox we argue that some of the critics can be handled if the MDP does not encode the history of the execution in the state space. However, the phenomena of causal chains and screening off can still occur in our definition. In the later Chapter 3.3 we describe a possible approach for limiting their impact by introducing quality measures for PR causes. These quality measures can then be used to compare different PR causes.

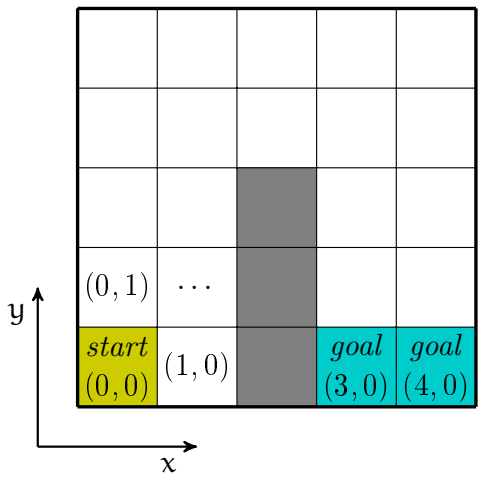

Figure 3.10: A discrete grid of a tiny Racetrack

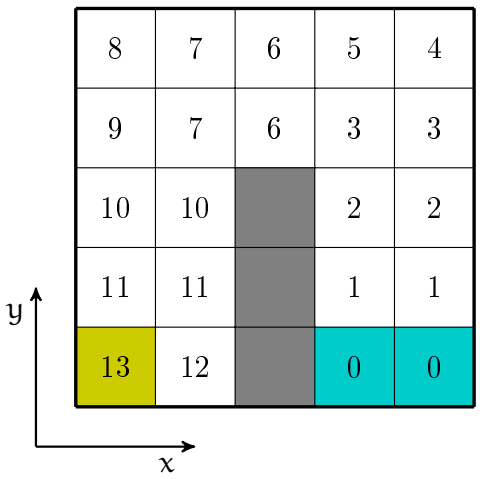

Figure 3.11: The value function for Racetrack

3.2.2 Racetrack MDP

Taken from literature on formal systems as described at the start of this section, we will now take a closer look at the Racetrack MDP as introduced in [BBS95]. A Racetrack consists of a track grid with boundaries and walls, a set of starting coordinates and a set of goal coordinates. Intuitively, the corresponding MDP describes the movement of a car along the track.

The control of the car is described by the enabled actions while the probabilistic part comes from an unreliability, e.g. it may happen that a chosen acceleration is not applied to the car. For simplicity we will stick to a specific track. In this example we address how to find and describe PR causes with simple methods, but we will also notice that more sophisticated algorithms for checking and deriving causality relations are needed. We will start with the formal definition of Racetrack and explain the intuitive idea thereafter.

We consider the track grid depicted in Figure 3.10. In this track there are 25 coordinates (x, y) from which one is the $start = (0,0)$ in yellow and two others are the $goal = \{(3,0),\ (4,0)\}$ in blue. The boundary is described by all coordinates which are not part of the track. The set of these coordinates is denoted by $Wall$. Additionally there are three impassable coordinates marked as gray fields which we add to the $Wall$ coordinates. Thus,

$$Wall = \{\ (\mathrm{x}, \mathrm{y}) \mid \mathrm{x} \notin \{0,1,2,3,4\} \text{ or } \mathrm{y} \notin \{0,1,2,3,4\}\ \} \cup \{(2,0),\ (2,1),\ (2,3)\}.$$

Moreover, we also include a value function which roughly describes the distance to $goal$. Each coordinate $(\mathrm{x}, \mathrm{y}) \notin Wall$ has the value $val(\mathrm{x}, \mathrm{y})$ depicted in Figure 3.11 of the track. Coordinates $(\mathrm{x}, \mathrm{y}) \in Wall$ have $val(\mathrm{x}, \mathrm{y}) = \infty$.

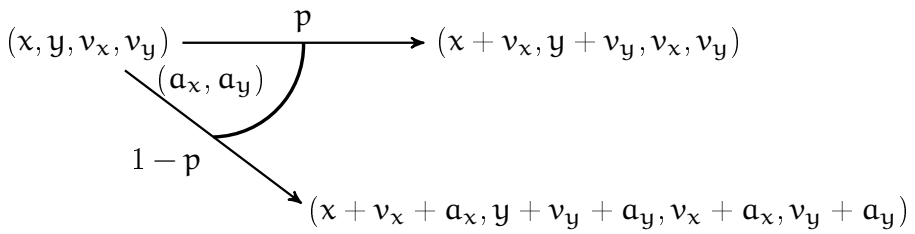

Figure 3.12: Visual representation of the transition function P from a Racetrack

We define the associated MDP $\mathcal{M} = (S, Act, P, init)$ to Figure 3.10 and Figure 3.11 by the set of states

$$S = \{\, (x, y, v_x, v_y) \mid x, y \in \{0, 1, 2, 3, 4\},\ v_x, v_y \in \{-2, -1, 0, 1, 2\} \,\} \cup \{crash\},$$

the initial state $init = (0, 0, 0, 0)$ and the set of actions

$$Act = \{(a_x, a_y) \mid a_x, a_y \in \{-1, 0, 1\}\} \cup \{\beta\}.$$

Here, a state $s \in S$ consists of coordinates x and y as well as velocities v_x and v_y in both directions. An action $\alpha \in Act \setminus \{\beta\}$ consists of values for the chosen acceleration a_x and a_y in both directions. For a given state $s = (x, y, v_x, v_y) \in S$ we define the set of allowed accelerations in s as

$$A(s) = \{(a_x, a_y) \in Act \mid val(x + v_x + a_x, y + v_y + a_y) > val(x, y)\} \setminus A_{crash}(s),$$

where $A_{crash}(s)$ is the set of actions $(a_x, a_y) \in Act$ such that for $x' = v_x + a_x$ and $y' = v_y + a_y$ we have for some $\lambda \in (0, 1)$

$$(x + \lambda \cdot x',\ y + \lambda \cdot y') \in \bigcup_{(x_w, y_w) \in Wall} \left\{ (x_l, y_l) \;\middle|\; \begin{array}{l} x_l \in [x_w - 0.5,\ x_w + 0.5] \\ y_l \in [y_w - 0.5,\ y_w + 0.5] \end{array} \right\}.$$

So, A_{crash} describes accelerations which lead to a crash. For a state $s = (x, y, v_x, v_y)$ the set of enabled actions is

$$Act(s) = \begin{cases} A(s) & \text{if } A(s) \neq \emptyset \\ \{\beta\} & \text{else} \end{cases},$$

where, β is only active, whenever a crash is inevitable. For $A(s) \neq \emptyset$ the transition function P is defined according to Figure 3.12, where for a state $s = (x, y, v_x, v_y)$ the action $(a_x, a_y) \in A(s) \subseteq Act$ is taken. On the other hand, for any state $s \in S$ with $\beta \in Act(s)$ we have $P(s, \beta, crash) = 1$. All states with coordinates $(x, y) \in goal$ are terminal and have no outgoing transitions. By abuse of notation we use coordinates (x, y) to describe sets of states in the MDP $\mathcal{M}$, i.e. (x, y) stands for the set $\{(x, y, v_x, v_y) \mid v_x, v_y \in \{-2, -1, 0, 1, 2\}\}$.

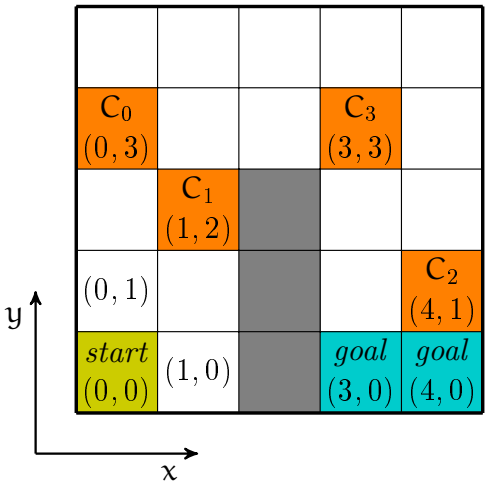

Figure 3.13: The coordinates considered as cause candidates

With this technical definition in mind, let us describe the idea behind the Racetrack MDP. A car is starting in the coordinates $start$ with no velocity on the x or y axis. In each state a controller can decide to accelerate with a_x in on the x-axis and with a_y on the y axis. However, by definition of $Act(s)$ the controller is required to accelerate in such a way that the intended acceleration is moving towards goal by decreasing the value of the current coordinate in each step.

In particular, the car can not stand still for more than one step. Moreover, as $val(x, y) = \infty$ for any coordinate in $Wall$ and by the definition of $A_{crash}(x, y)$ the controller is also required to avoid directly crashing into the wall. This means, that whenever a $Wall$ coordinate is next ot the car, the controller can not accelerate in the direction of this $Wall$ coordinate. This way we restrict the scheduler of $\mathfrak{M}$ to those which actively try to reach the goal.

The position and velocity of the car will update according to the chosen acceleration with probability $(1 - p)$ and with probability p the MDP ignores the acceleration, as if the road is slippery. In the later case, the position of the car will still update according to the current velocity. The probability $p \in [0, 1]$ then corresponds to the unreliability of the system. Whenever the update ends up crashing the car into a wall we enter the terminal state $crash$. We also bounded the number of states by bounding the velocity to an absolute of 2 since accelerating to a higher velocity is not possible by the value function anyway. Intuitively, higher velocity than 2 would lead to a crash in any scenario and thus the corresponding states would be unreachable. In fact there are more unreachable states. However, computing these would need a computer based analysis, which we omit here.

We now apply the framework of PR causes here for the effect of reaching goal. For this we consider a unreliability of $p = 0.6$ as if the road is almost completely made of ice. We then consider the effect set $Eff = \{goal\}$. For now we will consider fixed sets of cause candidates,

namely

$$\begin{aligned}
\mathrm{C}_0 &= \{(0,3,v_x,v_y) \mid v_x, v_y \in \{-2,-1,0,1,2\}\},\\
\mathrm{C}_1 &= \{(1,2,v_x,v_y) \mid v_x, v_y \in \{-2,-1,0,1,2\}\},\\
\mathrm{C}_2 &= \{(4,1,v_x,v_y) \mid v_x, v_y \in \{-2,-1,0,1,2\}\},\\
\mathrm{C}_3 &= \{(3,3,v_x,v_y) \mid v_x, v_y \in \{-2,-1,0,1,2\}\},
\end{aligned}$$

which correspond to the coordinates as depicted in Figure 3.13. For C_0 we have

$$\Pr^{\max}(\Diamond \mathrm{C}_0) = 1,$$

realized by an MD scheduler $\mathfrak{S}$ choosing $(0,1)$ in start as long as there is no velocity leading to eventually accelerating to velocity 1 on the y-axis. By then choosing the action $(0,0)$ we ensure to not accelerate further leading to reaching C_0 with probability 1. Afterwards the scheduler $\mathfrak{S}$ can be arbitrary. Thus, for any such scheduler $\mathfrak{S}$ we have

$$\Pr^{\mathfrak{S}}(\Diamond \mathit{goal} \mid \Diamond \mathrm{C}_0) = \Pr^{\mathfrak{S}}(\Diamond \mathit{goal})$$

and C_0 is no GPR cause and therefore also no SPR cause (cf. Rem. 18 and Prop. 15).

For C_1 the situation is not as clear as no scheduler can reach C_1 with probability 1. Moreover, no scheduler $\mathfrak{S}$ can realize $\Pr^{\mathfrak{S}}(\Diamond \mathit{goal}) = 1$. However, we can construct a scheduler which refutes the GPR condition **(G)** . Consider the scheduler $\mathfrak{S}$ choosing action $(0,1)$ in start until eventually accelerating such that it reaches state $(0,1,0,1)$ with probability 1. In this state action $(1,0)$ is taken transitioning to state $\mathbf{c}_1 = (1,2,1,1) \in \mathrm{C}_1$ with probability 0.4 and with probability 0.6 to $\mathbf{s} = (0,2,0,1)$. In the later case, the scheduler tries to maximize the probability to reach goal after reaching $\mathbf{s}$. If $\mathbf{c}_1$ is reached then $\mathfrak{S}$ chooses action $(0,1)$ leading to $(1,4,0,2)$ with probability 0.4 and to crash with probability 0.6. In $(1,4,0,2)$ only action β is enabled leading to crash. Therefore, we have

$$\Pr^{\mathfrak{S}}(\Diamond \mathit{goal} \mid \Diamond \mathrm{C}_1) = 0 < \Pr^{\mathfrak{S}}(\Diamond \mathit{goal}),$$

and $\mathfrak{S}$ is refuting the GPR condition **(G)** . By the existence of such a scheduler, C_1 is no GPR or SPR cause. The underlying idea here can be used for checking PR causality in general as the constructed scheduler $\mathfrak{S}$ is a kind of worst-case scheduler which minimizes the probability to reach goal after a state $\mathbf{c} \in \mathit{Cause} = \mathrm{C}_1$ is reached. We will prove the soundness of this idea when considering the algorithmic approach to checking the PR conditions (Lemma 36).

For the analysis of C_2 we consider the following sufficient condition for **(G)** :

$$\Pr^{\min}(\Diamond \mathit{goal} \mid \Diamond \mathrm{C}_2) > \Pr^{\max}(\Diamond \mathit{goal}).$$

If this condition holds then for each scheduler $\mathfrak{S}$ of $\mathfrak{M}$ it holds that

$$\Pr^{\mathfrak{S}}(\Diamond \mathit{goal} \mid \Diamond \mathrm{C}_2) \geqslant \Pr^{\min}(\Diamond \mathit{goal} \mid \Diamond \mathrm{C}_2) > \Pr^{\max}(\Diamond \mathit{goal}) \geqslant \Pr^{\mathfrak{S}}(\Diamond \mathit{goal}).$$

Using the probabilistic model checker PRISM [KNP11] with the extension of conditional probabilities as in [Bai+17] we compute the following values

$$\Pr^{\min}(\Diamond \mathit{goal} \mid \Diamond \mathbf{C}_2) = 0.4260 > 0.4096 = \Pr^{\max}(\Diamond \mathit{goal}).$$

This implies GPR causality of $\mathbf{C}_2$. The last cause candidate consider here is $\mathbf{C}_3$ for which

$$\Pr^{\min}(\Diamond \mathit{goal} \mid \Diamond \mathbf{C}_3) = 0.1600 < 0.4096 = \Pr^{\max}(\Diamond \mathit{goal}).$$

Obviously, this does not necessarily mean that $\mathbf{C}_3$ is not a GPR cause. Here, we need stronger sufficient and necessary conditions for PR causality, which shows that more sophisticated methods of checking the GPR condition **(G)** are needed. These methods will be addressed in Chapter 4.

The analysis of singleton causes is simpler. Recall, for singletons the conditions **(S)** and **(G)** are equivalent. We now consider the state $\mathbf{c}_3 = (3, 3, 0, -1) \in \mathbf{C}_3$. For this state we have

$$\Pr^{\min}(\Diamond \mathit{goal} \mid \Diamond \mathbf{c}_3) = \Pr^{\min}_{\mathbf{c}_3}(\Diamond \mathit{goal}) = 0.6102 > 0.4096 = \Pr^{\max}(\Diamond \mathit{goal})$$

and thus $\mathbf{c}_3$ is a singleton PR cause. The idea to compare the minimal probability of reaching the effect starting in the cause state in question to the maximal probability to reach the effect from the initial state is used in the later Algorithm 41 for checking the SPR condition **(S)** efficiently.

Let us go back to $\mathbf{C}_2$ which we have proven to be a GPR cause. We will show that there can also be states $\mathbf{c}_2 \in \mathbf{C}_2$ which are not a singleton SPR cause. Indeed, consider the state $\mathbf{c}_2 = (4, 1, 1, -1) \in \mathbf{C}_3$. In $\mathbf{c}_2$ there is only one enabled action, namely $(-1, 0)$ leading to goal with probability 0.4 and to crash with probability 0.6. So, consider any scheduler $\mathfrak{S}$ reaching $\mathbf{c}_2$ with positive probability. We then have

$$\Pr^{\mathfrak{S}}(\Diamond \mathit{goal} \mid \mathbf{c}_2) = \Pr^{\mathfrak{S}}_{\mathbf{c}_2}(\Diamond \mathit{goal}) = 0.4 < 0.4096 = \Pr^{\max}(\Diamond \mathit{goal}).$$

Although this does not imply that $\mathbf{c}_2$ is not a SPR cause, it gives us an indication for SPR causality as we see later in Section 4.1.2.

As these considerations show, the notions of GPR and SPR causes are non-trivial and require careful analysis. Thus, there are some algorithmic questions which need to be addressed. On one hand, we need to check the conditions **(S)** and **(G)** for a given cause candidate, as already stated above. On the other hand, we want to compute suitable SPR- or GPR causes for a given effect. However, as we already see in the Racetrack MDP, there can be many candidates for PR causes and it is not clear which one should be the result of such a computation. For example consider the coordinates $\mathbf{C}_2$ and $\mathbf{C}_3$ as possible PR causes. We can intuitively see that $\mathbf{C}_2$ is more precise, since it is "closer" to goal and thus reaching $\mathbf{C}_2$ can more precisely predict that the run will reach goal. However, reaching $\mathbf{C}_3$ seems to cover more runs which will eventually reach goal and thus covers more runs than $\mathbf{C}_2$. We will make these ideas more precise in the following section by introducing measures which rate the quality of a cause with respect to a certain criteria.

3.3 Quality Measures for Probability-Raising Causes

As stated in the beginning of this monograph the goal of introducing the probability-raising causality framework into MDPs is to have a tool which strengthens the understanding of probabilistic systems without the need of full knowledge about the system. To this end, we need to make quantitative statements about the explanatory power of a cause.

As Remark 17 suggests, it might happen that no probability-raising causes exist. However, the existence of PR causes can be characterized by the following result.

Lemma 25 (Criterion for the Existence of PR Causes). *Let $\mathcal{M}$ be an MDP and Eff a nonempty set of terminal states. The following statements are equivalent:*

(a) There is a set $\mathit{Cause} \subseteq S \setminus \mathit{Eff}$ which constitutes an SPR cause for Eff in $\mathcal{M}$.

(b) There is a set $\mathit{Cause} \subseteq S \setminus \mathit{Eff}$ which constitutes a GPR cause for Eff in $\mathcal{M}$.

(c) There is a state $\mathbf{c}_0 \in S \setminus \mathit{Eff}$ such that the singleton $\{\mathbf{c}_0\}$ is an SPR cause (and therefore a GRP cause) for Eff in $\mathcal{M}$.

As the proof needs some preparatory technical results, we will postpone it to Chapter 4. However, this result shows that the for existence of PR causes for an effect Eff in an MDP $\mathcal{M}$ it suffices if there is a single state which raises the probability of reaching Eff for all schedulers. A naive existence check would then test the SPR condition (also adressed later) for each state and yield a singleton state as a certificate for the existence. Furthermore, it is clear that the explanatory power of a single state can be limited, depending on the structure of the MDP.

Example 26 (Singleton Causes may not Cover the Effect very Well). For an example of this consider again the Racetrack MDP $\mathcal{M}$ from Section 3.2.2 described by Figures 3.10 and 3.12. For the effect $\mathit{Eff} = \{\mathit{goal}\}$ we consider the singleton $\mathbf{c} = (3, 1, 0, -1)$ describing the left position in front of goal moving towards the goal with a velocity of 1. It is not possible to crash from this situation and thus we have $\Pr^{\min}(\Diamond \mathit{Eff} \mid \Diamond \mathbf{c}) = 1$. Therefore, $\mathbf{c}$ is a singleton PR cause. In perspective of how well this singleton is predicting the effect, we can observe that it is sufficient to reach $\mathbf{c}$ (by the above equation) for the effect to happen. However, it is clear that this singleton only describes a very specific scenario. ◁

The goal of this chapter is to identify notions that measure how "good" causes are. As singletons or small cause sets might have poor coverage in the sense that the probability for paths that reach an effect state without visiting a cause state before ("uncovered effects") can be large, one intuitive approach would be to rate the quality of the cause by the size of the cause set in terms of cardinality. Even though this might be a usable criterion, actual usability of this still depends on the structure of the MDP.

Path hits	Eff	$\neg\mathit{Eff}$
Cause	True positive (tp) Cause correctly predicted Eff	False positive (fp) Cause falsely predicted Eff
$(\neg\mathit{Cause})$	False negative (fn) Cause falsely not predicted Eff	True negative (tn) Cause correctly not predicted Eff

Figure 3.14: Confusion matrix for Cause as a binary classifier for Eff

3.3.1 PR Causes as Binary Classifiers in MDPs

In our approach we consider the set Cause as a binary classifier which predicts whether or not a path reaches the set Eff. This classification by Cause results in the so called confusion matrix which is shown in Figure 3.14.

With this, we take inspiration from statistical analysis to consider quality measures for binary classifiers (see e.g. [Pow11]). In the context of this thesis, we focus on three distinct measures. This includes the *recall* as a measure for the relative coverage (proportion of covered effects among all effect scenarios), the *coverage ratio* (quotient of covered and uncovered effects) as well as the *f-score*. For the definition of the f-score, the precision as an additional measure is needed. The f-score is a standard measure for binary classifiers defined by the harmonic mean of precision and recall. It can be seen as a compromise to achieve both good precision and good recall.

Formally, as before we are given an MDP $\mathcal{M} = (\mathrm{S}, \mathit{Act}, \mathrm{P}, \mathit{init})$ and a set of terminal effect states $\mathit{Eff} \subseteq \mathrm{S}$. Furthermore, we suppose all states $\mathrm{s} \in \mathrm{S}$ are reachable from init. Given a set $\mathit{Cause} \subseteq \mathrm{S} \setminus \mathit{Eff}$, we define the following abbreviations:

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Cause} \wedge \Diamond \mathit{Eff}) = \mathit{tp}^{\mathfrak{S}}_{\mathcal{M}}, \qquad \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Cause} \wedge \neg\Diamond \mathit{Eff}) = \mathit{fp}^{\mathfrak{S}}_{\mathcal{M}},$$
$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\neg\Diamond \mathit{Cause} \wedge \Diamond \mathit{Eff}) = \mathit{fn}^{\mathfrak{S}}_{\mathcal{M}}, \qquad \Pr^{\mathfrak{S}}_{\mathcal{M}}(\neg\Diamond \mathit{Cause} \wedge \neg\Diamond \mathit{Eff}) = \mathit{tn}^{\mathfrak{S}}_{\mathcal{M}}.$$

We omit the subscript $\mathcal{M}$ if the MDP is clear from the context.

Remark 27 (Classifiers in DTMCs). If the MDP $\mathcal{M}$ in question is actually a DTMC, meaning there are only probabilistic choices and no non-determinism, then we note that all the defined values above are probabilities of conjunctions of reachability properties. Therefore, given a set Cause, these values can be computed in polynomial time and the confusion matrix can be derived. ◁

3.3.2 Quality Measures for a given PR Cause

We define the *precision*, *recall* and *coverage ratio* of a PR cause Cause under a scheduler $\mathfrak{S}$ by:

$$\mathit{precision}^{\mathfrak{S}}(\mathit{Cause}) \;=\; \frac{\mathit{tp}^{\mathfrak{S}}}{\mathit{tp}^{\mathfrak{S}} + \mathit{fp}^{\mathfrak{S}}}, \quad \mathit{recall}^{\mathfrak{S}}(\mathit{Cause}) \;=\; \frac{\mathit{tp}^{\mathfrak{S}}}{\mathit{tp}^{\mathfrak{S}} + \mathit{fn}^{\mathfrak{S}}}, \quad \mathit{covrat}^{\mathfrak{S}}(\mathit{Cause}) \;=\; \frac{\mathit{tp}^{\mathfrak{S}}}{\mathit{fn}^{\mathfrak{S}}}.$$

For these definitions we make some respective assumptions on the scheduler. We assume

- $\mathit{tp}^{\mathfrak{S}} + \mathit{fp}^{\mathfrak{S}} > 0$ for $\mathit{precision}^{\mathfrak{S}}(\mathit{Cause})$,
- $\mathit{tp}^{\mathfrak{S}} + \mathit{fn}^{\mathfrak{S}} > 0$ for $\mathit{recall}^{\mathfrak{S}}(\mathit{Cause})$ and
- $\mathit{fn}^{\mathfrak{S}} > 0$ for $\mathit{covrat}^{\mathfrak{S}}(\mathit{Cause})$.

If we have $\mathit{fn}^{\mathfrak{S}} = 0$ and $\mathit{tp}^{\mathfrak{S}} + \mathit{fn}^{\mathfrak{S}} > 0$ for some scheduler $\mathfrak{S}$, we define $\mathit{covrat}^{\mathfrak{S}}(\mathit{Cause}) = +\infty$. This makes sense since we can converge to such a scheduler $\mathfrak{S}$ with a sequence of randomized schedulers $\mathfrak{S}_0 \ldots$ for which $\mathit{fn}^{\mathfrak{S}_i} > 0$ for $i \in \mathbb{N}$. The coverage ratio of such a sequence converges to $+\infty$.

Note that recall and $\mathit{precision}$ are also conditional probabilities as we have

$$\mathit{precision}^{\mathfrak{S}}(\mathit{Cause}) \;=\; \Pr^{\mathfrak{S}}(\Diamond \mathit{Eff} \mid \Diamond \mathit{Cause}), \quad \mathit{recall}^{\mathfrak{S}}(\mathit{Cause}) \;=\; \Pr^{\mathfrak{S}}(\Diamond \mathit{Cause} \mid \Diamond \mathit{Eff}).$$

Finally, the *f-score* of Cause *under a scheduler* $\mathfrak{S}$ is the harmonic mean of the precision and recall. Formally we define

$$\mathit{fscore}^{\mathfrak{S}}(\mathit{Cause}) \;=\; \frac{2 \cdot \mathit{tp}^{\mathfrak{S}}}{2 \cdot \mathit{tp}^{\mathfrak{S}} + \mathit{fp}^{\mathfrak{S}} + \mathit{fn}^{\mathfrak{S}}} \;=\; 2 \cdot \frac{\mathit{precision}^{\mathfrak{S}}(\mathit{Cause}) \cdot \mathit{recall}^{\mathfrak{S}}(\mathit{Cause})}{\mathit{precision}^{\mathfrak{S}}(\mathit{Cause}) + \mathit{recall}^{\mathfrak{S}}(\mathit{Cause})}$$

Here, we assume $\mathit{tp}^{\mathfrak{S}} + \mathit{fp}^{\mathfrak{S}} + \mathit{fn}^{\mathfrak{S}} > 0$. For a GPR cause Cause this already implies $\mathit{tp}^{\mathfrak{S}} + \mathit{fn}^{\mathfrak{S}} > 0$ as we can reformulate the GPR condition for $\mathcal{M}$ to

$$\frac{\mathit{tp}^{\mathfrak{S}}}{\mathit{tp}^{\mathfrak{S}} + \mathit{fp}^{\mathfrak{S}}} \;>\; \mathit{tp}^{\mathfrak{S}} + \mathit{fn}^{\mathfrak{S}}. \tag{GPR*}$$

To lift the definitions of the quality measures under a scheduler to the quality measure of a cause, we consider the worst-case scheduler:

Definition 28 (Quality Measures for Causes). Let Cause be a PR cause. We define

$$\mathit{recall}(\mathit{Cause}) \;=\; \inf_{\mathfrak{S}} \; \mathit{recall}^{\mathfrak{S}}(\mathit{Cause}) \;=\; \Pr^{\min}_{\mathcal{M}}(\Diamond \mathit{Cause} \mid \Diamond \mathit{Eff})$$

when ranging over all schedulers $\mathfrak{S}$ with $\mathit{tp}^{\mathfrak{S}} + \mathit{fn}^{\mathfrak{S}} > 0$ or equivalently $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff}) > 0$. Likewise, the *coverage ratio* and *f-score* of Cause are defined by the worst-case coverage ratio resp. f-score when ranging over schedulers for which $\mathit{covrat}^{\mathfrak{S}}(\mathit{Cause})$ resp. $\mathit{fscore}^{\mathfrak{S}}(\mathit{Cause})$ is defined:

$$\mathit{covrat}(\mathit{Cause}) \;=\; \inf_{\mathfrak{S}} \; \mathit{covrat}^{\mathfrak{S}}(\mathit{Cause}), \qquad \mathit{fscore}(\mathit{Cause}) \;=\; \inf_{\mathfrak{S}} \; \mathit{fscore}^{\mathfrak{S}}(\mathit{Cause}).$$

◁

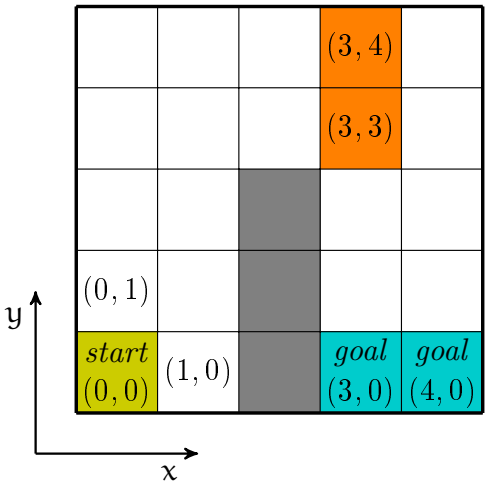

Figure 3.15: Racetrack coordinates with high recall value

Here, the choice to use the worst-case analysis over the non-determinism for the definition of quality measures is motivated by the fact that the Definition 14 of PR causes uses a universal quantification over scheduler which reach the cause set. By defining quality measures as the infimum over all possible scheduler, we ensure that the value of the measure is a lower bound for the quality of the cause for a given scheduler. In other circumstances (e.g. control over non-determinism c.f. Sec. 5.1), it also makes sense to use average-case or best-case analysis. However, for our case we will mainly focus on the worst-case definitions here.

Example 29 (Quality Measures in Racetrack). Let us consider the quality of causes in an Example. For this we again consider the Racetrack MDP $\mathcal{M}$ from Section 3.2.2 visualized in Figures 3.10, 3.11 and 3.13. Given the unreliability of the system is $\mathrm{p} = 0.6$, we already established the set C_2 as a GPR cause for $\mathit{Eff} = \{\mathit{goal}\}$. For this MDP we have

$$\mathit{recall}(\mathrm{C}_2) = \Pr^{\min}(\Diamond \mathit{goal} \mid \Diamond \mathrm{C}_2) = 0.058, \quad \mathit{precision}(\mathrm{C}_2) = \Pr^{\min}(\Diamond \mathit{Eff} \mid \Diamond \mathrm{C}_2) = 0.4260,$$

where the conditional probabilities were computed by an implementation of PRISM [KNP11] using the extension introduced in [Bai+17]. Since the unreliability of the track is fairly high, this shows that C_2 predicts the effect reasonably well with a precision above 0.4. However, with a recall below 0.1 it does not cover the effect very well, as there are, relative to the probabilities, many paths which reach Eff but do not reach Cause.

Let us also consider the coordinates $(3, 3)$ and $(3, 4)$ given by the Figure 3.15. That is, we consider the set $\mathrm{C} = \{(\mathrm{x}, \mathrm{y}, v_{\mathrm{x}}, v_{\mathrm{y}}) \in S \mid (\mathrm{x}, \mathrm{y}) = (3, 3) \text{ or } (\mathrm{x}, \mathrm{y}) = (3, 4)\}$ which includes C_3 and get

$$\mathit{recall}(\mathrm{C}) = \Pr^{\min}(\Diamond \mathit{goal} \mid \Diamond \mathrm{C}) = 0.3318, \quad \mathit{precision}(\mathrm{C}) = \Pr^{\min}(\Diamond \mathit{Eff} \mid \Diamond \mathrm{C}) = 0.062.$$

Since the recall of $\mathbf{C}$ is relatively high, we can see that $\mathbf{C}$ covers paths that reach Eff better than $\mathbf{C}_2$ from before. However, reaching $\mathbf{C}$ does not predict that the effect will be reached, since its precision is relatively low. This motivates the usage of fscore as it is desirable to balance the described properties of covering the effect and precisely predicting it. Deriving a quality-optimal cause here can lead to better understanding of the MDP. ◁

Besides the quality measures defined so far, which we will address in detail, there is a vast landscape of further quality measures for binary classifiers in the literature (for an overview, see, e.g., [Pow11]). One prominent example which has been claimed to be superior to the f-score recently [CJ20] is *Matthews correlation coefficient* (MCC). In terms of the entries of a confusion matrix (as in Figure 3.14), it is defined as

$$\mathrm{mcc}^{\mathfrak{S}}(\mathit{Cause}) = \frac{\mathit{tp}^{\mathfrak{S}} \cdot \mathit{tn}^{\mathfrak{S}} - \mathit{fp}^{\mathfrak{S}} \cdot \mathit{fn}^{\mathfrak{S}}}{\sqrt{(\mathit{tp}^{\mathfrak{S}} + \mathit{fp}^{\mathfrak{S}}) \cdot (\mathit{tp}^{\mathfrak{S}} + \mathit{fn}^{\mathfrak{S}}) \cdot (\mathit{tn}^{\mathfrak{S}} + \mathit{fp}^{\mathfrak{S}}) \cdot (\mathit{tn}^{\mathfrak{S}} + \mathit{fn}^{\mathfrak{S}})}} \tag{MCC}$$

In contrast to the f-score (as well as recall and coverage ratio), it makes use of all four entries of the confusion matrix. In our setting, we can assign the MCC to a Cause by again taking the infimum of the value over all schedulers $\mathfrak{S}$ for which $\mathrm{mcc}^{\mathfrak{S}}$ is defined. The MCC of a set Cause and the (GPR) condition have a strong connection as the MCC of Cause is positive if and only if Cause satisfies (GPR) for all scheduler of $\mathfrak{M}$. We will show this in the later Lemma 50, which arguments also apply to the values $\mathit{tp}^{\mathfrak{S}}, \mathit{tn}^{\mathfrak{S}}, \mathit{fp}^{\mathfrak{S}}$ and $\mathit{fn}^{\mathfrak{S}}$.

Like the MCC, a lot of the quality measures studied in the literature are algebraic functions (intuitively speaking, built from polynomials, fractions and root functions) in the entries of the confusion matrix (cf. [Pow11]). Therefore, we will also briefly consider general algebraic quality measures. For this, consider any algebraic function $\mathrm{f} : \mathbb{R}^4 \to \mathbb{R}$. That is $\mathrm{f}(\mathrm{x})$ satisfies some polynomial equation where the coefficients are polynomials in the entries of x. Then, we define for a set $\mathit{Cause} \subseteq S \setminus \mathit{Eff}$ and a scheduler $\mathfrak{S}$ of $\mathfrak{M}$ the value $\mathrm{f}^{\mathfrak{S}}(\mathit{Cause}) = \mathrm{f}(\mathit{tp}^{\mathfrak{S}}, \mathit{tn}^{\mathfrak{S}}, \mathit{fp}^{\mathfrak{S}}, \mathit{fn}^{\mathfrak{S}})$, if the later is well defined. Moreover, most of the standard quality measures have a simpler form as they can be represented by a ratio of two linear functions (cf. [Pow11]). Examples for such measures are the recall or the f-score.

With quality measures we have a tool to evaluate how "good" a given cause is with respect to the measure. Furthermore, by having different measures we have the flexibility to consider different aspects of the relation between cause and effect. For instance, having a value of $\mathit{recall}(\mathit{Cause}) = 1$ means that any path that reaches the effect Eff *necessarily* has to reach Cause beforehand. An example for this is depicted by the DTMC in Figure 3.16 where all paths leading to $\mathit{Eff} = \{\mathit{eff}\}$ reach $\mathit{Cause}_0 = \{\mathbf{c}_0\}$ before eff. On the opposite side of the spectrum, a value of $\mathit{precision}(\mathit{Cause}) = 1$ means that it is *sufficient* to know a path reaches Cause as afterwards Eff is reached with probability 1. This is also depicted in Figure 3.16 as the probability to reach eff after reaching $\mathit{Cause}_1 = \{\mathbf{c}_1\}$ is 1. For a more in depth consideration of necessity and sufficiency in transition systems see [Bai+22a].

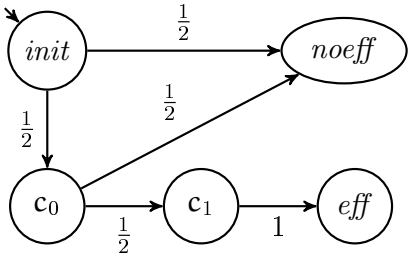

Figure 3.16: A DTMC where $\mathbf{c}_0$ is a necessary cause and $\mathbf{c}_1$ a sufficient cause for eff

As stated in the introduction of the chapter, a naive existence check for PR causes would only result in singleton causes as witness. Therefore, quality measures serve as a criterion to which a computed cause can be tailored. For the computation of causes we will mainly compute quality-optimal causes (Section 4.3).

A Summary and Outlook on Probability-Raising Causality

In this chapter we defined different notions of probabilistic cause in Markov decision processes by combining the probability-raising principle with temporal priority of causes. For this, we considered events as reachability properties of sets of states Eff and Cause. This resulted in two different notions of PR cause (Def 14) for which their respective PR condition needs to hold for all scheduler in the MDP. The notion of global PR cause is considering the set of states Cause as a unit in the (GPR) condition and is thus it also takes the distribution with which Cause is reached into account. The strict PR cause considers each state in Cause separately and requires probability-raising for each of them. This way, the (SPR) condition strictly imposes probability-raising regardless of how Cause is reached. We have discussed the soundness and meaning of such causes using conceptual examples inspired by philosophical literature on probabilistic causality as well as using the Racetrack MDP.

In the Racetrack MDP in particular it has become clear that there is a need for more sophisticated methods to check cause candidates Cause for the probability-raising conditions (SPR) and (GPR). Moreover, up to this point it is unclear how to derive an GPR (resp. SPR) cause for given a set of effect states Eff. In particular, in most cases there can be many different PR causes for a given effect and it is not clear, which one should be the result of a possible computation. To address this, we introduced quality measures for probability-raising causes by considering them as binary classifiers on paths. In the following chapter we will address algorithms for both checking and deriving quality-optimal causes.

4 Algorithmic Considerations

As stated in the introduction, this thesis is focused on building a fundamental framework for probability-raising causality in MDPs in order to enable further applications. Therefore, the thesis has a strong emphasis on algorithmic studies of cause-effect relations in MDPs using the probability-raising principle. In previous chapters we laid the groundwork for this analysis by giving a fixed setting for the MDPs in question as well as by defining and discussing GPR and SPR causes (cf. Definition 14) extensively in Section 3.1 and 3.2. Generally speaking, a given probabilistic cause-effect relation for events $\mathbf{C}$ and $\mathbf{E}$ poses two natural questions:

1. Given two events $\mathbf{C}$ and $\mathbf{E}$, does $\mathbf{C}$ cause $\mathbf{E}$ to happen?
2. Given an event $\mathbf{E}$, which events $\mathbf{C}$ cause $\mathbf{E}$ to happen?

Considering probability-raising causes in MDPs as in Definition 14, these questions translate to the following algorithmic queries:

1. Given an MDP $\mathfrak{M}$, a set of terminal states Eff and a set of states Cause, check whether Cause is a GPR (resp. SPR) cause for Eff in $\mathfrak{M}$.
2. Given an MDP $\mathfrak{M}$ and a set of terminal states Eff, compute a set of states Cause which constitutes a GPR (resp. SPR) cause for Eff in $\mathfrak{M}$.

For the second query we will see, that this is equivalent with checking whether there exists a singleton set $\{\mathbf{c}\}$ which is an SPR cause for Eff in $\mathfrak{M}$. Therefore, the quality measures defined in Chapter 3.3 can be used to compute more sophisticated causes than singletons by searching and computing quality-optimal causes for a given quality measure. These problems will be studied for GPR and SPR causes and the two distinct quality measures of recall and fscore, which comes down to 6 different algorithmic problems. We summarize these problems and our results for general MDPs in Table 4.1. Since the coverage ratio can be expressed in terms of the recall and covrat-optimal causes are also recall-optimal (and vice versa) we omit its corresponding algorithmic problems.

State-based PR-causality	Check PR condition Cause and Eff are given	Compute PR cause Cause for given effect Eff	
		recall-optimal	fscore-optimal
GPR cause	$\in \mathrm{coNP}$ (Thm.53)	poly-space(Rem.76) NP-hard (Thm.77)	poly-space(Rem.76) NP-hard (Thm.77)
SPR cause	$\in \mathrm{P}$ (Thm.44)	poly-time(Thm.70)	poly-space(Rem.72) $\in \mathrm{NP} \cap \mathrm{coNP}$ (Thm.75)

Table 4.1: Complexity results for MDPs with fixed effect set Eff

PR causality in DTMCs	Check PR condition Cause and Eff are given	Compute PR cause Cause for given effect Eff	
		recall-optimal	fscore-optimal
GPR cause	$\in \mathrm{P}$ (Rem.30)	NP-complete (Thm.77)	NP-complete (Thm.77)
SPR cause	$\in \mathrm{P}$ (Rem.30)	poly-time(Thm.70)	poly-time(Thm.73)

Table 4.2: Complexity results for DTMCs with fixed effect set Eff

Note that for hardness results regarding quality-optimal causes we consider the associated threshold problem. So, the problem is whether there exists a PR cause Cause of the corresponding variant such that $\mathit{quality}(\mathit{Cause}) \geqslant \vartheta$ for some $\vartheta \in \mathbb{Q}$. In the case where we deal with a DTMC we improve most results as shown in Table 4.2.

In this chapter prove and visualize these results. Since GPR and SPR causes differ, so do their complexity results. Thus, we also discuss the implications of these differences on the relation between both variants of probability-raising causality.

4.1 Checking the Probability-Raising Conditions

Throughout this section we are given an MDP $\mathcal{M} = (S, \mathit{Act}, \mathrm{P}, \mathit{init})$, a non-empty set of terminal states $\mathit{Eff} \subset S$ and a non-empty set $\mathit{Cause} \subseteq S \setminus \mathit{Eff}$. Here, we give algorithms for checking whether Cause is a GPR or SPR cause for Eff in $\mathcal{M}$ (cf. Def. 14). Both variants of PR cause include a minimality condition for the set Cause stating that for all $\mathbf{c} \in \mathit{Cause}$ we have $\Pr^{\max}_{\mathcal{M}}((\neg \mathit{Cause}) \,\mathrm{U}\, \mathbf{c}) > 0$. However, this condition only depends on the graph structure of $\mathcal{M}$ and for a state $\mathbf{c} \in \mathit{Cause}$ the condition is verifiable by standard model checking techniques in polynomial time. So, we focus on checking the probability-raising conditions (G) and (S) .

Remark 30 (Checking GPR and SPR in DTMCs). Suppose $\mathcal{M}$ is a DTMC, i.e. does not contain non-deterministic choices. Then both conditions **(G)** and **(S)** can be checked by computing the corresponding (conditional) probabilities, which can be done in polynomial time [BK08; Bai+14; Mär20]. ◁

By Remark 30, the interesting case is checking the probability-raising conditions in general MDPs. The decision problem for SPR causality is closely related to the existence of PR causes and decidable in P (Section 4.1.2). However, deciding **(G)** for GPR causality is more complex and polynomially reducible to the non-solvability of a quadratic program (Section 4.1.3). Since quadratic programming is proven to be in NP [Vav90], we can use similar techniques to show that deciding **(G)** is in coNP.

4.1.1 Canonical MDP for given Cause Set

Before we show and discuss the algorithms for checking GPR and SPR causality, we start with a model transformation that results in an end-component free MDP in which both Cause and Eff can only be visited once. Furthermore, the transformation will ensure that there are exactly *four* terminal states which correspond to the entries of the confusion matrix (Figure 3.14). This will allow us to express probabilities as (sums of) reachability probabilities of single terminal states. We use the following abbreviations as already introduced in Chapter 3.3:

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Cause} \wedge \Diamond \mathit{Eff}) = tp^{\mathfrak{S}}_{\mathcal{M}}, \qquad \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Cause} \wedge \neg\Diamond \mathit{Eff}) = fp^{\mathfrak{S}}_{\mathcal{M}},$$
$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\neg\Diamond \mathit{Cause} \wedge \Diamond \mathit{Eff}) = fn^{\mathfrak{S}}_{\mathcal{M}}, \qquad \Pr^{\mathfrak{S}}_{\mathcal{M}}(\neg\Diamond \mathit{Cause} \wedge \neg\Diamond \mathit{Eff}) = tn^{\mathfrak{S}}_{\mathcal{M}}.$$

We omit the subscript $\mathcal{M}$ if the MDP is clear from the context. Recall that we abbreviate end-components of $\mathcal{M}$ by ECs and maximal end-components by MECs.

Definition 31 (Canonical MDP). Given an MDP $\mathcal{M} = (S, \mathit{Act}, \mathrm{P}, \mathit{init})$, a non-empty set of terminal states $\mathit{Eff} \subset S$ and a non-empty set $\mathit{Cause} \subseteq S \setminus \mathit{Eff}$, we transform the original $\mathcal{M}$ to the *canonical MDP* $\mathcal{M}_{\mathit{Cause}} = (S', \mathit{Act} \cup \{\alpha_{\min}, \alpha_{\max}, \tau\}, \mathrm{P}', \mathit{init})$ for Cause in the following way:

(i) All outgoing transitions from states $\mathfrak{c} \in \mathit{Cause}$ are deleted and instead two fresh actions $\alpha_{\min}$ and $\alpha_{\max}$ are added. These actions then have the following transitions

$$\mathrm{P}'(\mathfrak{c}, \alpha_{\min}, \mathit{TP}) = \mathfrak{p}_{\mathfrak{c},\min} = \Pr^{\min}_{\mathcal{M},\mathfrak{c}}(\Diamond \mathit{Eff}), \qquad \mathrm{P}'(\mathfrak{c}, \alpha_{\min}, \mathit{FP}) = 1 - \mathfrak{p}_{\mathfrak{c},\min},$$
$$\mathrm{P}'(\mathfrak{c}, \alpha_{\max}, \mathit{TP}) = \mathfrak{p}_{\mathfrak{c},\max} = \Pr^{\max}_{\mathcal{M},\mathfrak{c}}(\Diamond \mathit{Eff}), \qquad \mathrm{P}'(\mathfrak{c}, \alpha_{\max}, \mathit{FP}) = 1 - \mathfrak{p}_{\mathfrak{c},\max},$$

where TP and FP are fresh terminal states. Denote the resulting MDP $\mathcal{M}_{<\mathit{Cause}>}$.

(ii) Collapse the maximal end-components (MECs) $\mathcal{E}$ of $\mathcal{M}_{<\mathit{Cause}>}$ by taking the MEC-quotient (cf. Def. 8) without the additional terminal state $\perp$ and the corresponding transitions. This results in states $s_{\mathcal{E}}$ representing the MECs. Denote the resulting MDP $\mathcal{M}_{(\mathrm{ii})}$.

(iii) The set $\mathit{Eff} \subset \mathrm{S}$ in $\mathcal{M}_{(\mathrm{ii})}$ is collapsed to the fresh state FN and all other terminal states in $\mathcal{M}_{(\mathrm{ii})}$ to the fresh state TN. States $\mathrm{s}_{\mathcal{E}}$ representing MECs in $\mathcal{M}_{(\mathrm{ii})}$ get the additional transition $\mathrm{P}(\mathrm{s}_{\mathcal{E}}, \tau, \mathit{TN}) = 1$. The resulting MDP is $\mathcal{M}_{\mathit{Cause}}$.

Therefore, $\mathrm{S}' = \mathrm{S} \cup \{\mathit{TP}, \mathit{FP}, \mathit{FN}, \mathit{TN}\} \cup \{\mathrm{s}_{\mathcal{E}} \mid \mathcal{E} \text{ is EC of } \mathcal{M}_{<\mathit{Cause}>}\}$. In $\mathcal{M}_{\mathit{Cause}}$ we consider a new effect set $\mathit{Eff}' = \{\mathit{TP}, \mathit{FN}\}$. If $\mathit{Cause} = \{\mathrm{c}\}$ is a singleton we write $\mathcal{M}_{\mathrm{c}}$. ◁

Intuitively, the first step (i) of the transformation ensures, that each state $\mathrm{c} \in \mathit{Cause}$ is visited at most once, while still preserving all possible values for $\mathit{tp}^{\mathfrak{S}}_{\mathcal{M}}$ and $\mathit{fp}^{\mathfrak{S}}_{\mathcal{M}}$ by using randomization in $\mathcal{M}_{\mathit{Cause}}$. In the second (ii) and third step (iii) the transformation gets rid of end-components by collapsing them into single states. Instead, a fresh action τ is introduced which corresponds to the case that in the original MDP $\mathcal{M}$ a scheduler $\mathfrak{S}$ realizes a probability $\mathit{tn}^{\mathfrak{S}}_{\mathcal{M}}$ by staying in an end-component $\mathcal{E}$. By choosing the action τ in the state $\mathrm{s}_{\mathcal{E}}$ with a corresponding probability, a scheduler $\mathfrak{U}$ of $\mathcal{M}_{\mathit{Cause}}$ can realize such probabilities as well.

As a visualization of the transformation consider the abstract MDP $\mathcal{M}$ in Figure 4.1. Here, we only highlight some of the states in the MDP and omit transition probabilities. Note, that t_0 and t_1 are terminal states. For $\mathit{Cause} = \{\mathrm{c}\}$ and $\mathit{Eff} = \{\mathrm{e}_0, \mathrm{e}_1\}$ we consider $\mathcal{M}_{<\mathrm{c}>}$ shown in Figure 4.2 for the first transformation from step (i). All outgoing transitions from c are deleted and instead two actions $\alpha_{\min}$ and $\alpha_{\max}$ are added. The outgoing transitions colored in red (■) correspond to $\alpha_{\max}$ with $\mathrm{P}'(\mathrm{c}, \alpha_{\max}, \mathit{TP}) = \mathrm{p}_{\mathrm{c},\max}$ and $\mathrm{P}'(\mathrm{c}, \alpha_{\max}, \mathit{FP}) = 1 - \mathrm{p}_{\mathrm{c},\max}$. The transitions colored in blue (■) correspond to $\alpha_{\min}$ with $\mathrm{P}'(\mathrm{c}, \alpha_{\min}, \mathit{TP}) = \mathrm{p}_{\mathrm{c},\min}$ and $\mathrm{P}'(\mathrm{c}, \alpha_{\min}, \mathit{FP}) = 1 - \mathrm{p}_{\mathrm{c},\min}$. The further transformation to $\mathcal{M}_{\mathrm{c}}$ (step (ii) and (iii)) can be seen in Figure 4.3. Here, step (ii) collapses the MEC $\mathcal{E}$ of $\mathcal{M}_{<\mathrm{c}>}$ to state $\mathrm{s}_{\mathcal{E}}$ while keeping the outgoing action α. Furthermore, in step (iii) the effect states e_0 and e_1 are collapsed into FN and all other terminal states to TN. Additionally, the collapsed MEC gets a fresh action τ leading to TN with probability 1.

The soundness of the canonical MDP for a given set Cause with respect to (S) and (G) is shown in the following Lemma 32 and Corollary 33.

Lemma 32 ($\mathcal{M}_{\mathit{Cause}}$ preserves Confusion Matrix). *Given an MDP $\mathcal{M}$, a set of terminal states $\mathit{Eff} \subset \mathrm{S}$ and a set $\mathit{Cause} \subseteq \mathrm{S} \setminus \mathit{Eff}$, for each scheduler $\mathfrak{S}$ of $\mathcal{M}$ there is a scheduler $\mathfrak{U}$ of $\mathcal{M}_{\mathit{Cause}}$ and vice versa such that*

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

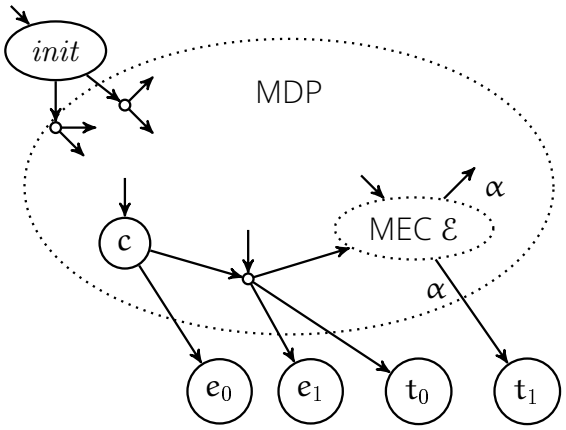

Figure 4.1: Example MDP $\mathcal{M}$

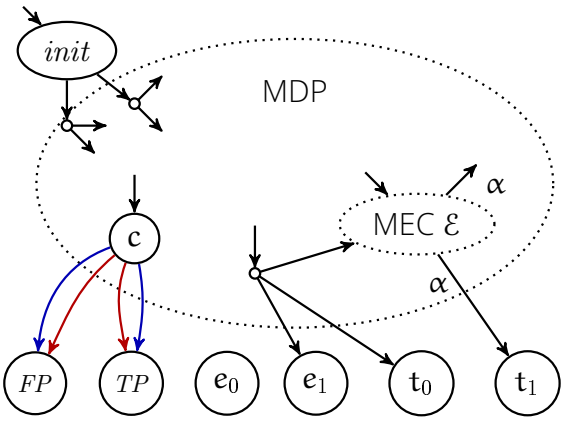

Figure 4.2: Intermediate MDP $\mathcal{M}_{<\mathbf{c}>}$

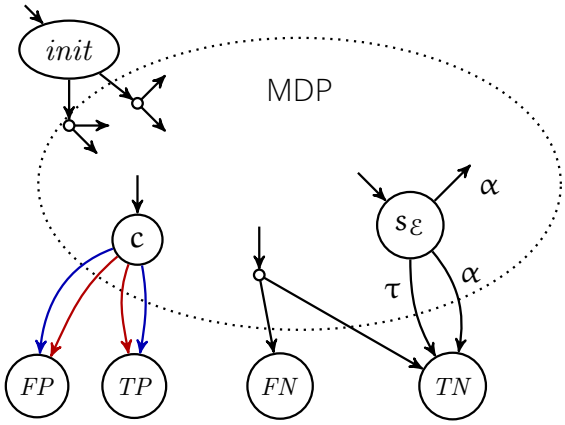

Figure 4.3: Transformed MDP $\mathcal{M}_{\mathbf{c}}$

Proof. We will proof the lemma by showing two claims:

Claim (i). *For each scheduler $\mathfrak{S}$ of $\mathcal{M}$ there is a scheduler $\mathfrak{T}$ of the MDP $\mathcal{M}_{<Cause>}$ [(i) of Def. 31] and vice versa such that*

$$tp^{\mathfrak{S}}_{\mathcal{M}} = \Pr^{\mathfrak{T}}_{\mathcal{M}_{<Cause>}}(\Diamond TP) = tp^{\mathfrak{T}}_{\mathcal{M}_{<Cause>}}, \qquad fp^{\mathfrak{S}}_{\mathcal{M}} = \Pr^{\mathfrak{T}}_{\mathcal{M}_{<Cause>}}(\Diamond FP) = fp^{\mathfrak{T}}_{\mathcal{M}_{<Cause>}},$$
$$fn^{\mathfrak{S}}_{\mathcal{M}} = fn^{\mathfrak{T}}_{\mathcal{M}_{<Cause>}}, \qquad tn^{\mathfrak{S}}_{\mathcal{M}} = tn^{\mathfrak{T}}_{\mathcal{M}_{<Cause>}},$$

" $\Longrightarrow$ ": Given a scheduler $\mathfrak{S}$ of $\mathcal{M}$ we define a corresponding scheduler $\mathfrak{T}$ of $\mathcal{M}_{<Cause>}$ in the following way: For all paths $\pi \in Paths_{fin}(\mathcal{M}_{<Cause>})$ ending in states $s \in S \setminus Cause$ we define for all actions $\alpha \in Act(s)$: $\mathfrak{T}(\pi,\alpha) = \mathfrak{S}(\pi,\alpha)$. When a path $\pi \in Paths_{fin}(\mathcal{M}_{<Cause>})$ reaches a state $\mathbf{c} \in Cause$ we let

$$\mathfrak{T}(\pi,\alpha_{\max}) = \frac{\Pr^{res(\mathfrak{S},\pi)}_{\mathcal{M},\mathbf{c}}(\Diamond Eff) - p_{\mathbf{c},\min}}{p_{\mathbf{c},\max} - p_{\mathbf{c},\min}} \quad \text{and} \quad \mathfrak{T}(\pi,\alpha_{\min}) = 1 - \mathfrak{T}(\pi,\alpha_{\max}),$$

if $p_{\mathbf{c},\max} > p_{\mathbf{c},\min}$. Otherwise, if $p_{\mathbf{c},\max} = p_{\mathbf{c},\min}$ then the actions $\alpha_{\max}$ and $\alpha_{\min}$ give equivalent distributions for the transition function and we let $\mathfrak{U}(\pi,\mathbf{c}_0) = \alpha_{\min}$.

In order to now show that $tp^{\mathfrak{S}}_{\mathcal{M}} = \Pr^{\mathfrak{T}}_{\mathcal{M}_{<Cause>}}(\Diamond TP)$ we need to show that for any path $\pi \in Paths_{fin}(\mathcal{M})$ that does not visit $Cause$ or only visits $Cause$ once with $last(\pi) \in Cause$ we have

$$\mathfrak{x} := \Pr^{res(\mathfrak{S},\pi)}_{\mathcal{M}}(\Diamond Cause \wedge \Diamond Eff) = \Pr^{res(\mathfrak{T},\pi)}_{\mathcal{M}_{<Cause>}}(\Diamond TP) =: \mathfrak{y},$$

where π is considered as a path in both MDPs $\mathcal{M}$ and $\mathcal{M}_{<Cause>}$ by abuse of notation. For paths not visiting a state $\mathfrak{c} \in Cause$ this is clear by construction. If π ends in a state $\mathfrak{c} \in Cause$ we have the following

$$\begin{aligned}\mathfrak{y} &= \frac{\mathfrak{x} - \mathfrak{p}_{\mathfrak{c},\min}}{\mathfrak{p}_{\mathfrak{c},\max} - \mathfrak{p}_{\mathfrak{c},\min}} \cdot \mathfrak{p}_{\mathfrak{c},\max} + \left(1 - \frac{\mathfrak{x} - \mathfrak{p}_{\mathfrak{c},\min}}{\mathfrak{p}_{\mathfrak{c},\max} - \mathfrak{p}_{\mathfrak{c},\min}}\right) \cdot \mathfrak{p}_{\mathfrak{c},\min} \\ &= \mathfrak{p}_{\mathfrak{c},\min} + \frac{\mathfrak{p}_{\mathfrak{c},\max} \cdot \mathfrak{x} - \mathfrak{p}_{\mathfrak{c},\max} \cdot \mathfrak{p}_{\mathfrak{c},\min} - \mathfrak{p}_{\mathfrak{c},\min} \cdot \mathfrak{x} + (\mathfrak{p}_{\mathfrak{c},\min})^2}{\mathfrak{p}_{\mathfrak{c},\max} - \mathfrak{p}_{\mathfrak{c},\min}} \\ &= \mathfrak{p}_{\mathfrak{c},\min} + \frac{(\mathfrak{p}_{\mathfrak{c},\max} - \mathfrak{p}_{\mathfrak{c},\min})(\mathfrak{x} - \mathfrak{p}_{\mathfrak{c},\min})}{\mathfrak{p}_{\mathfrak{c},\max} - \mathfrak{p}_{\mathfrak{c},\min}} = \mathfrak{x}.\end{aligned}$$

A analogue calculation shows $fp^{\mathfrak{S}}_{\mathcal{M}} = \Pr^{\mathfrak{T}}_{\mathcal{M}_{<Cause>}}(\Diamond FP)$. The remaining equalities follow from the fact that no path satisfying $(\neg Cause)$ in the original MDP $\mathcal{M}$ is influenced by the construction of $\mathcal{M}_{<Cause>}$.

" $\Longleftarrow$ ": Let $\mathfrak{T}$ be a scheduler of $\mathcal{M}_{<Cause>}$. We define a corresponding scheduler $\mathfrak{S}$ of $\mathcal{M}$ by mimicking $\mathfrak{T}$ before a state $\mathfrak{c} \in Cause$ was visited. Whenever a path visits a state $\mathfrak{c} \in Cause$ the scheduler $\mathfrak{T}$ chooses $\alpha_{\max}$ with probability $\mathfrak{p} \in [0,1]$ and $\alpha_{\min}$ with probability $(1-\mathfrak{p})$. The corresponding scheduler $\mathfrak{S}$ then chooses a maximizing action $\alpha \in Act(\mathfrak{c})$ for $\Diamond Eff$ with probability $\mathfrak{p}$ and a minimizing action $\beta \in Act(\mathfrak{c})$ for $\Diamond Eff$ with probability $1-\mathfrak{p}$. Afterwards, $\mathfrak{S}$ realizes a maximizing MD scheduler for $\Diamond Eff$ if a maximizing action α was taken and otherwise it realizes a minimizing MD scheduler for $\Diamond Eff$. This transformation adds two additional memory cells to $\mathfrak{S}$. One memory cell keeps track of whether $Cause$ was visited or not and one memory cell keeps track whether a maximizing or minimizing action was taken in the first visited $\mathfrak{c} \in Cause$. Furthermore, $\mathfrak{S}$ uses randomization. The claimed equalities then follow by construction of $\mathfrak{S}$.

Claim (ii + iii). *For each scheduler $\mathfrak{T}$ of $\mathcal{M}_{<Cause>}$ [(i) of Def. 31] there is a scheduler $\mathfrak{U}$ of $\mathcal{M}_{Cause}$ and vice versa such that*

$$\begin{aligned} tp^{\mathfrak{T}}_{\mathcal{M}_{<Cause>}} &= \Pr^{\mathfrak{U}}_{\mathcal{M}_{Cause}}(\Diamond TP), & fp^{\mathfrak{T}}_{\mathcal{M}_{<Cause>}} &= \Pr^{\mathfrak{U}}_{\mathcal{M}_{Cause}}(\Diamond FP), \\ fn^{\mathfrak{T}}_{\mathcal{M}_{<Cause>}} &= \Pr^{\mathfrak{U}}_{\mathcal{M}_{Cause}}(\Diamond FN), & tn^{\mathfrak{T}}_{\mathcal{M}_{<Cause>}} &= \Pr^{\mathfrak{U}}_{\mathcal{M}_{Cause}}(\Diamond TN). \end{aligned}$$

" $\Longrightarrow$ ": Given a scheduler $\mathfrak{T}$ of $\mathcal{M}_{<Cause>}$ we define $\mathfrak{U}$ in the following way: For paths which do not visit a state $\mathfrak{s}_{\mathcal{E}}$ representing an MEC $\mathfrak{U}$ mimics $\mathfrak{T}$. Whenever a path enters an MEC $\mathcal{E}$ of $\mathcal{M}_{<Cause>}$ the corresponding path π in $\mathcal{M}_{Cause}$ ends in the state $\mathfrak{s}_{\mathcal{E}}$. The scheduler $\mathfrak{T}$ stays in the MEC with probability $\mathfrak{p} \in [0,1]$ and we define $\mathfrak{U}(\pi,\tau) = \mathfrak{p}$, where τ is the action leading

to TN with probability 1. With the remaining probability $1-\mathrm{p}$ scheduler $\mathfrak{T}$ chooses external actions with some distribution. According to the same distribution weighted with $1-\mathrm{p}$ we let $\mathfrak{U}$ choose the same external actions.

Now, the first two equalities $tp^{\mathfrak{T}}_{\mathcal{M}_{<Cause>}} = \Pr^{\mathfrak{U}}_{\mathcal{M}_{Cause}}(\Diamond TP)$ and $fp^{\mathfrak{T}}_{\mathcal{M}_{<Cause>}} = \Pr^{\mathfrak{U}}_{\mathcal{M}_{Cause}}(\Diamond FP)$ follow from Lemma 9 as both values are reachability probabilities of terminal states. The same argumentation goes for the equality $fn^{\mathfrak{T}}_{\mathcal{M}_{<Cause>}} = \Pr^{\mathfrak{U}}_{\mathcal{M}_{Cause}}(\Diamond FN)$ as here only terminal states of $\mathcal{M}_{<Cause>}$ are collapsed into a fresh terminal state FN. Therefore, the same result for MECs Lemma 9 can be used. For the last equality $tn^{\mathfrak{T}}_{\mathcal{M}_{<Cause>}} = \Pr^{\mathfrak{U}}_{\mathcal{M}_{Cause}}(\Diamond TN)$ we note that this directly corresponds to either reaching any terminal state $\mathrm{t} \notin Eff$ in $\mathcal{M}_{<Cause>}$ or staying in an end-component indefinitely. The probability for the first case is redistributed to TN by collapsing exactly those states. For the second case the constructed scheduler $\mathfrak{U}$ chooses τ in a state $\mathrm{s}_{\mathcal{E}}$ representing the end-component with compatible probability. As the MEC quotient does not influence the probabilities in other ways [Alf97; Alf99] the claim follows.

" $\Longleftarrow$ ": Given a scheduler $\mathfrak{U}$ of $\mathcal{M}_{Cause}$ a corresponding scheduler $\mathfrak{T}$ for $\mathcal{M}_{<Cause>}$ can be defined using the correspondence from Lemma 9 which directly implies the claimed equalities.

The Lemma now follows from the combination of both claims. □

Corollary 33 ($\mathcal{M}_{Cause}$ preserves Probability-Raising). *Given an MDP $\mathcal{M}$, a set of terminal states $Eff \subset \mathrm{S}$ and a set $Cause \subseteq \mathrm{S} \setminus Eff$, a scheduler $\mathfrak{S}$ of $\mathcal{M}$ is satisfies* (GPR) *(resp.* (SPR) *) for $Cause$ and Eff in $\mathcal{M}$ if and only if a corresponding scheduler $\mathfrak{U}$ from Lemma 32 satisfies* (GPR) *(resp.* (SPR) *) for $Cause$ and $\{TP, FN\}$ in $\mathcal{M}_{Cause}$.*

Proof. For (GPR) this follows directly from Lemma 32 as we can express the (GPR) condition in terms of the confusion matrix, c.f. (GPR*). Thus, we consider the case where $\mathfrak{S}$ is satisfying (SPR) for $Cause$ and Eff in $\mathcal{M}$. Its corresponding scheduler $\mathfrak{U}$ of $\mathcal{M}_{Cause}$ is defined as described in the proof of Lemma 32. Since $\mathfrak{S}$ is satisfying SPR in $\mathcal{M}$, we have $\Pr^{\mathfrak{U}}_{\mathcal{M}_{Cause}}(\Diamond Cause) > 0$ trivially. In $\mathcal{M}_{Cause}$ any state $\mathrm{c} \in Cause$ can only be visited once. Thus, all paths visiting such a state c satisfy $((\neg Cause)\,\mathrm{U}\,\mathrm{c})$. Therefore we have

$$\Pr^{\mathfrak{U}}_{\mathcal{M}_{Cause}}(\Diamond Eff \mid \Diamond \mathrm{c}) = \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond Eff \mid (\neg Cause)\,\mathrm{U}\,\mathrm{c}) > \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond Eff) = \Pr^{\mathfrak{U}}_{\mathcal{M}_{Cause}}(\Diamond\{\mathsf{TP}, \mathsf{FN}\}),$$

where the second equality follows from Lemma 32.

For the other direction, the reversal of the above argumentation yields the claim. □

In a straight forward manner, Corollary 33 implies, that every GPR (resp. SPR) cause $Cause$ for Eff in $\mathcal{M}$ is a GPR (resp. SPR) cause for $\{TP, FN\}$ in $\mathcal{M}_{Cause}$. To see this, note that the transformation of scheduler from $\mathcal{M}_{Cause}$ to $\mathcal{M}$ can be done for any scheduler $\mathfrak{U}$ of $\mathcal{M}_{Cause}$.

Remark 34 (Assumptions to $\mathcal{M}$ when $Cause$ is given). Let us collect the assumptions that hold for $\mathcal{M} = \mathcal{M}_{Cause}$:

(A1) $\mathcal{M}$ does not contain end-components.

(A2) There are exactly four terminal states TP, FN, FP, TN with

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond Cause \wedge \Diamond Eff) = \Pr^{\mathfrak{U}}_{\mathcal{M}}(\Diamond TP), \qquad \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond Cause \wedge \neg\Diamond Eff) = \Pr^{\mathfrak{U}}_{\mathcal{M}}(\Diamond FP),$$
$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\neg\Diamond Cause \wedge \Diamond Eff) = \Pr^{\mathfrak{U}}_{\mathcal{M}}(\Diamond FN), \quad \Pr^{\mathfrak{S}}_{\mathcal{M}}(\neg\Diamond Cause \wedge \neg\Diamond Eff) = \Pr^{\mathfrak{U}}_{\mathcal{M}}(\Diamond TN).$$

(A3) Each state $\mathbf{c} \in Cause$ has exactly two enabled actions $\alpha_{\min}$ and $\alpha_{\max}$ with

$$\mathrm{P}(\mathbf{c}, \alpha_{\min}, TP) = \mathfrak{p}_{\min,\mathbf{c}}, \qquad \mathrm{P}(\mathbf{c}, \alpha_{\min}, FP) = 1 - \mathfrak{p}_{\min,\mathbf{c}},$$
$$\mathrm{P}(\mathbf{c}, \alpha_{\max}, TP) = \mathfrak{p}_{\max,\mathbf{c}}, \qquad \mathrm{P}(\mathbf{c}, \alpha_{\max}, FP) = 1 - \mathfrak{p}_{\max,\mathbf{c}},$$

where $\mathfrak{p}_{\min,\mathbf{c}} \leqslant \mathfrak{p}_{\max,\mathbf{c}}$.

When checking a cause candidate set $Cause \subseteq S \setminus Eff$ for GPR or SPR causality, we show that it suffices to only activate the minimizing actions $\alpha_{\min}$ in $\mathcal{M}_{Cause}$. Intuitively, this comes from the universal quantification over scheduler in **(G)** and **(S)** . This quantification means that in order to find a scheduler for which **(G)** or **(S)** does not hold, one can only consider the "worst-case" scheduler which minimize the probability of reaching effect after visiting the cause. As we use this result often, we introduce a notation for the corresponding MDP.

Definition 35 (MDP with Minimal Effect Probabilities from Cause Candidates). Given a set of states $Cause \subseteq S$ we write $\mathcal{M}^{\min}_{Cause}$ for the MDP resulting from the canonical causal MDP $\mathcal{M}_{Cause}$ by removing the action $\alpha_{\max}$ in all states in $Cause$. We call $\mathcal{M}^{\min}_{Cause}$ the *minimizing MDP* for $Cause$. In case $Cause = \{\mathbf{c}\}$ is a singleton we write $\mathcal{M}^{\min}_{\mathbf{c}}$. ◁

Lemma 36 (Minimal Effect Probabilities preserve PR Causes). *Let $\mathcal{M} = (S, Act, \mathrm{P}, init)$ be an MDP and $Eff \subseteq S$ a non-empty set of terminal states. Then $Cause \subseteq S \setminus Eff$ is a GPR cause (resp. SPR cause) for Eff in $\mathcal{M}$ if and only if $Cause$ is a GPR cause (resp. SPR cause) for $\{TP, FN\}$ in $\mathcal{M}^{\min}_{Cause}$.*

We will prove Lemma 36 by first providing the implications from $\mathcal{M}^{\min}_{Cause}$ to $\mathcal{M}$ separately.. This will be done for SPR causes in Lemma 37 and for GPR causes in Lemma 38. Note that we only have to consider the behavior of scheduler after visiting $Cause$ since the soundness with respect to the MEC-quotient was proven by Corollary 33.

Lemma 37 (Criterion for SPR Causes). *Suppose $Cause$ is an SPR cause for $\{TP, FN\}$ in $\mathcal{M}^{\min}_{Cause}$. Then, $Cause$ is also an SPR cause for Eff in $\mathcal{M}$.*

Proof. By Lemma 32 and Corollary 33 we assume $\mathcal{M} = \mathcal{M}_{Cause}$ throughout the proof. Naturally, a scheduler $\mathfrak{S}$ of $\mathcal{M}$ corresponds to scheduler $\mathfrak{U}$ of $\mathcal{M}^{\min}_{Cause}$, where $\mathfrak{U}$ behaves identically

to $\mathfrak{S}$ except for $\mathbf{c} \in \mathit{Cause}$ in which $\mathfrak{U}$ can only choose $\alpha_{\min}$. Vice versa, a scheduler $\mathfrak{U}$ of $\mathcal{M}^{\min}_{\mathit{Cause}}$ can directly be interpreted as a scheduler for $\mathcal{M}_{\mathit{Cause}}$. We assume that Cause is an SPR cause for $\{TP, FN\}$ in $\mathcal{M}^{\min}_{\mathit{Cause}}$ and need to prove that (S) holds for Cause in $\mathcal{M}$. Thus, given a state $\mathbf{c} \in \mathit{Cause}$ and a scheduler $\mathfrak{S}$ of $\mathcal{M}$ with $\Pr^{\mathfrak{S}}_{\mathcal{M}}((\neg \mathit{Cause}) \,\mathrm{U}\, \mathbf{c}) > 0$ need to show that

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff} \mid (\neg \mathit{Cause}) \,\mathrm{U}\, \mathbf{c}) > \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff})$$

Recall, that states in Eff are terminal. We use the abbreviations $\psi_{\mathbf{c}} = (\neg \mathit{Cause}) \,\mathrm{U}\, \mathbf{c}$ and $\mathfrak{p}_{\min,\mathbf{c}} = \Pr^{\min}_{\mathcal{M},\mathbf{c}}(\Diamond \mathit{Eff})$ and let $\mathfrak{U}_{\mathbf{c}}$ denote the set of all schedulers $\mathfrak{U}$ of $\mathcal{M}$ such that

- $\Pr^{\mathfrak{U}}_{\mathcal{M}}(\psi_{\mathbf{c}}) > 0$ and
- $\Pr^{\mathit{res}(\mathfrak{U},\pi)}_{\mathcal{M},\mathbf{c}}(\Diamond \mathit{Eff}) = \mathfrak{p}_{\min,\mathbf{c}}$ for each finite $\mathfrak{U}$-path π from init to $\mathbf{c}$.

Clearly, for $\mathfrak{U} \in \mathfrak{U}_{\mathbf{c}}$ we have $\Pr^{\mathfrak{U}}_{\mathcal{M}}(\Diamond \mathbf{c} \wedge \Diamond \mathit{Eff}) = \Pr^{\mathfrak{U}}_{\mathcal{M}}(\Diamond \mathbf{c}) \cdot \mathfrak{p}_{\min,\mathbf{c}}$. By assumption Cause is an SPR cause for $\{TP, FN\}$ in $\mathcal{M}^{\min}_{\mathit{Cause}}$ and thus we have:

$$\mathfrak{p}_{\min,\mathbf{c}} > \Pr^{\mathfrak{U}}_{\mathcal{M}^{\min}_{\mathit{Cause}}}(\Diamond \mathit{Eff}) \quad \text{for all schedulers } \mathfrak{U} \in \mathfrak{U}_{\mathbf{c}}. \tag{4.1}$$

Suppose $\mathfrak{S}$ is a scheduler for $\mathcal{M}$ with $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\psi_{\mathbf{c}}) > 0$. Then

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\psi_{\mathbf{c}} \wedge \Diamond \mathit{Eff}) \geqslant \Pr^{\mathfrak{S}}_{\mathcal{M}}(\psi_{\mathbf{c}}) \cdot \mathfrak{p}_{\min,\mathbf{c}}.$$

We now turn to condition (SPR) for $\mathbf{c}$ and $\mathfrak{S}$. Note that there exists a scheduler $\mathfrak{U} = \mathfrak{U}_{\mathfrak{S}} \in \mathfrak{U}_{\mathbf{c}}$ dependent on $\mathfrak{S}$ with

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\psi_{\mathbf{c}}) = \Pr^{\mathfrak{U}}_{\mathcal{M}}(\psi_{\mathbf{c}}) \quad \text{and} \quad \Pr^{\mathfrak{S}}_{\mathcal{M}}((\neg\psi_{\mathbf{c}}) \wedge \Diamond \mathit{Eff}) = \Pr^{\mathfrak{U}}_{\mathcal{M}}((\neg\psi_{\mathbf{c}}) \wedge \Diamond \mathit{Eff}).$$

To see this, consider the scheduler $\mathfrak{U}$ which behaves identical to $\mathfrak{S}$ as long as $\mathbf{c}$ is not reached. As soon as $\mathfrak{U}$ has reached $\mathbf{c}$, scheduler $\mathfrak{U}$ switches mode and behaves as an MD scheduler minimizing the probability to reach the set Eff. The SPR condition (S) holds for $\mathbf{c}$ and $\mathfrak{S}$ if and only if

$$\frac{\Pr^{\mathfrak{S}}_{\mathcal{M}}(\psi_{\mathbf{c}} \wedge \Diamond \mathit{Eff})}{\Pr^{\mathfrak{S}}_{\mathcal{M}}(\psi_{\mathbf{c}})} \;>\; \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff}) \tag{4.2}$$

by definition of conditional probabilities under a scheduler. Using

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff}) \;=\; \Pr^{\mathfrak{S}}_{\mathcal{M}}(\psi_{\mathbf{c}} \wedge \Diamond \mathit{Eff}) + \Pr^{\mathfrak{S}}_{\mathcal{M}}((\neg\psi_{\mathbf{c}}) \wedge \Diamond \mathit{Eff}),$$

we can equivalently convert condition (4.2) for $\mathbf{c}$ and $\mathfrak{S}$ to

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\psi_{\mathbf{c}} \wedge \Diamond \mathit{Eff}) \cdot \frac{1 - \Pr^{\mathfrak{S}}_{\mathcal{M}}(\psi_{\mathbf{c}})}{\Pr^{\mathfrak{S}}_{\mathcal{M}}(\psi_{\mathbf{c}})} \;>\; \Pr^{\mathfrak{S}}_{\mathcal{M}}((\neg\psi_{\mathbf{c}}) \wedge \Diamond \mathit{Eff}). \tag{4.3}$$

The remaining task is to derive (4.3) from (4.1). Applying (4.1) to scheduler $\mathfrak{U} = \mathfrak{U}_{\mathfrak{S}}$ as defined above yields:

$$\mathfrak{p}_{\min,\mathbf{c}} > \Pr^{\mathfrak{U}}_{\mathcal{M}}(\psi_{\mathbf{c}} \wedge \Diamond \mathit{Eff}) + \Pr^{\mathfrak{U}}_{\mathcal{M}}((\neg\psi_{\mathbf{c}}) \wedge \Diamond \mathit{Eff}) = \Pr^{\mathfrak{S}}_{\mathcal{M}}(\psi_{\mathbf{c}}) \cdot \mathfrak{p}_{\min,\mathbf{c}} + \Pr^{\mathfrak{S}}_{\mathcal{M}}((\neg\psi_{\mathbf{c}}) \wedge \Diamond \mathit{Eff}).$$

We conclude:

$$\begin{aligned} \Pr^{\mathfrak{S}}_{\mathcal{M}}(\psi_{\mathbf{c}} \wedge \Diamond \mathit{Eff}) \cdot \frac{1 - \Pr^{\mathfrak{S}}_{\mathcal{M}}(\psi_{\mathbf{c}})}{\Pr^{\mathfrak{S}}_{\mathcal{M}}(\psi_{\mathbf{c}})} \;&\geqslant\; \Pr^{\mathfrak{S}}_{\mathcal{M}}(\psi_{\mathbf{c}}) \cdot \mathfrak{p}_{\min,\mathbf{c}} \cdot \frac{1 - \Pr^{\mathfrak{S}}_{\mathcal{M}}(\psi_{\mathbf{c}})}{\Pr^{\mathfrak{S}}_{\mathcal{M}}(\psi_{\mathbf{c}})} \\ &=\; \big(1 - \Pr^{\mathfrak{S}}_{\mathcal{M}}(\psi_{\mathbf{c}})\big) \cdot \mathfrak{p}_{\min,\mathbf{c}} \\ &>\; \Pr^{\mathfrak{S}}_{\mathcal{M}}((\neg\psi_{\mathbf{c}}) \wedge \Diamond \mathit{Eff}). \end{aligned}$$

Thus, (4.3) holds for $\mathbf{c}$ and $\mathfrak{S}$ which proves the claim. □

Lemma 38 (Criterion for GPR Causes). *Suppose Cause is a GPR cause for $\{TP, FN\}$ in $\mathcal{M}^{\min}_{\mathit{Cause}}$. Then, Cause is a GPR cause for Eff in $\mathcal{M}$.*

Proof. Again, by Lemma 32 and Corollary 33 we assume $\mathcal{M} = \mathcal{M}_{\mathit{Cause}}$ throughout the proof. From the assumption that Cause is a GPR cause for $\{TP, FN\}$ in $\mathcal{M}^{\min}_{\mathit{Cause}}$, we can conclude that the condition (GPR) holds for all schedulers $\mathfrak{U}$ that satisfy

$$\Pr^{\mathfrak{U}}_{\mathcal{M}}(\Diamond \mathit{Cause}) > 0 \qquad \text{and} \qquad \Pr^{\mathit{res}(\mathfrak{U},\pi)}_{\mathcal{M},\mathbf{c}}(\Diamond \mathit{Eff}) = \Pr^{\min}_{\mathcal{M},\mathbf{c}}(\Diamond \mathit{Eff})$$

for each finite $\mathfrak{U}$-path π from the initial state init to a state $\mathbf{c} \in \mathit{Cause}$. The idea of this proof is to show that for an arbitrary scheduler $\mathfrak{S}$ with $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Cause}) > 0$ we can always find a scheduler $\mathfrak{U}$ as above which is "worse" with respect to (GPR). For our argumentation we use several new notations which are bound to this proof:

- $\Sigma_{>0}$ denotes the set of all schedulers $\mathfrak{S}$ such that $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Cause}) > 0$ and
- $\Sigma_{>0,\min}$ denotes the set of all schedulers with $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Cause}) > 0$ such that

  $$\Pr^{\mathit{res}(\mathfrak{S},\pi)}_{\mathcal{M},\mathbf{c}}(\Diamond \mathit{Eff}) \;=\; \Pr^{\min}_{\mathcal{M},\mathbf{c}}(\Diamond \mathit{Eff})$$

  for each finite $\mathfrak{S}$-path π from the initial state init to a state $\mathbf{c} \in \mathit{Cause}$.

Note that the scheduler in $\Sigma_{>0}$ are exactly those which are considered in **(G)** . With this notation it suffices to show that for each scheduler $\mathfrak{S} \in \Sigma_{>0}$ there is a scheduler $\mathfrak{S}' \in \Sigma_{>0,\min}$ such that if (GPR) holds for $\mathfrak{S}'$ then (GPR) holds for $\mathfrak{S}$. For $\mathbf{c} \in \mathit{Cause}$, let $\Pi_{\mathbf{c}}$ denote the set of finite paths which start in init and end in $\mathbf{c}$ without visiting any other state in Cause

$$\Pi_{\mathbf{c}} = \left\{ \pi \in \mathit{Paths}_{\mathit{fin}}(\mathcal{M}) \;\middle|\; \begin{array}{l} \pi = \mathit{init}\, \alpha_0\, s_1\, \alpha_1 \ldots \alpha_{n-1}\, \mathbf{c} \text{ and} \\ \{s_1, \ldots, s_{n-1}\} \cap (\mathit{Cause} \cup \mathit{Eff}) = \varnothing \end{array} \right\}$$

Given such a path $\pi \in \Pi_{\mathbf{c}}$ and a scheduler any $\mathfrak{S}$ of $\mathcal{M}$, let

$$w^{\mathfrak{S}}_{\pi} \;=\; \Pr^{\mathit{res}(\mathfrak{S},\pi)}_{\mathcal{M},\mathbf{c}}(\Diamond \mathit{Eff})$$

and let $\mathrm{p}_\pi^{\mathfrak{S}}$ denote the probability for (the cylinder set of) π under $\mathfrak{S}$. Then the probability of reaching $\mathbf{c} \in \mathit{Cause}$ without traversing Cause beforehand is the sum of all such probabilities $\mathrm{p}_\pi^{\mathfrak{S}}$ for paths $\pi \in \Pi_{\mathbf{c}}$, that is

$$\Pr_{\mathcal{M}}^{\mathfrak{S}}((\neg \mathit{Cause}) \,\mathrm{U}\, \mathbf{c}) \;=\; \sum_{\pi \in \Pi_{\mathbf{c}}} \mathrm{p}_\pi^{\mathfrak{S}}.$$

Moreover, we can reformulate the (conditional) probabilities from (GPR):

$$\Pr_{\mathcal{M}}^{\mathfrak{S}}(\Diamond \mathit{Eff}) \;=\; \Pr_{\mathcal{M}}^{\mathfrak{S}}((\neg \mathit{Cause}) \,\mathrm{U}\, \mathit{Eff}) + \sum_{\mathbf{c} \in \mathit{Cause}} \sum_{\pi \in \Pi_{\mathbf{c}}} \mathrm{p}_\pi^{\mathfrak{S}} \cdot w_\pi^{\mathfrak{S}} \text{ and}$$

$$\Pr_{\mathcal{M}}^{\mathfrak{S}}(\Diamond \mathit{Eff} \mid \Diamond \mathit{Cause}) \;=\; \frac{1}{\Pr_{\mathcal{M}}^{\mathfrak{S}}(\Diamond \mathit{Cause})} \cdot \sum_{\mathbf{c} \in \mathit{Cause}} \sum_{\pi \in \Pi_{\mathbf{c}}} \mathrm{p}_\pi^{\mathfrak{S}} \cdot w_\pi^{\mathfrak{S}}$$

This allows us to reformulate (GPR) for $\mathfrak{S} \in \Sigma_{>0}$. Thus, (GPR) holds for $\mathfrak{S} \in \Sigma_{>0}$ if and only if

$$\begin{aligned}
&\Pr_{\mathcal{M}}^{\mathfrak{S}}(\Diamond \mathit{Eff} \mid \Diamond \mathit{Cause}) \;>\; \Pr_{\mathcal{M}}^{\mathfrak{S}}(\Diamond \mathit{Eff}) \\
\iff &\frac{1}{\Pr_{\mathcal{M}}^{\mathfrak{S}}(\Diamond \mathit{Cause})} \cdot \sum_{\mathbf{c} \in \mathit{Cause}} \sum_{\pi \in \Pi_{\mathbf{c}}} \mathrm{p}_\pi^{\mathfrak{S}} \cdot w_\pi^{\mathfrak{S}} \;>\; \Pr_{\mathcal{M}}^{\mathfrak{S}}((\neg \mathit{Cause}) \,\mathrm{U}\, \mathit{Eff}) + \sum_{\mathbf{c} \in \mathit{Cause}} \sum_{\pi \in \Pi_{\mathbf{c}}} \mathrm{p}_\pi^{\mathfrak{S}} \cdot w_\pi^{\mathfrak{S}} \\
\iff &\sum_{\mathbf{c} \in \mathit{Cause}} \sum_{\pi \in \Pi_{\mathbf{c}}} \mathrm{p}_\pi^{\mathfrak{S}} \cdot w_\pi^{\mathfrak{S}} \;>\; \Pr_{\mathcal{M}}^{\mathfrak{S}}(\Diamond \mathit{Cause}) \cdot \left(\Pr_{\mathcal{M}}^{\mathfrak{S}}((\neg \mathit{Cause}) \,\mathrm{U}\, \mathit{Eff}) + \cdot \sum_{\mathbf{c} \in \mathit{Cause}} \sum_{\pi \in \Pi_{\mathbf{c}}} \mathrm{p}_\pi^{\mathfrak{S}} \cdot w_\pi^{\mathfrak{S}} \right) \\
\iff &\big(1 - \Pr_{\mathcal{M}}^{\mathfrak{S}}(\Diamond \mathit{Cause})\big) \cdot \sum_{\mathbf{c} \in \mathit{Cause}} \sum_{\pi \in \Pi_{\mathbf{c}}} \mathrm{p}_\pi^{\mathfrak{S}} \cdot w_\pi^{\mathfrak{S}} \;>\; \Pr_{\mathcal{M}}^{\mathfrak{S}}(\Diamond \mathit{Cause}) \cdot \Pr_{\mathcal{M}}^{\mathfrak{S}}((\neg \mathit{Cause}) \,\mathrm{U}\, \mathit{Eff}).
\end{aligned} \tag{4.4}$$

Thus, let $\mathfrak{S} \in \Sigma_{>0}$ and consider an MD scheduler $\mathfrak{T}$ that minimizes the probability to reach Eff from every state. In particular, $\mathrm{p}_{\min,\mathbf{c}} = w_\pi^{\mathfrak{T}} \leqslant w_\pi^{\mathfrak{S}}$ for every state $\mathbf{c} \in \mathit{Cause}$ and every path $\pi \in \Pi_{\mathbf{c}}$ (recall that $\mathrm{p}_{\min,\mathbf{c}} = \Pr_{\mathcal{M},\mathbf{c}}^{\min}(\Diamond \mathit{Eff})$). Moreover, the scheduler $\mathfrak{S}$ can be transformed into a scheduler $\mathfrak{S}_{\mathfrak{T}} \in \Sigma_{>0,\min}$ which is "equivalent" to $\mathfrak{S}$ with respect to the (GPR) condition. More concretely, let $\mathfrak{S}_{\mathfrak{T}}$ denote the scheduler which behaves identical to $\mathfrak{S}$ as long as no state $\mathbf{c} \in \mathit{Cause}$ has been visited. After a state $\mathbf{c} \in \mathit{Cause}$ has been visited $\mathfrak{S}_{\mathfrak{T}}$ behaves identical to $\mathfrak{T}$. Thus, for each $\pi \in \Pi_{\mathbf{c}}$ we have

$$\mathrm{p}_\pi^{\mathfrak{S}} = \mathrm{p}_\pi^{\mathfrak{S}_{\mathfrak{T}}} \qquad \text{and} \qquad \mathit{res}(\mathfrak{S}_{\mathfrak{T}}, \pi) = \mathfrak{T}.$$

This yields that the probability to reach $\mathbf{c} \in \mathit{Cause}$ from init is the same for both $\mathfrak{S}$ and $\mathfrak{S}_{\mathfrak{T}}$, i.e., $\Pr_{\mathcal{M}}^{\mathfrak{S}}(\Diamond \mathbf{c}) = \Pr_{\mathcal{M}}^{\mathfrak{S}_{\mathfrak{T}}}(\Diamond \mathbf{c})$. Therefore, the reachability probabilities for Cause are also the same for both scheduler, that is $\Pr_{\mathcal{M}}^{\mathfrak{S}}(\Diamond \mathit{Cause}) = \Pr_{\mathcal{M}}^{\mathfrak{S}_{\mathfrak{T}}}(\Diamond \mathit{Cause})$. This means $\mathfrak{S}_{\mathfrak{T}} \in \Sigma_{>0}$ and hence $\mathfrak{S}_{\mathfrak{T}} \in \Sigma_{>0,\min}$ by construction. Moreover, $\mathfrak{S}$ and $\mathfrak{S}_{\mathfrak{T}}$ reach Eff without visiting Cause with the same probability, i.e., $\Pr_{\mathcal{M}}^{\mathfrak{S}}((\neg \mathit{Cause}) \,\mathrm{U}\, \mathit{Eff}) = \Pr_{\mathcal{M}}^{\mathfrak{S}_{\mathfrak{T}}}((\neg \mathit{Cause}) \,\mathrm{U}\, \mathit{Eff})$.

Combining the equalities yields that if (4.4) holds for $\mathfrak{S}_{\mathfrak{T}}$ then (4.4) holds for $\mathfrak{S}$. As (4.4) holds for $\mathfrak{S}_{\mathfrak{T}}$ by assumption, this completes the proof. □

Proof of Lemma 36. As above, we assume $\mathcal{M} = \mathcal{M}_{Cause}$ by Lemma 32 and Corollary 33. By Lemma 37 and 38 we have that each GPR (resp. SPR) cause for $\{TP, FN\}$ in $\mathcal{M}^{\min}_{Cause}$ is also an GPR (resp. SPR) cause for Eff in $\mathcal{M}$. Furthermore, by structure of $\mathcal{M} = \mathcal{M}_{Cause}$ and $\mathcal{M}^{\min}_{Cause}$ we have for all $\mathbf{c} \in \mathit{Cause}$ that $\Pr^{\max}_{\mathcal{M}}((\neg \mathit{Cause}) \,\mathrm{U}\, \mathbf{c}) = \Pr^{\max}_{\mathcal{M}^{\min}_{Cause}}((\neg \mathit{Cause}) \,\mathrm{U}\, \mathbf{c}) > 0$. So, let Cause be an GPR (resp. SPR) cause for Eff in $\mathcal{M}$. By assumption $\mathcal{M} = \mathcal{M}_{Cause}$ the transformation to $\mathcal{M}^{\min}_{Cause}$ only restricts the class of considered scheduler for the PR conditions (G) and (S) . Thus, Cause is also an GPR (resp. SPR) cause for $\{TP, FN\}$ in $\mathcal{M}^{\min}_{Cause}$. □

This result allows us to assume $\mathcal{M} = \mathcal{M}^{\min}_{Cause}$ when considering the decision problems for GPR and SPR causes. Note, that all of these transformations do have an impact on both randomization and memory requirements on the scheduler. Given a scheduler $\mathfrak{U}$ of $\mathcal{M}_{Cause}$, a corresponding scheduler $\mathfrak{S}$ of $\mathcal{M}$ inherits randomization for the balancing between $\alpha_{\min}$ and $\alpha_{\max}$. Furthermore, the two modes "before Cause" and "after Cause" need additional memory cells. However, the traversal of end-components does not need further memory as long as the action τ leading to TN is taken with probability 0 or 1 by the following result.

Theorem 39 (Traversing ECs Memoryless). *Let $\mathcal{M}$ be an MDP with pairwise disjoint action sets for all states. Then, for each MR scheduler $\mathfrak{U}$ for the MEC-quotient of $\mathcal{M}$ with $\mathfrak{U}(s_{\mathcal{E}}, \tau) \in \{0, 1\}$ for each MEC $\mathcal{E}$ of $\mathcal{M}$ there is an MR scheduler $\mathfrak{S}$ for $\mathcal{M}$ such that every action α of $\mathcal{M}$ that does not belong to an MEC of $\mathcal{M}$, has the same expected frequency under $\mathfrak{U}$ and $\mathfrak{S}$.*

Proof. Let $\mathfrak{U}$ be an MR scheduler for the MEC-quotient $MEC(\mathcal{M})$ with $\mathfrak{U}(s_{\mathcal{E}}, \tau) \in \{0, 1\}$ for each MEC $\mathcal{E}$ of $\mathcal{M}$. We consider the following extension $\mathcal{M}'$ of $\mathcal{M}$: The state space of $\mathcal{M}$ is extended by a new terminal state $\perp$ and a fresh action τ is enabled in each state $s \in S$ that belongs to a MEC of $\mathcal{M}$. Action τ leads to $\perp$ with probability 1. All remaining transition probabilities are as in $\mathcal{M}$. So, $\mathcal{M}'$ is obtained from $\mathcal{M}$ by allowing a transition to a new terminal state $\perp$ from each state that belongs to a MEC.

Now, we first provide a finite-memory scheduler $\mathfrak{S}$ for $\mathcal{M}'$ that leaves each MEC $\mathcal{E}$ for which $\mathfrak{U}(s_{\mathcal{E}}, \tau) = 0$ via the state action pair (s, α) with probability $\mathfrak{U}(s_{\mathcal{E}}, \alpha)$. Recall that we assume pairwise disjoint action sets $Act(s)$. Thus, each action is enabled in at most one state. Moreover, for a state $s_{\mathcal{E}}$ in $MEC(\mathcal{M})$ representing the MEC $\mathcal{E}$ of $\mathcal{M}$ the enabled actions are precisely the enabled actions in states of $\mathcal{E}$ which leave $\mathcal{E}$ with non-zero probability. (Section 2.1.6)

We now define the scheduler $\mathfrak{S}$:

- For $s \in S$ and $\alpha \in Act$ in $\mathcal{M}$ not belonging to a MEC we set $\mathfrak{S}(s, \alpha) = \mathfrak{U}(s, \alpha)$, that is the scheduler $\mathfrak{S}$ is mimicking $\mathfrak{U}$ memoryless.
- For $s \in S$ in an EC $\mathcal{E}$ of $\mathcal{M}$ with $\mathfrak{U}(s_{\mathcal{E}}, \tau) = 1$ we define $\mathfrak{S}(s, \tau) = 1$.

If a MEC $\mathcal{E}$ of $\mathcal{M}$ with $\mathfrak{U}(s_{\mathcal{E}}, \tau) = 0$ is entered, $\mathfrak{S}$ makes use of finitely many memory modes by the following:

Enumerate the state action pairs (s, α) where s belongs to $\mathcal{E}$, but α does not belong to $\mathcal{E}$, and for which $\mathfrak{U}(s_{\mathcal{E}}, \alpha) > 0$ by $(s_1, \alpha_1), \ldots, (s_k, \alpha_k)$. Further, let $p_i = \mathfrak{U}(s_{\mathcal{E}}, \alpha_i) > 0$ for all $1 \leqslant i \leqslant k$. By assumption $\sum_{1 \leqslant i \leqslant k} p_i = 1$. When entering $\mathcal{E}$, the scheduler works in k memory modes $1, \ldots, k$ until an action α that does not belong to $\mathcal{E}$ is scheduled starting in memory mode 1. In each memory mode i, $\mathfrak{S}$ follows an MD scheduler for $\mathcal{E}$ that reaches s_i with probability 1 from all states of $\mathcal{E}$. Once, s_i is reached, $\mathfrak{S}$ chooses action α_i with probability

$$q_i = \frac{p_i}{1 - \sum_{j<i} p_j}.$$

Now $\mathfrak{S}$ leaves $\mathcal{E}$ via (s_k, α_k) with probability 1 if it reaches the last memory mode k. As $\mathfrak{S}$ behaves MD in each mode, it leaves the end-component $\mathcal{E}$ after finitely many steps in expectation. Furthermore, for each $i \leqslant k$, it leaves $\mathcal{E}$ via (s_i, α_i) with probability $(1 - \sum_{j<i} p_j) \cdot q_i = p_i$. Since the behavior of $\mathfrak{U}$ in $MEC(\mathcal{M})$ is mimicked by $\mathfrak{S}$ in $\mathcal{M}'$, we conclude that the expected frequency of actions in $\mathcal{M}$ which do not belong to an end-component is the same in $\mathcal{M}'$ under $\mathfrak{S}$ and in $MEC(\mathcal{M})$ under $\mathfrak{U}$.

The expected frequency of each state-action pair of $\mathcal{M}'$ under $\mathfrak{S}$ is finite, since each MEC of $\mathcal{M}'$ is left after finitely many steps in expectation. In the terminology of [Kal20], the scheduler $\mathfrak{S}$ is *transient*. By [Kal20, Theorem 9.16], this implies that there is an MR scheduler $\mathfrak{T}$ for $\mathcal{M}'$ under which the expected frequency of state-action pairs is the same as under $\mathfrak{S}$ and thus the expected frequency in $\mathcal{M}'$ of actions α of $\mathcal{M}$ that do not belong to an end-component is the same as under $\mathfrak{U}$ in $MEC(\mathcal{M})$.

Finally, we modify the scheduler $\mathfrak{S}$ for $\mathcal{M}'$ to get a a scheduler $\mathfrak{T}$ for $\mathcal{M}$. For each end-component $\mathcal{E}$ of $\mathcal{M}$ with $\mathfrak{U}(s_{\mathcal{E}}, \tau) = 1$, we fix a memoryless scheduler $\mathfrak{T}_{\mathcal{E}}$ that does not leave the end-component. Now, whenever a state $s \in S$ in $\mathcal{E}$ is visited, the modified scheduler switches to the behavior of $\mathfrak{T}_{\mathcal{E}}$ instead of choosing action τ with probability 1. Clearly, this does not affect the expected frequency of actions of $\mathcal{M}$ that do not belong to an end-component and hence the modified scheduler is as claimed in the theorem. □

Remark 40 (Complexity of Traversing ECs). The proof of Theorem 39 above provides an algorithm how to obtain the scheduler $\mathfrak{S}$ from $\mathfrak{U}$. The number of memory modes of the intermediately constructed finite-memory scheduler is bounded by the number of state-action pairs of $\mathcal{M}$. Further, in each memory mode during the traversal of a MEC, the scheduler behaves in a memoryless deterministic way. Hence, the induced Markov chain is of size polynomial in the size of the MDP $\mathcal{M}$ and the representation of the scheduler $\mathfrak{U}$. Therefore, also the expected frequencies of all state-action pairs under the intermediate finite-memory scheduler and hence under $\mathfrak{S}$ can be computed in time polynomial in the size of the MDP $\mathcal{M}$ and the representation of the scheduler $\mathfrak{U}$. So, also the scheduler $\mathfrak{S}$ itself can be derived from these expected frequencies in polynomial time. ◁

4.1.2 Checking the SPR Condition and the Existence of PR Causes

With all of the preliminary transformation to $\mathcal{M}^{\min}_{Cause}$ in mind we now turn to checking a cause candidate for strict probability-raising causality. For this, we restate the SPR condition for an MDP $\mathcal{M} = (S, Act, P, init)$ with non-empty set of terminal effect states $\mathit{Eff} \subset S$ and a non-empty set $\mathit{Cause} \subseteq S \setminus \mathit{Eff}$:

(S) For each state $c \in \mathit{Cause}$ and each scheduler $\mathfrak{S}$ of $\mathcal{M}$ with $\Pr^{\mathfrak{S}}_{\mathcal{M}}((\neg \mathit{Cause}) \,\mathrm{U}\, c) > 0$:

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff} \mid (\neg \mathit{Cause}) \,\mathrm{U}\, c) > \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff}). \tag{SPR}$$

The basis of both checking the existence of PR causes or checking the SPR condition **(S)** for a given cause candidate is the following polynomial time algorithm to check whether **(S)** holds in a given state $c \in S$ for all schedulers $\mathfrak{S}$ of $\mathcal{M}$ with $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond c) > 0$:

Algorithm 41 (Checking SPR for Singletons).

Input MDP $\mathcal{M}$, state $c \in S$, non-empty set of terminal states $\mathit{Eff} \subset S$

Task Decide whether (SPR) holds in c for all schedulers $\mathfrak{S}$.

0. Transform $\mathcal{M}$ to $\mathcal{N} = \mathcal{M}^{\min}_{c}$ and compute $p_{\min,c} = \Pr^{\min}_{\mathcal{M},c}(\Diamond \mathit{Eff})$ and $q_s = \Pr^{\max}_{\mathcal{N},s}(\Diamond \mathit{Eff})$ for each state s in $\mathcal{N}$.
1. If $p_{\min,c} > q_{init}$, then return "yes, (SPR) holds for c".
2. If $p_{\min,c} < q_{init}$, then return "no, (SPR) does not hold for c".
3. Suppose $p_{\min,c} = q_{init}$. Let $A(s) = \{\alpha \in Act_{\mathcal{N}}(s) \mid q_s = \sum_{t \in S_{\mathcal{N}}} P_{\mathcal{N}}(s, \alpha, t) \cdot q_t\}$ for each non-terminal state s in $\mathcal{N}$. Let $\mathcal{N}^{\max}_{c}$ denote the sub-MDP of $\mathcal{N} = \mathcal{M}^{\min}_{c}$ induced by the state-action pairs (s, α) where $\alpha \in A(s)$.
   - 3.1 If c is reachable from $init$ in $\mathcal{N}^{\max}_{c}$, then return "no, (SPR) does not hold for c".
   - 3.2 If c is not reachable from $init$ in $\mathcal{N}^{\max}_{c}$, then return "yes, (SPR) holds for c".

Note that the set of actions $A(s)$ in case 3. of the algorithm consists of exactly the maximizing actions for $\Diamond \mathit{Eff}$ in s. Therefore, the MDP $\mathcal{N}^{\max}_{c}$ is a sub-MDP of $\mathcal{M}^{\min}_{c}$ where only maximizing actions for $\Diamond \mathit{Eff}$ are enabled.

As the construction of the minimizing MDP $\mathcal{M}^{\min}_{c}$ suggests, the two values compared by the algorithm are instances of worst-case scheduler as the the probability to reach Eff is minimized whenever c has been reached, while it is maximized if c was not seen yet. If in such a scenario we are in case 1. then $p_{\min,c} > q_{init}$ then c obviously satisfies (SPR). In the case 2. $p_{\min,c} < q_{init}$ we can build a scheduler which refuses (SPR) for c. Lastly, in the corner case 3. $p_{\min,c} = q_{init}$ a treatment by a reachability analysis is needed, as seen in the shortly following Example 42.

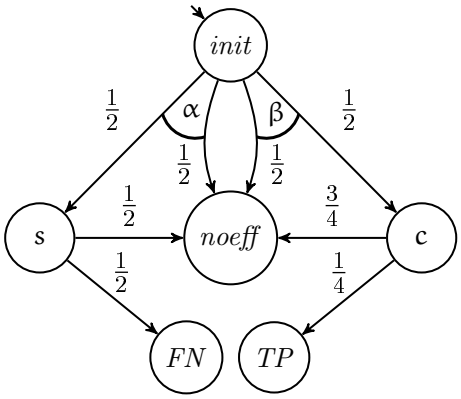

Figure 4.4: Example MDP $\mathcal{M}_{\mathbf{c}}^{\min}$

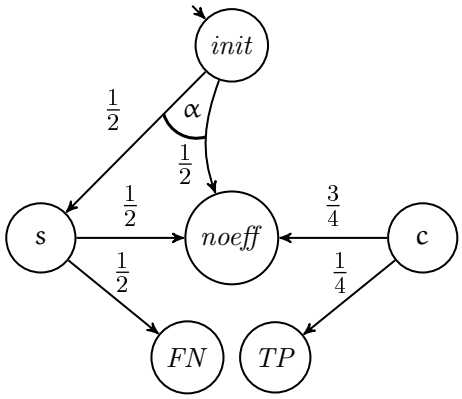

Figure 4.5: Transformed MDP $\mathcal{N}_{\mathbf{c}}^{\max}$

We have already applied the idea of this algorithm in Section 3.2.2 where we consider the Racetrack MDP given by Figures 3.10 and 3.11. For the state $\mathbf{c}_3 = (3,3,1,0)$ describing the position $(\mathrm{x},\mathrm{y}) = (3,3)$ in the grid with a velocity of 1 on the x-axis we show that $\mathbf{c}_3$ is satisfying (SPR) by calculating

$$\mathrm{p}_{\min,\mathbf{c}_3} = \Pr_{\mathbf{c}_3}^{\min}(\lozenge \mathit{Eff}) = 0.6102 > 0.4096 = \Pr^{\max}(\lozenge \mathit{Eff}) > \Pr_{\mathcal{N}}^{\max}(\lozenge \mathit{Eff})$$

using the PRISM model checker [KNP11], where $\mathcal{N} = \mathcal{M}_{\mathbf{c}}^{\min}$

Example 42 (MDP with only Maximizing Actions). For a visualization of the transformation to $\mathcal{N}_{\mathbf{c}}^{\max}$ consider the MDP $\mathcal{M} = \mathcal{M}_{\mathbf{c}}^{\min}$ from Figure 4.4. For $\mathit{Cause} = \{\mathbf{c}\}$ and $\mathit{Eff} = \{\mathit{TP}, \mathit{FN}\}$ we are in case 3. of Algorithm 41 as

$$\Pr_{\mathcal{M},\mathbf{c}}^{\min}(\lozenge \mathit{Eff}) = \mathrm{p}_{\min,\mathbf{c}} = \frac{1}{4} = \Pr_{\mathcal{M}}^{\max}(\lozenge \mathit{Eff}) = \mathrm{q}_{\mathit{init}}.$$

The only non-deterministic choice is in the state init. We have $\mathrm{A}(\mathit{init}) = \{\alpha\}$ since α is the only maximizing action for $\lozenge \mathit{Eff}$ in init. Thus, in the resulting MDP $\mathcal{N}_{\mathbf{c}}^{\max}$, depicted in Figure 4.5, all other actions in init are deleted. Here, we are in case 3.2 since $\mathbf{c}$ is not reachable from the initial state in $\mathcal{N}_{\mathbf{c}}^{\max}$. ◁

Lemma 43. *Algorithm 41 is sound and runs in polynomial time.*

Proof. First, we show the soundness of Algorithm 41. By the virtue of Corollary 33 and Lemma 36 stating the soundness of the transformation $\mathcal{M}$ to $\mathcal{M}_{\mathbf{c}}^{\min}$ it suffices to show that Algorithm 41 returns the correct answers "yes" or "no" when the task is to check whether the singleton $\mathit{Cause} = \{\mathbf{c}\}$ is an SPR cause in $\mathcal{N} = \mathcal{M}_{\mathbf{c}}^{\min}$. Recall the notation $\mathrm{q}_{\mathrm{s}} = \Pr_{\mathcal{N},\mathrm{s}}^{\max}(\lozenge \mathit{Eff})$. Note that $(\neg \mathit{Cause})\,\mathrm{U}\,\mathbf{c}$ is equivalent to $\lozenge \mathbf{c}$ as $\mathbf{c} \in \mathit{Cause}$ is a singleton.

For every scheduler $\mathfrak{S}$ of $\mathcal{N}$ we have $\Pr_{\mathcal{N},\mathbf{c}}^{\mathfrak{S}}(\lozenge \mathit{Eff}) = \mathrm{p}_{\min,\mathbf{c}}$ by construction. For a scheduler $\mathfrak{S}$ of $\mathcal{N}$ with $\Pr_{\mathcal{N}}^{\mathfrak{S}}(\lozenge \mathbf{c}) > 0$ we also have $\Pr_{\mathcal{N}}^{\mathfrak{S}}(\lozenge \mathit{Eff} \mid \lozenge \mathbf{c}) = \mathrm{p}_{\min,\mathbf{c}}$ as $\mathbf{c}$ has no cycles and only transitions to the terminal states $\mathit{TP} \in \mathit{Eff}$ and $\mathit{FP} \notin \mathit{Eff}$.

In the corner case of $\mathfrak{p}_{\min,\mathfrak{c}} = 0$ Algorithm 41 correctly answers "no" (case 2 or 3.1). Thus, we consider $\mathfrak{p}_{\min,\mathfrak{c}} > 0$. By the above argumentation, the SPR condition for $\mathfrak{c}$ reduces to $\mathfrak{p}_{\min,\mathfrak{c}} > \Pr^{\mathfrak{S}}_{\mathcal{N}}(\Diamond \mathit{Eff})$ for all schedulers $\mathfrak{S}$ of $\mathcal{N}$ with $\Pr^{\mathfrak{S}}_{\mathcal{N}}(\Diamond \mathfrak{c}) > 0$. We will now consider each distinct case of Algorithm 41:

1. (i.e., if $\mathfrak{p}_{\min,\mathfrak{c}} > \mathfrak{q}_{\mathit{init}}$) The answer "yes" is sound as then $\mathfrak{p}_{\min,\mathfrak{c}} > \Pr^{\max}_{\mathcal{N}}(\Diamond \mathit{Eff}) = \mathfrak{q}_{\mathit{init}}$.
2. (i.e., if $\mathfrak{p}_{\min,\mathfrak{c}} < \mathfrak{q}_{\mathit{init}}$) Let $\mathfrak{T}$ be an MD scheduler which realizes $\Pr^{\mathfrak{T}}_{\mathcal{N},\mathfrak{s}}(\Diamond \mathit{Eff}) = \mathfrak{q}_{\mathfrak{s}}$ for each state $\mathfrak{s}$ in $\mathcal{N}$. Furthermore, let $\mathfrak{S}$ be an MD scheduler with $\Pr^{\mathfrak{S}}_{\mathcal{N}}(\Diamond \mathfrak{c}) > 0$ and which chooses τ in states $\mathfrak{s}_{\mathcal{E}}$ representing MECs of $\mathcal{M}$ identical to $\mathfrak{T}$.

   Let λ be any real number with $1 > \lambda > \frac{\mathfrak{p}_{\min,\mathfrak{c}}}{\mathfrak{q}_{\mathit{init}}}$ and let $\mathcal{K}$ denote the sub-MDP of $\mathcal{N}$ with state space S where the enabled actions of are the actions scheduled under one of the schedulers $\mathfrak{T}$ or $\mathfrak{S}$. Let now $\mathfrak{U}$ be the MR scheduler $\lambda\mathfrak{T} \oplus (1-\lambda)\mathfrak{S}$ defined as in Lemma 7 for $\mathcal{K}$. Then we have for all terminal states $\mathfrak{t}$ of $\mathcal{N}$ and $\mathfrak{t} = \mathfrak{c}$

$$\Pr^{\mathfrak{U}}_{\mathcal{N}}(\Diamond \mathfrak{t}) = \lambda \Pr^{\mathfrak{T}}_{\mathcal{N}}(\Diamond \mathfrak{t}) + (1-\lambda)\Pr^{\mathfrak{S}}_{\mathcal{N}}(\Diamond \mathfrak{t}).$$

   Hence:

$$\Pr^{\mathfrak{U}}_{\mathcal{N}}(\Diamond \mathfrak{c}) \;\geqslant\; (1-\lambda) \cdot \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathfrak{c}) \;>\; 0 \text{ and}$$
$$\Pr^{\mathfrak{U}}_{\mathcal{N}}(\Diamond \mathit{Eff}) \geqslant\; \lambda \cdot \Pr^{\mathfrak{T}}_{\mathcal{M}}(\Diamond \mathit{Eff}) \;=\; \lambda \cdot \mathfrak{q}_{\mathit{init}} \;>\; \mathfrak{p}_{\min,\mathfrak{c}}$$

   Thus, scheduler $\mathfrak{U}$ is a witness why (SPR) does not hold for $\mathfrak{c}$.

3.1 (i.e., if $\mathfrak{p}_{\min,\mathfrak{c}} = \mathfrak{q}_{\mathit{init}}$ and $\mathfrak{c}$ reachable in $\mathcal{N}^{\max}_{\mathfrak{c}}$) Pick an MD scheduler $\mathfrak{S}$ of $\mathcal{N}^{\max}_{\mathfrak{c}}$ such that $\mathfrak{c}$ is reachable from init via $\mathfrak{S}$ and $\Pr^{\mathfrak{S}}_{\mathcal{N}^{\max}_{\mathfrak{c}},\mathfrak{s}}(\Diamond \mathit{Eff}) = \mathfrak{q}_{\mathfrak{s}}$ for all states $\mathfrak{s}$. Then we have

$$\Pr^{\mathfrak{S}}_{\mathcal{N}^{\max}_{\mathfrak{c}}}(\Diamond \mathit{Eff} \mid \Diamond \mathfrak{c}) = \mathfrak{p}_{\min,\mathfrak{c}} \;=\; \mathfrak{q}_{\mathit{init}} = \Pr^{\mathfrak{S}}_{\mathcal{N}^{\max}_{\mathfrak{c}}}. \tag{4.5}$$

   We can trace back $\mathfrak{S}$ to also define a scheduler $\mathfrak{U}$ in $\mathcal{N}$ by behaving identically in each state $\mathfrak{s} \in S$. Therefore, (SPR) does not hold for $\mathfrak{c}$ and $\mathfrak{U}$.

3.2 (i.e., if $\mathfrak{p}_{\min,\mathfrak{c}} = \mathfrak{q}_{\mathit{init}}$ and $\mathfrak{c}$ not reachable in $\mathcal{N}^{\max}_{\mathfrak{c}}$) We have $\Pr^{\mathfrak{S}}_{\mathcal{N}}(\Diamond \mathfrak{c}) = 0$ for all schedulers $\mathfrak{S}$ for $\mathcal{N}$ with $\Pr^{\mathfrak{S}}_{\mathcal{N}}(\Diamond \mathit{Eff}) = \mathfrak{p}_{\min,\mathfrak{c}} = \mathfrak{q}_{\mathit{init}}$. But then any scheduler $\mathfrak{U}$ of $\mathcal{N}$ with $\Pr^{\mathfrak{U}}_{\mathcal{N}}(\Diamond \mathfrak{c}) > 0$ has a higher probability of reaching Eff than the scheduler of $\mathcal{N}^{\max}_{\mathfrak{c}}$. For such scheduler this implies $\mathfrak{p}_{\min,\mathfrak{c}} = \mathfrak{q}_{\mathit{init}} > \Pr^{\mathfrak{U}}_{\mathcal{N}}(\Diamond \mathit{Eff})$ as required for (SPR).

The polynomial runtime of Algorithm 41 follows from the fact that minimal and maximal reachability probabilities and hence also the MDPs $\mathcal{M}^{\min}_{\mathfrak{c}}$ and its sub-MDP $\mathcal{N}^{\max}_{\mathfrak{c}}$ can be computed in polynomial time [BK08]. □

Whenever the given set of states Cause is not a singleton we can apply Algorithm 41 to all states $\mathfrak{c} \in \mathit{Cause}$ to check condition (S) . Together with standard algorithms to check the existence of a path satisfying $(\neg \mathit{Cause})\,\mathrm{U}\,\mathfrak{c}$ for every state $\mathfrak{c} \in \mathit{Cause}$, we obtain:

Theorem 44 (Checking SPR Causes). *The problem "given $\mathcal{M} = (S, Act, P, init)$ and a non-empty set of terminal states $Eff \subset S$, deciding whether a non-empty set $Cause \subseteq S \setminus Eff$ is an SPR cause for Eff in $\mathcal{M}$" is in* P.

Remark 45 (Memory Requirements for (S)). As the soundness proof for Algorithm 41 shows: If $Cause$ does not satisfy **(S)** , then there is an MR scheduler $\mathfrak{T}$ for $\mathcal{N} = \mathcal{M}^{\min}_{Cause}$ witnessing the violation of (SPR). Scheduler $\mathfrak{T}$ corresponds to a finite-memory (randomized) scheduler $\mathfrak{S}$ with two memory cells for $\mathcal{M}$: "before $Cause$" (where $\mathfrak{S}$ behaves as $\mathfrak{T}$) and "after $Cause$" (where $\mathfrak{T}$ behaves as an MD scheduler minimizing the effect probability). Note that the traversal of end-components does not need additional memory in the original MDP as the realization of $q_{init} = \Pr^{\max}_{\mathcal{N}}(\Diamond Eff)$ only needs memoryless deterministic scheduler. Therefore, Theorem 39 applies as then also τ will only be chosen deterministically with probability 0 or 1. ◁

Moreover, we can show that in general both memorization and randomization are needed to witness a violation of the SPR condition. Therefore, the requirements specified in Remark 45 are minimal.

Remark 46 (Memory needed for refuting PR Condition). Let $\mathcal{M}$ be the MDP in Figure 4.6 with $Eff = \{eff\}$ and $Cause = \{c\}$. In this MDP only state s has a non-deterministic choice. Here, $Cause$ is not an SPR cause, since we can consider the finite-memory deterministic scheduler $\mathfrak{T}$ that schedules β for the first visit of s and α for the second visit of s which yields:

$$\Pr^{\mathfrak{T}}_{\mathcal{M}}(\Diamond eff) = \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2}\cdot 1\cdot\frac{1}{4} = \frac{5}{16} > \frac{1}{4} = \Pr^{\mathfrak{T}}_{\mathcal{M}}(\Diamond eff \mid \Diamond c).$$

Denote the MR scheduler reaching c with positive probability as $\mathfrak{S}_\lambda$ with $\mathfrak{S}_\lambda(s)(\alpha) = \lambda$ and $\mathfrak{S}_\lambda(s)(\beta) = 1-\lambda$ for some $\lambda \in [0,1)$. Then, $\Pr^{\mathfrak{S}_\lambda}_{\mathcal{M},s}(\Diamond eff) > 0$ and

$$\Pr^{\mathfrak{S}_\lambda}_{\mathcal{M}}(\Diamond eff) = \frac{1}{2}\cdot\Pr^{\mathfrak{S}_\lambda}_{\mathcal{M},s}(\Diamond eff) < \Pr^{\mathfrak{S}_\lambda}_{\mathcal{M},s}(\Diamond eff) = \Pr^{\mathfrak{S}_\lambda}_{\mathcal{M},c}(\Diamond eff) = \Pr^{\mathfrak{S}_\lambda}_{\mathcal{M}}(\Diamond eff|\Diamond c).$$

Thus, the (SPR) (and thus (GPR)) condition holds for $Cause$ and Eff under all memoryless schedulers reaching $Cause$ with positive probability, although $Cause$ is not an SPR cause. ◁

Remark 47 (Randomization needed for refuting PR Condition). Consider the MDP $\mathcal{M}$ in Figure 4.7 with $Eff = \{FN, TP\}$ and $Cause = \{c\}$. Here the state FN is not covered by the cause whereas TP is, hence their names. The two MD schedulers $\mathfrak{S}_\alpha$ and $\mathfrak{S}_\beta$ which select α resp. β for the initial state $init$ are the only deterministic schedulers. As $\mathfrak{S}_\alpha$ does not reach c, it is irrelevant for the PR conditions. On the other hand $\mathfrak{S}_\beta$ satisfies (SPR) and (GPR) since

$$\Pr^{\mathfrak{S}_\beta}_{\mathcal{M}}(\Diamond Eff \mid \Diamond c) = \frac{1}{2} > \frac{1}{4} = \Pr^{\mathfrak{S}_\beta}_{\mathcal{M}}(\Diamond Eff).$$

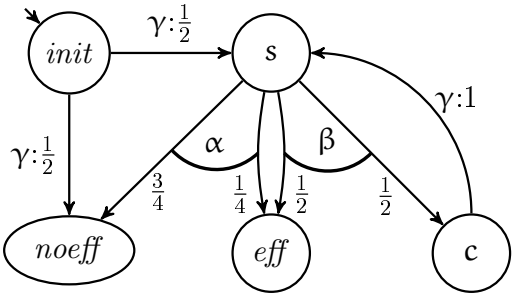

Figure 4.6: MDP $\mathfrak{M}$ from Remark 46

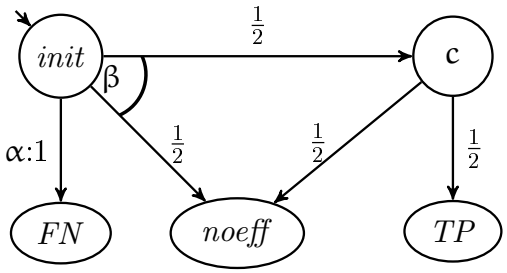

Figure 4.7: MDP $\mathfrak{M}$ from Remark 47

The MR scheduler $\mathfrak{T}$ which selects α and β with probability $\frac{1}{2}$ in *init* also reaches $\mathbf{c}$ with positive probability but violates (SPR) and (GPR) as

$$\Pr^{\mathfrak{T}}_{\mathfrak{M}}(\Diamond \mathit{Eff} \mid \Diamond \mathbf{c}) \;=\; \frac{1}{2} \;<\; \frac{5}{8} \;=\; \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \;=\; \Pr^{\mathfrak{T}}_{\mathfrak{M}}(\Diamond \mathit{Eff}).$$

Therefore, randomization might be needed to witness a violation of PR conditions. ◁

As previously stated, the algorithm for checking the (SPR) condition for a singleton $\{\mathbf{c}\}$ is also the basis to decide whether there exists any kind of probability-raising cause, strict or global, in the MDP for a given effect. This fact comes from equivalences for the existence as stated in the following Lemma 48.

Lemma 48 (Criterion for the Existence of PR Causes). *Let $\mathfrak{M}$ be an MDP and Eff a nonempty set of terminal states. The following statements are equivalent:*

(a) There is a set $\mathit{Cause} \subseteq S \setminus \mathit{Eff}$ which constitutes an SPR cause for Eff in $\mathfrak{M}$.

(b) There is a set $\mathit{Cause} \subseteq S \setminus \mathit{Eff}$ which constitutes a GPR cause for Eff in $\mathfrak{M}$.

(c) There is a state $\mathbf{c}_0 \in S \setminus \mathit{Eff}$ such that the singleton $\{\mathbf{c}_0\}$ is an SPR cause (and therefore a GRP cause) for Eff in $\mathfrak{M}$.

Proof. Obviously, statement (c) implies statements (a) and (b). The implication "(a) $\Longrightarrow$ (b)" follows from the fact that **(S)** implies **(G)** (Proposition 15). Thus, we turn to the proof of "(b) $\Longrightarrow$ (c)". For this, we assume that we are given a GPR cause Cause for Eff in $\mathfrak{M}$. For $\mathbf{c} \in \mathit{Cause}$, let $\mathfrak{p}_{\min,\mathbf{c}} = \Pr^{\min}_{\mathbf{c}}(\Diamond \mathit{Eff})$. We then select a state which contributes the maximal increase in probability to reach effect. Thus, we pick a state $\mathbf{c}_0 \in \mathit{Cause}$ such that $\mathfrak{p}_{\min,\mathbf{c}_0} = \max\{\mathfrak{p}_{\min,\mathbf{c}} : \mathbf{c} \in \mathit{Cause}\}$. Let $\mathfrak{S}$ be a scheduler for $\mathfrak{M}$ which minimizes the effect probability whenever it visits a state in Cause, and which visits Cause with positive probability. By Lemma 36 it is sufficient to consider such scheduler for the satisfaction of the PR conditions **(S)** and **(G)** . Then, the conditional probability $\Pr^{\mathfrak{S}}(\Diamond \mathit{Eff} \mid \Diamond \mathit{Cause})$ is a weighted average of the values $\mathfrak{p}_{\min,\mathbf{c}}$ for $\mathbf{c} \in \mathit{Cause}$ as shown in the proof of Proposition 15. We thus have

$$\mathfrak{p}_{\min,\mathbf{c}_0} \;\geqslant\; \Pr^{\mathfrak{S}}(\Diamond \mathit{Eff} \mid \Diamond \mathit{Cause}) \;>\; \Pr^{\mathfrak{S}}(\Diamond \mathit{Eff}).$$

As $\mathfrak{S}$ was arbitrary with the aforementioned properties, we conclude that $\{\mathbf{c}_0\}$ is both an SPR and a GPR cause for Eff. □

With these results we see that it is sufficient to check whether there exists a state $\mathbf{c} \in \mathrm{S} \setminus \mathit{Eff}$ such that $\{\mathbf{c}\}$ is an SPR cause for Eff in $\mathcal{M}$. This gives rise to a decision procedure for the existence of PR causes using Algorithm 41.

Lemma 49 (Checking the Existence of PR Causes). *Given an MDP $\mathcal{M}$ and a set of terminal states Eff, the existence of GPR and SPR causes can be decided with Algorithm 41 in P.*

Proof. We can use Algorithm 41 for each state $\mathrm{S} \setminus \mathit{Eff}$ to check whether some state is a singleton SPR cause. As Algorithm 41 runs in polynomial time, we conclude that we can decide whether there exists an SPR and/or GPR cause for Eff in $\mathcal{M}$ in P. □

Furthermore, we notice that the resulting algorithm only yields singletons as witnesses for the existence of PR causes for Eff in $\mathcal{M}$. As discussed in the introduction to quality measures (Chapter 3.3), this motivates the computation of quality-optimal causes. On the other hand, if there is no singleton PR cause, then the algorithm yields for each state $\mathbf{s} \in \mathrm{S} \setminus \mathit{Eff}$ a scheduler witnessing the violation of (SPR) for $\mathbf{s}$. Such a scheduler would have at most two memory-cells for "before" and "after cause" and may need memorization by the same argumentation as in Remark 45.

4.1.3 Checking the GPR Condition

In this section, we are concerned with checking the global probability-raising condition **(G)** . Recall, we are given an MDP $\mathcal{M} = (\mathrm{S}, \mathit{Act}, \mathrm{P}, \mathit{init})$, a non-empty set of terminal states $\mathit{Eff} \subset \mathrm{S}$ and a non-empty cause candidate $\mathit{Cause} \subseteq \mathrm{S} \setminus \mathit{Eff}$. So, in this section we address the decision problem whether Cause is a global probability-raising cause for Eff in $\mathcal{M}$. We first note that the minimality condition (for all $\mathbf{c} \in \mathit{Cause} : \Pr^{\max}_{\mathcal{M}}((\neg \mathit{Cause}) \,\mathrm{U}\, \mathbf{c}) > 0$) can again be checked in polynomial time using standard graph algorithms. Therefore, we also assume this condition holds for all $\mathbf{c} \in \mathit{Cause}$ throughout this section. We will concentrate on an algorithm to check the global probability-raising condition **(G)** which states:

(G) For each scheduler $\mathfrak{S}$ of $\mathcal{M}$ we have

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff} \mid \Diamond \mathit{Cause}) > \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff}). \tag{GPR}$$

In parallel to checking SPR causes, deciding **(G)** leads to two major results. On one hand, we prove a coNP complexity upper bound for the decision problem regarding GPR causality, as stated in Theorem 53. For this, we rely on the minimizing MDP (Def. 35, Lem. 36) which crucially has no non-trivial end-components. This allows us to express the violation of (GPR) by a scheduler $\mathfrak{S}$ in $\mathcal{M}$ in terms of expected frequencies of state-action pairs under a

scheduler $\mathfrak{U}$ in $\mathcal{M}^{\min}_{Cause}$. With this, we can express the existence of a scheduler witnessing the violation of (G) in $\mathcal{M}_{Cause}$ by a quadratic constraint system. On the other hand, in the case where $Cause$ is not a GPR cause for Eff the decision procedure above allows us to restrict the quantification in (G) to MR scheduler of $\mathcal{M}^{\min}_{Cause}$. Tracing back the memory requirements and randomization of a scheduler violating (G) in $\mathcal{M}^{\min}_{Cause}$ through the transformation to a violating scheduler $\mathfrak{S}$ of $\mathcal{M}$ gives us the second result, stated in Theorem 56.

In the remaining section we thus assume $\mathcal{M} = \mathcal{M}^{\min}_{Cause}$ (Notation 89) and denote the considered effect set $\{TP, FN\}$ in $\mathcal{M}^{\min}_{Cause}$ with Eff by abuse of notation. Therefore, $\mathcal{M}$ has no end-components and in each state $\mathbf{c} \in Cause$ there is only a single enabled action $\alpha_{\min}$ which leads to $TP \in \mathit{Eff}$ with $\mathfrak{p}_{\min,\mathbf{c}}$ and to the terminal state $FP \notin \mathit{Eff}$ with $1 - \mathfrak{p}_{\min,\mathbf{c}}$. With these assumptions, condition (G) can be reformulated as follows:

Lemma 50 (Alternative GPR Condition). *In $\mathcal{M} = \mathcal{M}^{\min}_{Cause}$ the set $Cause$ satisfies (G) if and only if for each scheduler $\mathfrak{S}$ with $\Pr^{\mathfrak{S}}(\Diamond Cause) > 0$ the following condition holds:*

$$0 > \Pr^{\mathfrak{S}}(\Diamond FN) \cdot \Pr^{\mathfrak{S}}(\Diamond FP) - \Pr^{\mathfrak{S}}(\Diamond TN) \cdot \Pr^{\mathfrak{S}}(\Diamond TP) \tag{GPR-1}$$

Proof. We fix a scheduler $\mathfrak{S}$ of $\mathcal{M}$ with $\Pr^{\mathfrak{S}}(\Diamond Cause) > 0$ and use the following abbreviations

$$\Pr^{\mathfrak{S}}(\Diamond TP) = \mathfrak{tp}, \qquad \Pr^{\mathfrak{S}}(\Diamond FP) = \mathfrak{fp},$$
$$\Pr^{\mathfrak{S}}(\Diamond FN) = \mathfrak{fn}, \qquad \Pr^{\mathfrak{S}}(\Diamond TN) = \mathfrak{tn}.$$

By construction, the four states TP, FP, FN, TN are the only terminal states in the MEC-free MDP $\mathcal{M}$ and thus we have

$$1 = \mathfrak{tp} + \mathfrak{fp} + \mathfrak{fn} + \mathfrak{tn}. \tag{1}$$

Now (GPR) is equivalent to (GPR-1) by the following:

$$\begin{aligned}
\text{(GPR)} &\iff \Pr^{\mathfrak{S}}(\Diamond \mathit{Eff} \mid \Diamond Cause) > \Pr^{\mathfrak{S}}(\Diamond \mathit{Eff}) \\
&\iff \frac{\Pr^{\mathfrak{S}}(\Diamond \mathit{Eff} \wedge \Diamond Cause)}{\Pr^{\mathfrak{S}}(\Diamond Cause)} > \Pr^{\mathfrak{S}}(\Diamond \mathit{Eff}) \\
&\iff \frac{\mathfrak{tp}}{\mathfrak{tp} + \mathfrak{fp}} > \mathfrak{tp} + \mathfrak{fn} \\
&\iff 0 > \mathfrak{tp} + \mathfrak{fn} - \frac{\mathfrak{tp}}{\mathfrak{tp} + \mathfrak{fp}} \\
&\iff 0 > \mathfrak{tp}^2 + \mathfrak{tp} \cdot \mathfrak{fp} + \mathfrak{fn} \cdot \mathfrak{tp} + \mathfrak{fn} \cdot \mathfrak{fp} - \mathfrak{tp} \\
&\iff 0 > \mathfrak{fn} \cdot \mathfrak{fp} + \mathfrak{tp} \cdot (\mathfrak{tp} + \mathfrak{fp} + \mathfrak{fn} - 1) \\
&\overset{(1)}{\iff} 0 > \mathfrak{fn} \cdot \mathfrak{fp} - \mathfrak{tn} \cdot \mathfrak{tp} \\
&\iff 0 > \Pr^{\mathfrak{S}}(\Diamond FN) \cdot \Pr^{\mathfrak{S}}(\Diamond FP) - \Pr^{\mathfrak{S}}(\Diamond TN) \cdot \Pr^{\mathfrak{S}}(\Diamond TP) \\
&\iff \text{(GPR-1)}
\end{aligned}$$

□

Recall, in $\mathcal{M} = \mathcal{M}^{\min}_{Cause}$ a terminal state of $\mathcal{M}$ is reached almost surely after finitely many steps in expectation for all schedulers. Let T be one of the sets $\{TP\}$, $\{FN\}$, $Cause$, or a singleton $\{\mathbf{c}\}$ with $\mathbf{c} \in Cause$. Then T is visited at most once in each run of $\mathcal{M}$ and we have $\Pr^{\mathfrak{S}}(\Diamond \mathsf{T}) = \mathit{freq}_{\mathfrak{S}}(\mathsf{T})$ for each scheduler $\mathfrak{S}$. This allows us to express the violation of **(G)** in terms of a quadratic constraint system over variables for the expected frequencies of state-action pairs. For the set of state-action pairs $StAct$ of $\mathcal{M}$ we consider the following constraint system over the variables $\mathrm{x}_{\mathbf{s},\alpha}$ for each $(\mathbf{s},\alpha) \in StAct$ where we use the short form notation $\mathrm{x}_{\mathbf{s}} = \sum_{\alpha \in Act(\mathbf{s})} \mathrm{x}_{\mathbf{s},\alpha}$:

$$\mathrm{x}_{\mathbf{s},\alpha} \;\geqslant\; 0 \qquad \text{for all } (\mathbf{s},\alpha) \in StAct \tag{S1}$$

$$\mathrm{x}_{init} \;=\; 1 + \sum_{(\mathbf{t},\alpha) \in StAct} \mathrm{x}_{\mathbf{t},\alpha} \cdot \mathsf{P}(\mathbf{t},\alpha,init) \tag{S2}$$

$$\mathrm{x}_{\mathbf{s}} \;=\; \sum_{(\mathbf{t},\alpha) \in StAct} \mathrm{x}_{\mathbf{t},\alpha} \cdot \mathsf{P}(\mathbf{t},\alpha,\mathbf{s}) \qquad \text{for all } \mathbf{s} \in \mathsf{S} \setminus \{init\} \tag{S3}$$

We use results for MDPs without ECs introduced in Section 2.1.4. Given a vector $\mathrm{x} \in \mathbb{R}^{StAct}$, then x is a solution to (S1) and the balance equations (S2) and (S3) if and only if there is a (possibly history-dependent) scheduler $\mathfrak{S}$ for $\mathcal{M}$ with $\mathrm{x}_{\mathbf{s},\alpha} = \mathit{freq}_{\mathfrak{S}}(\mathbf{s},\alpha)$ for all $(\mathbf{s},\alpha) \in StAct$ if and only if there is an MR scheduler $\mathfrak{S}$ for $\mathcal{M}$ with $\mathrm{x}_{\mathbf{s},\alpha} = \mathit{freq}_{\mathfrak{S}}(\mathbf{s},\alpha)$ for all $(\mathbf{s},\alpha) \in StAct$.

The violation of (GPR-1) in Lemma 50 can be reformulated in terms of the frequency-variables as follows:

$$\mathrm{x}_{FN} \cdot \mathrm{x}_{FP} - \mathrm{x}_{TN} \cdot \mathrm{x}_{TP} \;\geqslant\; 0 \tag{non-GPR}$$

Finally, we can reformulate the condition $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond Cause) > 0$ by frequency variables where we abbreviate $\mathrm{x}_{Cause} = \sum_{\mathbf{c} \in Cause} \mathrm{x}_{\mathbf{c}}$:

$$\mathrm{x}_{Cause} > 0 \tag{S4}$$

By construction we then get the following Lemma:

Lemma 51 (Quadratic System for GPR). *The set $Cause$ is a GPR cause for Eff in $\mathcal{M}$ if and only if the constructed quadratic system of inequalities (S1)-(S4) and* (non-GPR) *does not have a solution.*

Example 52 (GPR Check). Let us examine one instance of the resulting system of inequalities for the MDP $\mathcal{M}$ given by Figure 4.8. Here we consider the effect $Eff = \{fn, tp_1, tp_2\}$ and the cause candidate $Cause = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$. We notice that $\mathcal{M}$ is not of the form of the canonical MDP $\mathcal{M}_{Cause}$. Thus, transforming $\mathcal{M}$ will result in the MDP $\mathcal{M}_{Cause}$ from Figure 4.9.

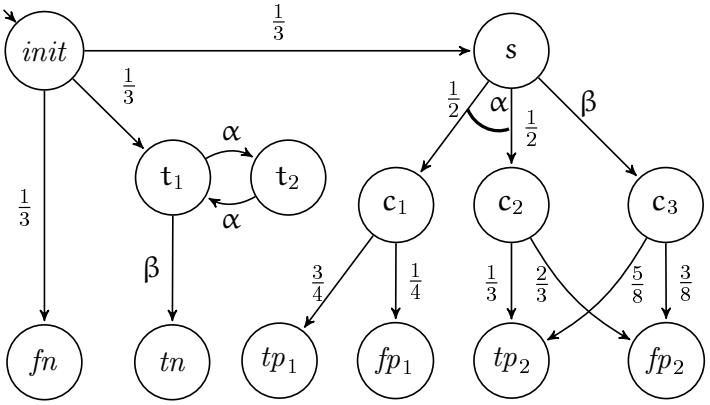

Figure 4.8: MDP $\mathcal{M}$ from Example 52

We then get the following constraint system:

$$\mathrm{x}_{\mathrm{s},\alpha} \geqslant 0 \text{ for all } (\mathrm{s},\alpha) \in \mathit{StAct} \qquad \mathrm{x}_{\mathit{init}} = 1 \qquad \mathrm{x}_{\mathit{fn}} = \frac{1}{3} \cdot \mathrm{x}_{\mathit{init}}$$

$$\mathrm{x}_{\mathcal{E},\beta} + \mathrm{x}_{\mathcal{E},\alpha} = \frac{1}{3} \cdot \mathrm{x}_{\mathit{init}} \qquad \mathrm{x}_{\mathit{tn}} = \mathrm{x}_{\mathcal{E},\beta} + \mathrm{x}_{\mathcal{E},\alpha} \qquad \mathrm{x}_{\mathrm{s},\alpha} + \mathrm{x}_{\mathrm{s},\beta} = \frac{1}{3} \cdot \mathrm{x}_{\mathit{init}}$$

$$\mathrm{x}_{\mathrm{c}_1} = \frac{1}{2} \cdot \mathrm{x}_{\mathrm{s},\alpha} \qquad \mathrm{x}_{\mathrm{c}_2} = \frac{1}{2} \cdot \mathrm{x}_{\mathrm{s},\alpha} \qquad \mathrm{x}_{\mathrm{c}_3} = \mathrm{x}_{\mathrm{s},\beta}$$

$$\mathrm{x}_{\mathit{TP}} = \frac{1}{3} \cdot \mathrm{x}_{\mathrm{c}_1} + \frac{3}{4} \cdot \mathrm{x}_{\mathrm{c}_2} + \frac{5}{8}\mathrm{x}_{\mathrm{c}_3} \qquad \mathrm{x}_{\mathit{FP}} = \frac{2}{3} \cdot \mathrm{x}_{\mathrm{c}_1} + \frac{1}{4} \cdot \mathrm{x}_{\mathrm{c}_2} + \frac{3}{8}\mathrm{x}_{\mathrm{c}_3} \qquad \mathrm{x}_{\mathrm{c}_1} + \mathrm{x}_{\mathrm{c}_2} + \mathrm{x}_{\mathrm{c}_3} > 0,$$

for which we maximize the function $\mathrm{f}(\mathrm{x}) = \mathrm{x}_{\mathit{FN}} \cdot \mathrm{x}_{\mathit{FP}} - \mathrm{x}_{\mathit{TN}} \cdot \mathrm{x}_{\mathit{TP}}$. If the maximal value $\mathrm{f}_{\max}$ of f satisfies $\mathrm{f}_{\max} \geqslant 0$ then GPR does not hold. From the constraint system we can derive

$$\mathrm{x}_{\mathit{FN}} = \frac{1}{3} \qquad \mathrm{x}_{\mathit{TN}} = \frac{1}{3}$$

$$\mathrm{x}_{\mathit{TP}} = \left(\frac{1}{6} + \frac{3}{8}\right) \cdot \mathrm{x}_{\mathrm{s},\alpha} + \frac{5}{8} \cdot \mathrm{x}_{\mathrm{s},\beta} \qquad \mathrm{x}_{\mathit{FP}} = \left(\frac{1}{3} + \frac{1}{8}\right) \cdot \mathrm{x}_{\mathrm{s},\alpha} + \frac{3}{8} \cdot \mathrm{x}_{\mathrm{s},\beta}.$$

As x_{FN} and x_{TN} are constant we minimize x_{TP} and maximize x_{FP} in order to maximize $\mathrm{f}(\mathrm{x})$. This means we set $\mathrm{x}_{\mathrm{s},\alpha} = \frac{1}{3}$ and $\mathrm{x}_{\mathrm{s},\beta} = 0$ since

$$\left(\frac{1}{6} + \frac{3}{8}\right) < \frac{5}{8} \qquad \text{and} \qquad \left(\frac{1}{3} + \frac{1}{8}\right) > \frac{3}{8}.$$

The corresponding value x' then satisfies

$$\mathrm{f}_{\max} = \mathrm{f}(\mathrm{x}') = \frac{1}{3} \cdot \frac{1}{3} \cdot \left(\frac{1}{3} + \frac{1}{8}\right) - \frac{1}{3} \cdot \frac{1}{3} \cdot \left(\frac{1}{6} + \frac{3}{8}\right) = \frac{-7}{702} < 0,$$

which means that Cause is indeed a GPR cause for Eff in $\mathcal{M}$. ◁

Theorem 53 (Complexity of Checking GPR). *Given an MDP $\mathcal{M} = (\mathrm{S}, \mathit{Act}, \mathrm{P}, \mathit{init})$, a non-empty set of terminal states $\mathit{Eff} \subset \mathrm{S}$ and a non-empty set $\mathit{Cause} \subseteq \mathrm{S} \setminus \mathit{Eff}$, deciding whether Cause is a GPR cause for Eff in $\mathcal{M}$ can be done in* coNP.

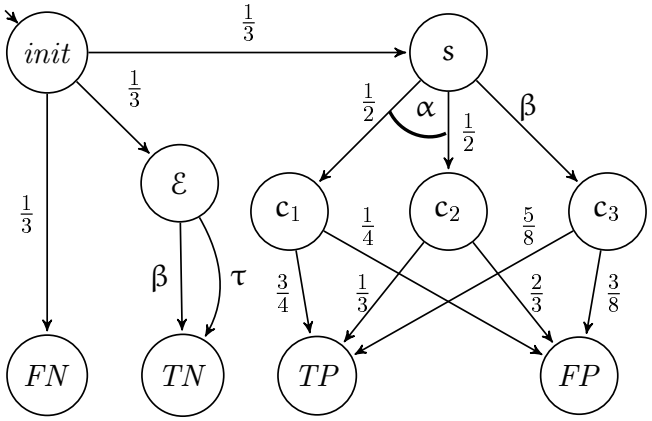

Figure 4.9: Transformed MDP $\mathcal{M}_{Cause}$

Proof. The quadratic system of inequalities can be constructed from $\mathcal{M}$, $Cause$, and Eff in polynomial time by applying the transformation $\mathcal{M} \to \mathcal{M}^{\min}_{Cause}$. Except for the strict inequality constraint in (S4), the system has the form of a quadratic program, for which the threshold problem can be decided in NP by [Vav90]. We will prove that also when including (S4), it can be checked in NP whether the system (S1)-(S4) and (non-GPR) has a solution which then corresponds to the violation of (GPR). Thus, deciding whether a set $Cause$ is a GPR cause can be done in coNP.

For this, we will provide a non-deterministic algorithm that runs in polynomial time and finds a solution to (S1)-(S4) and (non-GPR) if one exists. Some of the arguments are similar to the arguments used in [Vav90]. Additionally, we will rely on the implicit function theorem. First let us reformulate the condition (non-GPR). Since (non-GPR) comes from (GPR-1) which in turn comes from (GPR), we reformulate (GPR) into a form that uses $\Pr^{\mathfrak{S}}(\Diamond Cause)$. Given a scheduler $\mathfrak{S}$ of $\mathcal{M} = \mathcal{M}^{\min}_{Cause}$ we use the following equivalences:

$$
\begin{aligned}
& \Pr^{\mathfrak{S}}(\Diamond Eff \mid \Diamond Cause) \;>\; \Pr^{\mathfrak{S}}(\Diamond Eff) \\
\iff & \Pr^{\mathfrak{S}}(\Diamond Eff \wedge \Diamond Cause) \;>\; \Pr^{\mathfrak{S}}(\Diamond Eff) \cdot \Pr^{\mathfrak{S}}(\Diamond Cause) \\
\iff & \Pr^{\mathfrak{S}}(\Diamond TP) \;>\; \big(\Pr^{\mathfrak{S}}(\Diamond TP) + \Pr^{\mathfrak{S}}(\Diamond FN)\big) \cdot \Pr^{\mathfrak{S}}(\Diamond Cause) \\
\iff & (1 - \Pr^{\mathfrak{S}}(\Diamond Cause)) \cdot \Pr^{\mathfrak{S}}(\Diamond TP) \;>\; \Pr^{\mathfrak{S}}(\Diamond FN) \cdot \Pr^{\mathfrak{S}}(\Diamond Cause). \qquad \text{(GPR-2)}
\end{aligned}
$$

Since every cause state $\mathfrak{c} \in Cause$ can at most be reached once in every run of $\mathcal{M}$ we use $\Pr^{\mathfrak{S}}(\Diamond TP) = \sum_{\mathfrak{c} \in Cause} \Pr^{\mathfrak{S}}(\Diamond \mathfrak{c}) \cdot \mathfrak{p}_{\min,\mathfrak{c}}$ to express the violation of (GPR-2) terms of variables:

$$
\big(1 - \mathfrak{x}_{Cause}\big) \cdot \sum_{\mathfrak{c} \in Cause} \mathfrak{x}_{\mathfrak{c}} \cdot \mathfrak{p}_{\min,\mathfrak{c}} - \; \mathfrak{x}_{Cause} \cdot \mathfrak{x}_{FN} \;\leqslant\; 0 \qquad \text{(S5)}
$$

Assume that a solution to (S1)-(S5) exists. There are two possible cases:

Case 1 All solutions to (S1)-(S4) satisfy (S5). Then, in particular, the frequency values of an MD scheduler maximizing the probability to reach Cause are a solution to (S1)-(S4) and (non-GPR).

Case 2 There are solutions to (S1)-(S4) that violate (S5). Since the space of feasible points for conditions (S1)-(S4) is connected and the left hand side of (S5)

$$\left(1-\mathfrak{x}_{Cause}\right)\cdot\sum_{\mathfrak{c}\in Cause}\mathfrak{x}_{\mathfrak{c}}\cdot\mathfrak{p}_{\min,\mathfrak{c}}\ -\ \mathfrak{x}_{Cause}\cdot\mathfrak{x}_{FN}$$

is continuous there is a solution to (S1)-(S4) that satisfies

$$\left(1-\mathfrak{x}_{Cause}\right)\cdot\sum_{\mathfrak{c}\in Cause}\mathfrak{x}_{\mathfrak{c}}\cdot\mathfrak{p}_{\min,\mathfrak{c}}\ -\ \mathfrak{x}_{Cause}\cdot\mathfrak{x}_{FN}=0 \tag{S5'}$$

Now, let us take a closer look at Case 2: First of all, we add the equation

$$\mathfrak{x}_{Cause}=\sum_{\mathfrak{c}\in Cause}\mathfrak{x}_{\mathfrak{c}}. \tag{S6}$$

Thus, the variables are $\mathfrak{x}_{Cause}$ and $\mathfrak{x}_{\mathfrak{s},\alpha}$ for each $(\mathfrak{s},\alpha)\in StAct$. Obviously, this does not influence the satisfiability. We write $\mathfrak{x}$ for the vector of variables $\mathfrak{x}_{\mathfrak{s},\alpha}$ with $(\mathfrak{s},\alpha)\in StAct$.

In Case 2, there is a solution $(\mathfrak{x}^*,\mathfrak{x}^*_{Cause})$ such that the maximal possible number of variables is 0 and such that $\mathfrak{x}^*_{Cause}$ is maximal among all such solutions. Let X' be the set of variables that are 0 in $(\mathfrak{x}^*,\mathfrak{x}^*_{Cause})$. We remove all variables from X' from all constraints by setting them to 0 and call the resulting system (T1)-(T6), where (T5) is obtained from (S5'), while all other equations (T$\mathfrak{i}$) are obtained from (S$\mathfrak{i}$), by removing the chosen variables. We then collect the remaining variables in the vector $\mathfrak{v}=(\mathfrak{y},\mathfrak{y}_{Cause})$. Let $(\mathfrak{y}^*,\mathfrak{y}^*_{Cause})$ be the solution $(\mathfrak{x}^*,\mathfrak{x}^*_{Cause})$ after the variables in X' have been removed. Thus, all values in this vector are positive.

Define the function $\mathfrak{f}$ as the left hand side of (T5):

$$\mathfrak{f}(\mathfrak{y},\mathfrak{y}_{Cause})=\left(1-\mathfrak{y}_{Cause}\right)\cdot\sum_{\mathfrak{c}\in Cause}\mathfrak{y}_{\mathfrak{c}}\cdot\mathfrak{p}_{\min,\mathfrak{c}}\ -\ \mathfrak{y}_{Cause}\cdot\mathfrak{y}_{FN}$$

where the variables $\mathfrak{y}$ are as the original variables $\mathfrak{x}$ after the variables in X' have been removed. Now, we apply the implicit function theorem: Observe that

$$\frac{\partial\mathfrak{f}(\mathfrak{y},\mathfrak{y}_{Cause})}{\partial\mathfrak{y}_{Cause}}=-\sum_{\mathfrak{c}\in Cause}\mathfrak{y}_{\mathfrak{c}}\cdot\mathfrak{p}_{\min,\mathfrak{c}}-\mathfrak{y}_{FN}.$$

Evaluated at $(\mathfrak{y}^*,\mathfrak{y}^*_{Cause})$, this value is non-zero as all summands are negative and there are at least some of the variables in the abbreviation $\mathfrak{y}_{\mathfrak{c}}$ with $\mathfrak{c}\in Cause$ left, i.e., not removed because they were not 0 due to the original constraint (S5). So, we can apply the implicit function theorem, which guarantees us the existence of a function $\mathfrak{g}(\mathfrak{y})$ with $\mathfrak{g}(\mathfrak{y}^*)=\mathfrak{y}^*_{Cause}$ and for all $\mathfrak{y}'$ in an open ball B_1 around $\mathfrak{y}^*$, we have

$$\mathfrak{f}(\mathfrak{y}',\mathfrak{g}(\mathfrak{y}'))=0.$$

By the implicit function theorem, we can explicitly compute the gradient

$$\nabla g = \left(\frac{\partial g(y)}{\partial y_1}, \ldots, \frac{\partial g(y)}{\partial y_k}\right) = -\left(\frac{\partial f(y, y_{Cause})}{\partial y_{Cause}}\right)^{-1} \cdot \left(\frac{\partial f(y, y_{Cause})}{\partial y_1}, \ldots, \frac{\partial f(y, y_{Cause})}{\partial y_k}\right)$$

of the derivatives on B_1 for the appropriate k from the derivatives of f. Note that on B_1, the gradient ∇g is 0 iff

$$H(y, y_{Cause}) = \left(\frac{\partial f(y, y_{Cause})}{\partial y_1}, \ldots, \frac{\partial f(y, y_{Cause})}{\partial y_k}\right)$$

is 0. Furthermore, all entries of $H(y, y_{Cause})$ are linear in the variables v as the function f is quadratic. As the function g has a local maximum in y^*, we know that $\nabla g(y^*) = 0$.

Equations (T2), (T3), and (T6) are linear equations in the remaining variables v. We can rewrite these three equations with a matrix M and a vector b whose entries can easily be expressed in terms of the coefficients of the original system (again, after the set of variables X' has been removed) as $Mv = b$. The solutions to this equation form an r-dimensional affine space W. It can be written as

$$W = \{\underbrace{c_0 + c_1 \cdot z_1 + \cdots + c_r \cdot z_r}_{\stackrel{\text{def}}{=} h(z_1, \ldots, z_r)} \mid (z_1, \ldots, z_r) \in \mathbb{R}^r\}$$

for some vectors $c_0, c_1, \ldots, c_r$ which can be computed from M and b in polynomial time.

Let B_2 be an open ball in $\mathbb{R}^r$ such that $h(B_2) \subseteq B_1$ and such that $h(B_2)$ contains (y^*, y^*_{Cause}). We claim that $g \circ h \colon B_2 \to \mathbb{R}$ has an isolated local maximum at $z^* = h^{-1}(y^*, y^*_{Cause})$. It is clear that $g \circ h$ has a local maximum since g has a local maximum at (y^*, y^*_{Cause}). Suppose now, that $g \circ h$ does not have an isolated local maximum at $h^{-1}(y^*, y^*_{Cause})$. As h is an affine map and the graph of g is the solution to a quadratic equation, this is only possible if there is a direction $d \in \mathbb{R}^r \setminus \{0\}$ such that

$$g \circ h(z^*) = g \circ h(z^* + t \cdot d) \tag{4.6}$$

for all $t \in \mathbb{R}$. Due to the boundedness of the polyhedron described by conditions (T1)-(T4) and (T6) and since z^* lies in the interior of this polyhedron, this means that there must be a value $q \in \mathbb{R}$ such that $(h(z^* + q \cdot d), g \circ h(z^* + q \cdot d))$ provides a solution v to equations (T1)-(T4) and (T6) with an additional 0. By the definition of g, this solution furthermore satisfies (T5) and, by equation (4.6), it still satisfies (T4). This contradicts the choice of the original solution (x^*, x^*_{Cause}).

So, $g \circ h \colon B_2 \to \mathbb{R}$ has an isolated local maximum at z^*. This implies that on an open ball around z^*, the point z^* is the only solution to $\nabla g(h(z)) = 0$ and consequently, the only solution to $H(h(z)) = 0$.

Since $H(h(z))$ is a vector of linear expressions in z, this implies that z^* is the only solution on $\mathbb{R}^r$ to $H(h(z)) = 0$. This is the key result that we need to provide a non-deterministic polynomial-time algorithm to check the satisfiability of the original constraint system.

We now describe the algorithm: The algorithm begins by computing the frequency values of an MD scheduler as in Case 1 in polynomial time and checks whether the resulting vector of frequency values satisfies (S1)-(S4) and (non-GPR). If this is the case, the algorithm returns that the system is satisfiable.

If this is not the case, the algorithm tries to compute a solution to (S1)-(S4), and (S5') as in Case 2. The algorithm non-deterministically guesses a subset of the variables and removes them from all constraints by replacing them with 0.

Suppose, we guess the set $\mathbf{X}'$ as above. We show how we compute a solution then. After the variables from $\mathbf{X}'$ have been removed, $\mathbf{H}(\mathfrak{y}, \mathfrak{y}_{Cause})$ can be computed in polynomial time as all the derivatives of $\mathbf{f}$ are linear expressions in the variables which require basic arithmetic and can be computed in polynomial time. Likewise, $\mathbf{M}$ and $\mathfrak{b}$ can be computed in polynomial time from the original constraints after the guessed variables have been removed. The vectors $\mathfrak{c}_0, \mathfrak{c}_1, \ldots, \mathfrak{c}_\mathfrak{r}$ describing the solution space to $\mathbf{M}\nu = \mathfrak{b}$ can then also be computed in polynomial time.

Thus, also the vector $\mathbf{H}(\mathfrak{h}(z))$ of linear expressions in the variables z can be computed in polynomial time. The equation system $\mathbf{H}(\mathfrak{h}(z)) = 0$ has a unique solution if the guessed variables were indeed $\mathbf{X}'$. In this case, the solution z^* can be computed in polynomial time as well. If the guess of variables was not $\mathbf{X}'$, then either there is no unique solution to this equation system which can be detected in polynomial time, or the solution, which is computed in the sequel in polynomial time, might not satisfy the original constraints, which is checked in the end.

From z^*, we can compute $\mathfrak{y}^* = \mathfrak{h}(z^*)$ using the vectors $\mathfrak{c}_0, \mathfrak{c}_1, \ldots, \mathfrak{c}_\mathfrak{r}$. The solution $\mathfrak{x}^*$ is then obtained by plugging in 0s for the removed variables. Checking whether the resulting vector satisfies all constraints can also be done in polynomial time in the end. If $\mathbf{X}'$ was guessed correctly, this vector $\mathfrak{x}^*$ indeed forms a solution to the original constraints as we have seen.

In summary, the algorithm needs to guess the set $\mathbf{X}'$ of variables which are 0 in a solution to the original constraints with the maximal number of zeroes. All other steps are deterministic polynomial-time computations. Thus, satisfiability of (S1)-(S4) and (non-GPR) can be checked in NP. □

Memory Requirements of Schedulers in the Original MDP. We now address the memory requirements of a scheduler which refutes **(G)** if it exists. As stated before, every solution to the linear system of inequalities (S1), (S2), and (S3) corresponds to the expected frequencies of state-action pairs of an MR scheduler in $\mathfrak{M}^{\min}_{Cause}$. Hence:

Corollary 54 (MR-Scheduler refuting GPR). *For $\mathfrak{M} = \mathfrak{M}^{\min}_{Cause'}$, the set $Cause$ is not an GPR cause for Eff iff there exists an MR scheduler $\mathfrak{T}$ with $\Pr^{\mathfrak{T}}_{\mathfrak{M}}(\Diamond Cause) > 0$ violating* (GPR).

Here again the model transformation $\mathcal{M} \to \mathcal{M}^{\min}_{Cause}$ does affect the memory requirements of scheduler. In order to use Theorem 39 to lift a scheduler $\mathfrak{U}$ witnessing the violation of (GPR) to $\mathcal{M}$ we further restrict the MR schedulers necessary to witness non-causality in $\mathcal{M}^{\min}_{Cause}$. For the following lemma, recall that τ is the action of the MEC quotient used for the extra transition from states $s_{\mathcal{E}}$ representing MECs to a new trap state TN.

Lemma 55 (Deterministically traversing ECs). *Assume $\mathcal{M} = \mathcal{M}^{\min}_{Cause}$. Given an MR scheduler $\mathfrak{U}$ with $\Pr^{\mathfrak{U}}_{\mathcal{M}}(\Diamond Cause) > 0$ that violates* (GPR)*, an MR-scheduler $\mathfrak{V}$ with $\mathfrak{V}(s)(\tau) \in \{0,1\}$ for each state $s \in S$ with $\tau \in Act(s)$ that satifies $\Pr^{\mathfrak{V}}_{\mathcal{M}}(\Diamond Cause) > 0$ and violates* (GPR) *is computable in polynomial time.*

Proof. Let $\mathfrak{U}$ be a scheduler with $\Pr^{\mathfrak{U}}_{\mathcal{M}}(\Diamond Cause) > 0$ violating (GPR). As in the proof of Theorem 53 $\mathfrak{U}$ this is the case if and only if $\mathfrak{U}$ violates (GPR-2). Thus

$$(1 - \Pr^{\mathfrak{S}}(\Diamond Cause)) \cdot \Pr^{\mathfrak{S}}(\Diamond TP) \;\leqslant\; \Pr^{\mathfrak{S}}(\Diamond FN) \cdot \Pr^{\mathfrak{S}}(\Diamond Cause).$$

We will show how to transform $\mathfrak{U}$ into an MR scheduler $\mathfrak{V}$ that schedules the τ-transitions to TN with probability 0 or 1. We regard the set U of states u which have a τ-transition to TN (recall that then $P(u, \tau, TN) = 1$) and where $0 < \mathfrak{U}(u, \tau) < 1$. We now process the U-states in an arbitrary order, say $u_1, \ldots, u_k$, and generate a sequence $\mathfrak{V}_0 = \mathfrak{U}, \mathfrak{V}_1, \ldots, \mathfrak{V}_k = \mathfrak{V}$ of MR schedulers such that for $i \in \{1, \ldots, k\}$:

- $\mathfrak{V}_i$ refutes (GPR) (or equivalently condition (GPR-2)),
- $\mathfrak{V}_i$ agrees with $\mathfrak{V}_{i-1}$ for all states but u_i,
- $\mathfrak{V}_i(u_i)(\tau) \in \{0,1\}$.

Thus, the final scheduler $\mathfrak{V}_k = \mathfrak{V}$ satisfies the desired properties.

To explain how to derive $\mathfrak{V}_i$ from $\mathfrak{V}_{i-1}$, let $i \in \{1, \ldots, k\}$, $\mathfrak{T} = \mathfrak{V}_{i-1}$, $u = u_i$ and $y = 1 - \mathfrak{T}(u)(\tau)$. Then, $0 < y < 1$ (as $u \in U$ and by definition of U) and $y = \sum_{\alpha \in Act(u) \setminus \{\tau\}} \mathfrak{T}(u)(\alpha)$. For $x \in [0,1]$, let $\mathfrak{T}_x$ denote the MR scheduler that agrees with $\mathfrak{T}$ for all states but u, for which $\mathfrak{T}_x$'s decision is:

$$\mathfrak{T}_x(u)(\tau) = 1 - x, \qquad \mathfrak{T}_x(u)(\alpha) = \mathfrak{T}(u)(\alpha) \cdot \frac{x}{y} \quad \text{for } \alpha \in Act(u) \setminus \{\tau\}$$

Obviously, $\mathfrak{T}_y = \mathfrak{T}$. We now show that at least one of the two MR schedulers $\mathfrak{T}_0$ or $\mathfrak{T}_1$ also refutes (GPR). For this, we suppose by contraction that this is not the case, which means that (GPR) holds for both. Let $f : [0,1] \to \mathbb{Q}$ be defined by

$$f(x) \;=\; (1 - \Pr^{\mathfrak{T}_x}(\Diamond Cause)) \cdot \Pr^{\mathfrak{T}_x}(\Diamond TP) - \Pr^{\mathfrak{T}_x}(\Diamond FN) \cdot \Pr^{\mathfrak{T}_x}(\Diamond Cause)$$

As $\mathfrak{T} = \mathfrak{T}_y$ violates (GPR-2), while $\mathfrak{T}_0$ and $\mathfrak{T}_1$ satisfy (GPR-2) we obtain:

$$f(0), f(1) < 0 \qquad \text{and} \qquad f(y) \geqslant 0$$

We now split Cause into the set C of states $\mathfrak{c} \in \mathit{Cause}$ such that there is a $\mathfrak{T}$-path from init to $\mathfrak{c}$ that traverses $\mathfrak{u}$ and $\mathrm{D} = \mathit{Cause} \setminus \mathrm{C}$. By defining

$$\mathfrak{a}_{\mathfrak{x}} = \Pr^{\mathfrak{T}_{\mathfrak{x}}}(\Diamond \mathrm{C}), \qquad \mathfrak{a} = \Pr^{\mathfrak{T}}(\Diamond \mathrm{D}),$$
$$\mathfrak{b}_{\mathfrak{x}} = \Pr^{\mathfrak{T}_{\mathfrak{x}}}(\Diamond \mathfrak{u} \wedge \Diamond \mathit{FN}), \qquad \mathfrak{b} = \Pr^{\mathfrak{T}_{\mathfrak{x}}}(\neg\Diamond \mathfrak{u} \wedge \Diamond \mathit{FN}),$$

we get $\Pr^{\mathfrak{T}_{\mathfrak{x}}}_{\mathcal{M}}(\Diamond \mathit{Cause}) = \mathfrak{a}_{\mathfrak{x}} + \mathfrak{a}$ and $\Pr^{\mathfrak{T}_{\mathfrak{x}}}_{\mathcal{M}}(\Diamond \mathit{FN}) = \mathfrak{b}_{\mathfrak{x}} + \mathfrak{b}$. With $\mathfrak{a}_{\mathfrak{x},\mathfrak{c}} = \Pr^{\mathfrak{V}_{\mathfrak{x}}}_{\mathcal{M}}(\Diamond \mathfrak{c})$ for $\mathfrak{c} \in \mathrm{C}$ and $\mathfrak{a}_{\mathfrak{d}} = \Pr^{\mathfrak{V}}_{\mathcal{M}}(\Diamond \mathfrak{d})$ for $\mathfrak{d} \in \mathrm{D}$, let

$$\mathfrak{c}_{\mathfrak{x}} = \sum_{\mathfrak{c} \in \mathrm{C}} \mathfrak{p}_{\mathfrak{x},\mathfrak{c}} \cdot w_{\mathfrak{c}} \qquad \text{and} \qquad \mathfrak{c} = \sum_{\mathfrak{d} \in \mathrm{D}} \mathfrak{p}_{\mathfrak{d}} \cdot w_{\mathfrak{d}},$$

where $w_{\mathfrak{c}} = \Pr^{\min}_{\mathcal{M},\mathfrak{c}}(\Diamond \mathit{Eff})$ for each $\mathfrak{c} \in \mathit{Cause}$. As $\mathfrak{y}$ is fixed, the values $\mathfrak{a}_{\mathfrak{y}}, \mathfrak{a}_{\mathfrak{y},\mathfrak{c}}, \mathfrak{b}_{\mathfrak{y}}, \mathfrak{b}_{\mathfrak{y}}$ can be seen as constants. Moreover, $\mathfrak{a}_{\mathfrak{x}}, \mathfrak{a}_{\mathfrak{x},\mathfrak{c}}, \mathfrak{b}_{\mathfrak{x}}, \mathfrak{c}_{\mathfrak{x}}$ differ from $\mathfrak{a}_{\mathfrak{y}}, \mathfrak{a}_{\mathfrak{y},\mathfrak{c}}, \mathfrak{b}_{\mathfrak{y}}, \mathfrak{c}_{\mathfrak{y}}$ only by factor $\frac{\mathfrak{x}}{\mathfrak{y}}$. So:

$$\mathfrak{a}_{\mathfrak{x}} = \mathfrak{a}_{\mathfrak{y}}\frac{\mathfrak{x}}{\mathfrak{y}}, \qquad \mathfrak{a}_{\mathfrak{x},\mathfrak{c}} = \mathfrak{a}_{\mathfrak{y},\mathfrak{c}}\frac{\mathfrak{x}}{\mathfrak{y}}, \qquad \mathfrak{b}_{\mathfrak{x}} = \mathfrak{b}_{\mathfrak{y}}\frac{\mathfrak{x}}{\mathfrak{y}} \qquad \mathfrak{c}_{\mathfrak{x}} = \mathfrak{c}_{\mathfrak{y}}\frac{\mathfrak{x}}{\mathfrak{y}}.$$

Thus, $\mathfrak{f}(\mathfrak{x})$ has the following form:

$$\begin{aligned}
\mathfrak{f}(\mathfrak{x}) &= (\mathfrak{a}_{\mathfrak{x}}+\mathfrak{a})(\mathfrak{b}_{\mathfrak{x}}+\mathfrak{b}) - \big(1-(\mathfrak{a}_{\mathfrak{x}}+\mathfrak{a})\big)(\mathfrak{c}_{\mathfrak{x}}+\mathfrak{c}) \\
&= \underbrace{\mathfrak{a}_{\mathfrak{x}}\mathfrak{b}_{\mathfrak{x}}+\mathfrak{a}_{\mathfrak{x}}\mathfrak{c}_{\mathfrak{x}}}_{\alpha\mathfrak{x}^2} + \underbrace{\mathfrak{a}_{\mathfrak{x}}(\mathfrak{b}+\mathfrak{c}) + \mathfrak{b}_{\mathfrak{x}}\mathfrak{a} - \mathfrak{c}_{\mathfrak{x}}}_{\beta\mathfrak{x}} + \underbrace{\mathfrak{a}\mathfrak{b} - \mathfrak{c} + \mathfrak{a}\mathfrak{c}}_{\gamma} \\
&= \alpha\mathfrak{x}^2 + \beta\mathfrak{x} + \gamma
\end{aligned}$$

For the value α, we have $\alpha\mathfrak{x}^2 = \mathfrak{a}_{\mathfrak{x}}\mathfrak{b}_{\mathfrak{x}}+\mathfrak{a}_{\mathfrak{x}}\mathfrak{c}_{\mathfrak{x}}$ and hence $\alpha = \frac{1}{\mathfrak{y}^2}(\mathfrak{a}_{\mathfrak{y}}\mathfrak{b}_{\mathfrak{y}} + \mathfrak{a}_{\mathfrak{y}}\mathfrak{c}_{\mathfrak{y}}) > 0$. Thus, the second derivative $\mathfrak{f}''(\mathfrak{x}) = 2\alpha$ of $\mathfrak{f}$ is positive, which yields that $\mathfrak{f}$ has a global minimum at some point $\mathfrak{x}_0$ and is strictly decreasing for $\mathfrak{x} < \mathfrak{x}_0$ and strictly increasing for $\mathfrak{x} > \mathfrak{x}_0$. As $\mathfrak{f}(0)$ and $\mathfrak{f}(1)$ are both negative, we obtain $\mathfrak{f}(\mathfrak{x}) < 0$ for all $\mathfrak{x} \in [0,1]$, which contradicts $\mathfrak{f}(\mathfrak{y}) \geqslant 0$.

This concludes that at least one of the schedulers $\mathfrak{T}_0$ or $\mathfrak{T}_1$ witnesses the violation of (GPR). Thus, we can define $\mathfrak{V}_{\mathfrak{i}} \in \{\mathfrak{T}_0, \mathfrak{T}_1\}$ accordingly.

The number of states $\mathfrak{k}$ in $\mathfrak{U}$ is bounded by the number of states in $\mathfrak{S}$. In each iteration of the above construction, the function value $\mathfrak{f}(0)$ is sufficient to determine one of the schedulers $\mathfrak{T}_0$ and $\mathfrak{T}_1$ witnessing the violation of (GPR). So, the procedure has to compute the values in condition (GPR-1) for $\mathfrak{k}$-many MR schedulers and update the scheduler afterwards. As the update can easily be carried out in polynomial time, the run-time of all $\mathfrak{k}$ iterations is polynomial as well. □

So, the constructed scheduler $\mathfrak{V}$ of Lemma 55 and hence the scheduler with two memory modes whose existence is stated in Theorem 56 can be computed from a solution to the constraint system (S1)-(S4) and (non-GPR) from Section 4.1.3 in time polynomial in the size of the original MDP and the representation of the solution to (S1)-(S4) and (non-GPR).

With these results we can now prove the second main result of this section, Theorem 56, stating that if (GPR) does not hold there is a finite-memory scheduler with two memory cells refuting the GPR condition. For the following result we do *not* assume $\mathcal{M} = \mathcal{M}^{\min}_{Cause}$.

Theorem 56 (Requirements on Scheduler refuting GPR). *Let $\mathcal{M}$ be an MDP with a non-empty set of terminal states Eff and a non-empty set $\mathit{Cause} \subseteq S \setminus \mathit{Eff}$ which satisfies for all $\mathbf{c} \in \mathit{Cause}$: $\Pr^{\max}_{\mathcal{M}}((\neg \mathit{Cause}) \,\mathrm{U}\, \mathbf{c}) > 0$. If Cause is not a GPR cause for Eff in $\mathcal{M}$, then there is an MR scheduler for $\mathcal{M}^{\min}_{\mathit{Cause}}$ refuting the GPR condition **(G)** for Cause in $\mathcal{M}^{\min}_{\mathit{Cause}}$ and a finite-memory randomized scheduler $\mathfrak{S}$ for $\mathcal{M}$ with two memory cells refuting the GPR condition **(G)** for Cause in $\mathcal{M}$.*

Proof. By Corollary 54 Lemma 55 we conclude that Cause is not a GPR cause for Eff in $\mathcal{M}$ if and only if there is a MR-scheduler $\mathfrak{T}$ for $\mathcal{N} = \mathcal{M}^{\min}_{\mathit{Cause}}$ with $\Pr^{\mathfrak{T}}_{\mathcal{N}}(\Diamond \mathit{Cause}) > 0$ that violates (GPR). Moreover, $\mathfrak{T}$ chooses τ with probability 0 or 1. Therefore, by Theorem 39 we can lift $\mathfrak{T}$ to $\mathcal{M}$ to construct a scheduler $\mathfrak{S}$ of $\mathcal{M}$ satisfying the desired properties. □

Remark 57 (On Lower Bounds of GPR Checking). Solving systems of quadratic inequalities with linear side constraints is NP-hard in general (see, e.g., [GJ79]). For convex problems, in which the associated symmetric matrix occurring in the quadratic inequality has only non-negative eigenvalues, the problem is, however, solvable in polynomial time [KTK80]. Unfortunately, in general the quadratic constraint system describing a scheduler refuting (GPR) given by (S1)-(S4) and (non-GPR) is not of this form. We observe that even if for a simple system where only the values $\mathbf{x}_{FN}, \mathbf{x}_{FP}, \mathbf{x}_{TN}, \mathbf{x}_{TP}$ are variables by structure of the MDP, the symmetric matrix for (non-GPR) takes the form:

$$\mathrm{Q} = \frac{1}{2}\begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}, \tag{4.7}$$

such that for $\mathbf{x} = (\mathbf{x}_{TN}, \mathbf{x}_{FP}, \mathbf{x}_{FN}, \mathbf{x}_{TP})$ we have $\mathbf{x}^{\mathsf{T}} \cdot \mathrm{Q} \cdot \mathbf{x} = \mathbf{x}_{FN} \cdot \mathbf{x}_{FP} - \mathbf{x}_{TN} \cdot \mathbf{x}_{TP}$. Therefore, Q has -1 as a negative eigenvalue. Although it is not ruled out that the quadratic system (S1)-(S4) and (non-GPR) belongs to another class of efficiently solvable constraint systems, the NP-hardness result in [PV91] for the solvability of quadratic inequalities with linear side constraints where the quadratic matrix only has one negative eigenvalue might be an indication for the computational difficulty. ◁

To conclude the algorithms for checking the PR conditions, we have seen that checking the strict PR condition **(S)** is considerably simpler than checking the global PR condition **(G)** . This result is expected since by Proposition 15 there is an implication from SPR to GPR causality.

Checking SPR causality is done directly on a transformed MDP derived from $\mathcal{M}$ by comparing explicit probabilities. Since the MDP transformation and the computation of the corresponding probabilities can be done in polynomial time, the SPR check is in P.

Checking the GPR condition is more complex as it requires solving a quadratic program which is not necessarily convex. Moreover, we do not have lower bounds for checking the GPR condition. Although Remark 57 argues for an indication for the computational complexity of checking GPR causes, this is still dependent on the quadratic programming approach to the problem. Thus, there might be completely different approaches which may improve the complexity of checking GPR. Nonetheless, we will see in the later Section 4.3.2 that GPR causes do have an underlying hardness when we address the problem of computing quality-optimal GPR causes.

4.2 Computing the Quality of a PR Cause

Before addressing the algorithmic complexities of computing optimal causes for the respective quality measures, we will give algorithms on how to compute a quality measure for a given PR cause,

Throughout this section we are given an MDP $\mathcal{M} = (\mathrm{S}, Act, \mathrm{P}, init)$, a non-empty set of terminal effect states $\mathit{Eff} \subseteq \mathrm{S}$ and we assume a fixed PR cause of any kind $\mathit{Cause} \subseteq \mathrm{S} \setminus \mathit{Eff}$ is given. Furthermore, we suppose all states $s \in \mathrm{S}$ are reachable from $init$.

Our first observation is that all quality measures are preserved by the transformation from $\mathcal{M}$ to $\mathcal{M}^{\min}_{\mathit{Cause}}$ of Definition 35. This holds, since we defined quality values as a worst-case analysis by taking the infimum over all scheduler for which the quality measure exists. Intuitively, such scheduler will always minimize the probability of reaching Eff starting in Cause anyway. Therefore, we first show that the quality measures $\mathit{recall}, \mathit{covrat}$ and fscore of a fixed Cause are compatible with the model transformation $\mathcal{M} \to \mathcal{M}^{\min}_{\mathit{Cause}}$. Recall that these transformations have three stages:

1. $\mathcal{M} \to \mathcal{M}_{<\mathit{Cause}>}$ where all outgoing transitions from states $\mathrm{c} \in \mathit{Cause}$ are deleted and instead two actions $\alpha_{\min}$ and $\alpha_{\max}$ are added which transition to fresh states TP and FP with

$$\mathrm{P}_{\mathcal{M}_{<\mathit{Cause}>}}(\mathrm{c}, \alpha_{\min}, TP) = \mathrm{p}_{\mathrm{c},\min}, \qquad \mathrm{P}_{\mathcal{M}_{<\mathit{Cause}>}}(\mathrm{c}, \alpha_{\min}, FP) = 1 - \mathrm{p}_{\mathrm{c},\min},$$
$$\mathrm{P}_{\mathcal{M}_{<\mathit{Cause}>}}(\mathrm{c}, \alpha_{\max}, TP) = \mathrm{p}_{\mathrm{c},\max}, \qquad \mathrm{P}_{\mathcal{M}_{<\mathit{Cause}>}}(\mathrm{c}, \alpha_{\max}, FP) = 1 - \mathrm{p}_{\mathrm{c},\max}.$$

2. $\mathcal{M}_{<\mathit{Cause}>} \to \mathcal{M}_{\mathit{Cause}}$ where the end-components are collapsed by using the MEC-quotient. Additionally, terminal states are collapsed to either FN (effect states) or TN (non-effect states).

3. $\mathcal{M}_{\mathit{Cause}} \to \mathcal{M}^{\min}_{\mathit{Cause}}$ where the action $\alpha_{\max}$ in states $\mathrm{c} \in \mathit{Cause}$ is deleted.

For the first two transformations we already proved that these preserve the values of the confusion matrix by Lemma 32, which are exactly the values for which quality measures are defined.

Corollary 58 (Canonical MDP preserves Quality Measures). *If $Cause$ is an SPR or a GPR cause for Eff in $\mathfrak{M}$ then we have:*

$$\begin{aligned} recall_{\mathfrak{M}}(Cause) &= recall_{\mathfrak{M}_{Cause}}(Cause) \\ covrat_{\mathfrak{M}}(Cause) &= covrat_{\mathfrak{M}_{Cause}}(Cause) \\ fscore_{\mathfrak{M}}(Cause) &= fscore_{\mathfrak{M}_{Cause}}(Cause) \end{aligned}$$

Proof. By Lemma 32 for each scheduler $\mathfrak{S}$ of $\mathfrak{M}$ there is a scheduler $\mathfrak{U}$ of $\mathfrak{M}_{Cause}$ and vice versa such that

$$\begin{aligned} tp^{\mathfrak{S}}_{\mathfrak{M}} &= tp^{\mathfrak{U}}_{\mathfrak{M}_{Cause}}, & fp^{\mathfrak{S}}_{\mathfrak{M}} &= fp^{\mathfrak{U}}_{\mathfrak{M}_{Cause}}, \\ fn^{\mathfrak{S}}_{\mathfrak{M}} &= fn^{\mathfrak{U}}_{\mathfrak{M}_{Cause}}, & tn^{\mathfrak{S}}_{\mathfrak{M}} &= tn^{\mathfrak{U}}_{\mathfrak{M}_{Cause}}. \end{aligned}$$

As all quality measures are defined in terms of these values, we also have

$$\begin{aligned} recall^{\mathfrak{S}}_{\mathfrak{M}}(Cause) &= recall^{\mathfrak{U}}_{\mathfrak{M}_{Cause}}(Cause) \\ covrat^{\mathfrak{S}}_{\mathfrak{M}}(Cause) &= covrat^{\mathfrak{U}}_{\mathfrak{M}_{Cause}}(Cause) \\ fscore^{\mathfrak{S}}_{\mathfrak{M}}(Cause) &= fscore^{\mathfrak{U}}_{\mathfrak{M}_{Cause}}(Cause) \end{aligned}$$

for the corresponding schedulers $\mathfrak{S}$ and $\mathfrak{U}$. Therefore, the infimum over these values also stays the same and the claim holds. □

We are left to show that the quality measures are also preserved when we omit the maximal action $\alpha_{\max}$. Intuitively, this holds since we define $recall$, $covrat$ and $fscore$ as infima and thus scheduler realizing these values would only take action $\alpha_{\min}$ anyway.

Lemma 59 (Minimizing MDP preserves Quality Measures). *If $Cause$ is an SPR or GPR cause for Eff in $\mathfrak{M}$ then we have*

$$\begin{aligned} recall_{\mathfrak{M}}(Cause) &= recall_{\mathfrak{M}^{\min}_{Cause}}(Cause) \\ covrat_{\mathfrak{M}}(Cause) &= covrat_{\mathfrak{M}^{\min}_{Cause}}(Cause) \\ fscore_{\mathfrak{M}}(Cause) &= fscore_{\mathfrak{M}^{\min}_{Cause}}(Cause). \end{aligned}$$

Proof. By Corollary 58 we can assume $\mathfrak{M} = \mathfrak{M}_{Cause}$. We show the equality by proving inequality in both directions.

"$\leqslant$": A scheduler $\mathfrak{U}$ of $\mathfrak{M}^{\min}_{Cause}$ corresponds to a scheduler $\mathfrak{S}$ for $\mathfrak{M}$ which behaves as a memoryless deterministic scheduler and which minimizes the reachability probability of Eff from every state in $Cause$ and we have:

$$\begin{aligned} recall^{\mathfrak{S}}_{\mathfrak{M}}(Cause) &= recall^{\mathfrak{U}}_{\mathfrak{M}^{\min}_{Cause}}(Cause) \\ covrat^{\mathfrak{S}}_{\mathfrak{M}}(Cause) &= covrat^{\mathfrak{U}}_{\mathfrak{M}^{\min}_{Cause}}(Cause) \\ fscore^{\mathfrak{S}}_{\mathfrak{M}}(Cause) &= fscore^{\mathfrak{S}}_{\mathfrak{M}^{\min}_{Cause}}(Cause). \end{aligned}$$

We especially obtain $\mathit{recall}_{\mathcal{M}}(\mathit{Cause}) \leqslant \mathit{recall}_{\mathcal{M}^{\min}_{\mathit{Cause}}}(\mathit{Cause})$ and the analogous statements for covrat and fscore.

"$\geqslant$": Let $\mathfrak{S}$ be a scheduler of $\mathcal{M}$ and let $\mathfrak{U}$ be the corresponding scheduler of $\mathcal{M}^{\min}_{\mathit{Cause}}$ which behaves like $\mathfrak{S}$ up until reaching Cause. Recall and coverage under $\mathfrak{S}$ and $\mathfrak{U}$ have the form:

$$\mathit{recall}^{\mathfrak{S}}_{\mathcal{M}}(\mathit{Cause}) = \frac{x}{x+q} \qquad \mathit{covrat}^{\mathfrak{S}}_{\mathcal{M}}(\mathit{Cause}) = \frac{x}{q}$$
$$\mathit{recall}^{\mathfrak{T}}_{\mathcal{M}}(\mathit{Cause}) = \frac{y}{y+q} \qquad \mathit{covrat}^{\mathfrak{T}}_{\mathcal{M}}(\mathit{Cause}) = \frac{y}{q}$$

where $x \geqslant y$ and $q = \mathsf{fn}^{\mathfrak{S}}$ is independent of the choices in Cause. Thus, we get

$$\mathit{recall}^{\mathfrak{S}}_{\mathcal{M}}(\mathit{Cause}) \geqslant \mathit{recall}^{\mathfrak{U}}_{\mathcal{M}^{\min}_{\mathit{Cause}}}(\mathit{Cause}) \quad \text{and} \quad \mathit{covrat}^{\mathfrak{S}}_{\mathcal{M}}(\mathit{Cause}) \geqslant \mathit{covrat}^{\mathfrak{U}}_{\mathcal{M}^{\min}_{\mathit{Cause}}}(\mathit{Cause}).$$

As recall and covrat are infima over $\mathit{recall}^{\mathfrak{S}}$ and $\mathit{covrat}^{\mathfrak{S}}$ for all scheduler $\mathfrak{S}$ such that these values are defined, these inequalities continue to

$$\mathit{recall}_{\mathcal{M}}(\mathit{Cause}) \geqslant \mathit{recall}_{\mathcal{M}^{\min}_{\mathit{Cause}}}(\mathit{Cause}) \quad \text{and} \quad \mathit{covrat}_{\mathcal{M}}(\mathit{Cause}) \geqslant \mathit{covrat}_{\mathcal{M}^{\min}_{\mathit{Cause}}}(\mathit{Cause}).$$

Using the same argumentation we get

$$\mathit{precision}^{\mathfrak{S}}_{\mathcal{M}}(\mathit{Cause}) = \mathit{precision}^{\mathfrak{U}}_{\mathcal{M}^{\min}_{\mathit{Cause}}}(\mathit{Cause}).$$

The harmonic mean viewed as a function $f : \mathbb{R}^2_{>0} \to \mathbb{R}$, $f(x,y) = 2\frac{xy}{x+y}$ is monotonically increasing in both arguments, since $\frac{df}{dx} = \frac{y^2}{(x+y)^2} > 0$ and $\frac{df}{dy} = \frac{x^2}{(x+y)^2} > 0$. Thus, we obtain:

$$\mathit{fscore}^{\mathfrak{S}}_{\mathcal{M}}(\mathit{Cause}) \geqslant \mathit{fscore}^{\mathfrak{U}}_{\mathcal{M}^{\min}_{\mathit{Cause}}}(\mathit{Cause}),$$

which yields $\mathit{fscore}_{\mathcal{M}}(\mathit{Cause}) \geqslant \mathit{fscore}_{\mathcal{M}^{\min}_{\mathit{Cause}}}(\mathit{Cause})$. □

This now allows us to work under the assumption $\mathcal{M} = \mathcal{M}^{\min}_{\mathit{Cause}}$ when addressing problems concerning the quality measures for a fixed cause set.

However, for the computation of $\mathit{recall}(\mathit{Cause})$ we can rely on known methods for the minimal conditional probability $\mathit{recall}(\mathit{Cause}) = \Pr^{\min}_{\mathcal{M}}(\Diamond \mathit{Cause} \mid \Diamond \mathit{Eff})$ from literature (see [Bai+14; Mär20] for poly-time algorithms to compute conditional reachability probabilities). Moreover, the coverage ratio can be expressed in terms of the recall and thus we can use the same methods to efficiently compute $\mathit{covrat}(\mathit{Cause})$.

Lemma 60 (Computing Coverage Ratio). *The values $\mathit{covrat}(\mathit{Cause})$ and $\mathit{recall}(\mathit{Cause})$ as well as corresponding worst-case schedulers are computable in polynomial time.*

Proof. We fix a scheduler $\mathfrak{S}$ and recall the definitions

$$\mathit{covrat}^{\mathfrak{S}}(\mathit{Cause}) = \frac{\mathsf{tp}^{\mathfrak{S}}}{\mathsf{fn}^{\mathfrak{S}}} \quad \text{and} \quad \mathit{recall}^{\mathfrak{S}}(\mathit{Cause}) = \frac{\mathsf{tp}^{\mathfrak{S}}}{\mathsf{tp}^{\mathfrak{S}} + \mathsf{fn}^{\mathfrak{S}}}.$$

We get the following equations

$$\frac{1}{\frac{\mathsf{tp}^{\mathfrak{S}}}{\mathsf{tp}^{\mathfrak{S}}+\mathsf{fn}^{\mathfrak{S}}}} = \frac{\mathsf{tp}^{\mathfrak{S}}+\mathsf{fn}^{\mathfrak{S}}}{\mathsf{tp}^{\mathfrak{S}}} = \frac{\mathsf{tp}^{\mathfrak{S}}}{\mathsf{tp}^{\mathfrak{S}}} + \frac{\mathsf{fn}^{\mathfrak{S}}}{\mathsf{tp}^{\mathfrak{S}}} = 1 + \frac{\mathsf{fn}^{\mathfrak{S}}}{\mathsf{tp}^{\mathfrak{S}}} = 1 + \frac{1}{\frac{\mathsf{tp}^{\mathfrak{S}}}{\mathsf{fn}^{\mathfrak{S}}}},$$

which imply

$$\frac{1}{recall^{\mathfrak{S}}(Cause)} = 1 + \frac{1}{covrat^{\mathfrak{S}}(Cause)} \text{ thus } covrat^{\mathfrak{S}}(Cause) = 1/\left(\frac{1}{recall^{\mathfrak{S}}(Cause)} - 1\right).$$

Thus, a scheduler $\mathfrak{S}$ realizing the infimum for $covrat^{\mathfrak{S}}(Cause) = 1/\left(\frac{1}{recall^{\mathfrak{S}}(Cause)} - 1\right)$ also realizes the infimum for $recall^{\mathfrak{S}}(Cause)$. The infima above are also minima by [Bai+14] as they are realized by MD scheduler. In order to compute $covrat(Cause)$ we thus need to compute

$$\inf_{\mathfrak{S}} recall^{\mathfrak{S}}(Cause) = \Pr^{\min}_{\mathcal{M}}(\ \Diamond Cause \mid \Diamond Eff\ ),$$

which can be done in polynomial time by [Bai+14; Mär20]. □

Remark 61. Consider an SPR cause (resp. GPR) cause $Cause$ for Eff in $\mathcal{M}$ such that for any other GPR (resp. SPR) cause $Cause'$ we have

$$covrat(Cause) \geqslant covrat(Cause').$$

By the proof of Lemma 60 we then also get that $recall(Cause) \geqslant recall(Cause')$ since for each scheduler $\mathfrak{S}$ we have

$$\frac{1}{recall^{\mathfrak{S}}(Cause)} = 1 + \frac{1}{covrat^{\mathfrak{S}}(Cause)} \leqslant 1 + \frac{1}{covrat^{\mathfrak{S}}(Cause')} = \frac{1}{recall^{\mathfrak{S}}(Cause')}.$$

Vice versa, for an GPR (resp. SPR) cause $Cause$ for which $recall(Cause)$ is maximal among all GPR (resp. SPR) causes also $covrat(Cause)$ is maximal.

Note, that the analysis up to this point is also possible without assuming $\mathcal{M} = \mathcal{M}^{\min}_{Cause}$. Moreover, the reduction of both $recall$ and $covrat$ to minimal conditional probabilities also means that the corresponding worst-case scheduler are MD scheduler in $\mathcal{M}$ by [Bai+14].

However, for $fscore$ we will rely on the transformations and the assumption $\mathcal{M} = \mathcal{M}^{\min}_{Cause}$. The remainder of this subsection will focus on the computation of the f-score of a PR cause. The key observation is the observation that $fscore(Cause)$ can be expressed in terms of the supremum of a quotient of reachability probabilities for disjoint sets of terminal states.

Remark 62 (Ratio for Fscore). Assuming $\mathcal{M} = \mathcal{M}^{\min}_{Cause}$ and $fscore(Cause) > 0$, we have after some straight-forward transformations

$$fscore^{\mathfrak{S}}(Cause) = \frac{2\mathsf{tp}^{\mathfrak{S}}}{2\mathsf{tp}^{\mathfrak{S}} + \mathsf{fn}^{\mathfrak{S}} + \mathsf{fp}^{\mathfrak{S}}}.$$

Using this we get

$$\frac{2}{fscore^{\mathfrak{S}}(Cause)} - 2 = \frac{\mathsf{fp}^{\mathfrak{S}} + \mathsf{fn}^{\mathfrak{S}}}{\mathsf{tp}^{\mathfrak{S}}} = \frac{\Pr^{\mathfrak{S}}(\lozenge FP) + \Pr^{\mathfrak{S}}(\lozenge FN)}{\Pr^{\mathfrak{S}}(\lozenge TP)}$$

Thus, we conclude

$$fscore(Cause) \;=\; \frac{2}{\mathsf{X}+2} \quad \text{where} \quad \mathsf{X} \;=\; \sup_{\mathfrak{S}} \frac{\Pr^{\mathfrak{S}}(\lozenge FP) + \Pr^{\mathfrak{S}}(\lozenge FN)}{\Pr^{\mathfrak{S}}(\lozenge TP)}$$

where $\mathfrak{S}$ ranges over all schedulers with $\Pr^{\mathfrak{S}}(\lozenge TP) > 0$. ◁

Moreover, we can show that we can handle the corner case of $fscore(Cause) = 0$.

Lemma 63 (Fscore=0). *Let $Cause$ be an SPR or a GPR cause. Then, the following three statements are equivalent:*

(a) $fscore(Cause) = 0$.

(b) $\mathsf{X} = \sup_{\mathfrak{S}} \frac{\Pr^{\mathfrak{S}}(\lozenge FP)+\Pr^{\mathfrak{S}}(\lozenge FN)}{\Pr^{\mathfrak{S}}(\lozenge TP)} = \infty$, *($\mathfrak{S}$ ranges over MR schedulers with $\Pr^{\mathfrak{S}}(\lozenge TP) > 0$).*

(c) There is a scheduler $\mathfrak{S}$ such that $\Pr^{\mathfrak{S}}(\lozenge TP) = 0$ and $\Pr^{\mathfrak{S}}(\lozenge FP) + \Pr^{\mathfrak{S}}(\lozenge FN) > 0$.

Proof. "(a) $\implies$ (b)": Suppose $\mathsf{X} \neq \infty$. Then $\mathsf{X} \in \mathbb{R}_{\geqslant 0}$ and by Remark 62 we thus have $fscore(Cause) = \frac{2}{\mathsf{X}+2} > 0$.

"(b) $\implies$ (c)": By [Kal20][Thm 9.16] the MR schedulers of $\mathcal{M} = \mathcal{M}^{\min}_{Cause}$ form a convex space as the closed convex hull of the space of MD schedulers of $\mathcal{M}$. Thus, there is a sequence of MR scheduler $(\mathfrak{S}_i)_{i\in\mathbb{N}}$ which satisfy $\Pr^{\mathfrak{S}_i}(\lozenge TP) > 0$ for each $i \in \mathbb{N}$ and for which

$$\mathsf{X}(\mathfrak{S}_i) = \frac{\Pr^{\mathfrak{S}_i}(\lozenge FP) + \Pr^{\mathfrak{S}_i}(\lozenge FN)}{\Pr^{\mathfrak{S}_i}(\lozenge TP)}$$

converges to ∞ for $i \to \infty$. Since the space of MR scheduler in $\mathcal{M}$ is closed, this sequence converges to an MR scheduler $\mathfrak{S}$ for which $\Pr^{\mathfrak{S}}(\lozenge TP) = 0$ and $\Pr^{\mathfrak{S}}(\lozenge FP) + \Pr^{\mathfrak{S}}(\lozenge FN) > 0$.

"(c) $\implies$ (a)": Given an MR scheduler $\mathfrak{S}$ as described in (c), we can construct a sequence $(\mathfrak{S}_i)_{i\in\mathbb{N}}$ of MR scheduler of $\mathcal{M}$ which reach TP (and thus $Cause$) and for which $fscore^{\mathfrak{S}_i}(Cause)$ converges to 0 by Remark 62. Thus $fscore(Cause) = 0$. □

4.2.1 Max/Min Ratios for Disjoint Sets of Terminal States

We will now give a computation method that is generally applicable for computing extremal ratios of reachability probabilities in MDPs without end-components. Suppose we are given an MDP $\mathcal{M} = (S, Act, P, init)$ without ECs and disjoint subsets $U, V \subseteq S$ of terminal states. Given a scheduler $\mathfrak{S}$ with $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\lozenge V) > 0$ we define:

$$ratio^{\mathfrak{S}}_{\mathcal{M}}(U, V) \;=\; \frac{\Pr^{\mathfrak{S}}_{\mathcal{M}}(\lozenge U)}{\Pr^{\mathfrak{S}}_{\mathcal{M}}(\lozenge V)}.$$

The goal is an algorithm for computing the extremal values:

$$ratio^{\min}_{\mathcal{M}}(\mathrm{U},\mathrm{V}) = \inf_{\mathfrak{S}} ratio^{\mathfrak{S}}_{\mathcal{M}}(\mathrm{U},\mathrm{V}) \quad \text{and} \quad ratio^{\max}_{\mathcal{M}}(\mathrm{U},\mathrm{V}) = \sup_{\mathfrak{S}} ratio^{\mathfrak{S}}_{\mathcal{M}}(\mathrm{U},\mathrm{V}),$$

where $\mathfrak{S}$ ranges over all schedulers with $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathrm{V}) > 0$. For such ratios we call the set U the *nominator set* and V the *denominator set*.

To compute these values, we rely on a polynomial reduction to the classical *stochastic shortest path problem* [BT91]. Given the end-component free MDP $\mathcal{M} = (\mathrm{S}, Act, \mathrm{P}, init)$ we consider the MDP $\mathcal{N} = (\mathrm{S}, Act \cup \{\delta\}, \mathrm{P}', init)$ arising from $\mathcal{M}$ by adding reset transitions from all terminal states $\mathrm{t} \in \mathrm{S} \setminus \mathrm{V}$ not in the denominator set to $init$. Formally, consider the set of terminal states which are no denominator-states $\mathrm{T} = \{\mathrm{t} \in \mathrm{S} \setminus \mathrm{V} \mid \mathrm{t} \text{ is terminal}\}$. For $\mathrm{t} \in \mathrm{T}$ we have $Act_{\mathcal{N}}(\mathrm{t}) = \{\delta\}$ with $\mathrm{P}'(\mathrm{t}, \delta, init) = 1$ and for $\mathrm{s} \in \mathrm{S} \setminus \mathrm{T},\ \mathrm{s}' \in \mathrm{S}$ and $\alpha \in Act$ we have $\mathrm{P}'(\mathrm{s}, \alpha, \mathrm{s}') = \mathrm{P}(\mathrm{s}, \alpha, \mathrm{s}')$. Thus, exactly the denominator-states $\mathrm{s} \in \mathrm{V}$ are terminal in $\mathcal{N}$. With the fresh reset transitions the resulting MDP $\mathcal{N}$ might contain ECs, however these can not intersect V. Furthermore, we equip $\mathcal{N}$ with the weight function that assigns weight 1 to all states in the nominator set U and 0 to all other states. For a scheduler $\mathfrak{T}$ with $\Pr^{\mathfrak{T}}_{\mathcal{N}}(\Diamond \mathrm{V}) = 1$, let $\mathrm{E}^{\mathfrak{T}}_{\mathcal{N}}(\boxplus \mathrm{V})$ be the expected accumulated weight until reaching V under $\mathfrak{T}$. Intuitively, this is the expected number of resets through the nominator set U until reaching the denominator set V. We now use standard notation for the following:

$$\mathrm{E}^{\min}_{\mathcal{N}}(\boxplus \mathrm{V}) = \inf_{\mathfrak{T}} \mathrm{E}^{\mathfrak{T}}_{\mathcal{N}}(\boxplus \mathrm{V}) \qquad \text{and} \qquad \mathrm{E}^{\max}_{\mathcal{N}}(\boxplus \mathrm{V}) = \sup_{\mathfrak{T}} \mathrm{E}^{\mathfrak{T}}_{\mathcal{N}}(\boxplus \mathrm{V}),$$

where $\mathfrak{T}$ ranges over all schedulers with $\Pr^{\mathfrak{T}}_{\mathcal{N}}(\Diamond \mathrm{V}) = 1$.

We can rely on known results [BT91; Alf99; Bai+18] to obtain that both of these values $\mathrm{E}^{\min}_{\mathcal{N}}(\boxplus \mathrm{V})$ and $\mathrm{E}^{\max}_{\mathcal{N}}(\boxplus \mathrm{V})$ are computable in polynomial time. As $\mathcal{N}$ has only non-negative weights, the infimum $\mathrm{E}^{\min}_{\mathcal{N}}(\boxplus \mathrm{V})$ is finite and an MD scheduler realizing the minimal expectation exists. The value $\mathrm{E}^{\max}_{\mathcal{N}}(\boxplus \mathrm{V})$ is finite and the maximum is achieved by an MD scheduler as well if and only if $\mathcal{N}$ has no end-components containing states from the nominator set U. Otherwise, there is such an end-component containing U-states (in terms of [Bai+18] the corresponding EC is *pumping*). By this existence, there is also a scheduler $\mathfrak{S}$ with $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathrm{U}) > 0$ and $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathrm{V}) = 0$. Using convergence to this scheduler by MR scheduler (also cf. Lemma 7) we see that $\mathrm{E}^{\max}_{\mathcal{N}}(\boxplus \mathrm{V}) = +\infty$ in such cases.

The correspondence of the minimal and maximal expectations in the weighted MDP $\mathcal{N}$ with the ratio function is given by the next Theorem 64.

Theorem 64 (Ratio Reduction to SSP). *Let $\mathcal{M}$ be an MDP without ECs and U, V disjoint sets of terminal states in $\mathcal{M}$. Consider the MDP $\mathcal{N}$ from above. Then, $ratio^{\min}_{\mathcal{M}}(\mathrm{U},\mathrm{V}) = \mathrm{E}^{\min}_{\mathcal{N}}(\boxplus \mathrm{V})$ and $ratio^{\max}_{\mathcal{M}}(\mathrm{U},\mathrm{V}) = \mathrm{E}^{\max}_{\mathcal{N}}(\boxplus \mathrm{V})$. Furthermore, both values are computable in polynomial time, and there is an MD scheduler minimizing $ratio^{\mathfrak{S}}_{\mathcal{M}}(\mathrm{U},\mathrm{V})$, and an MD-scheduler maximizing $ratio^{\mathfrak{S}}_{\mathcal{M}}(\mathrm{U},\mathrm{V})$ if $ratio^{\max}_{\mathcal{M}}(\mathrm{U},\mathrm{V})$ is finite.*

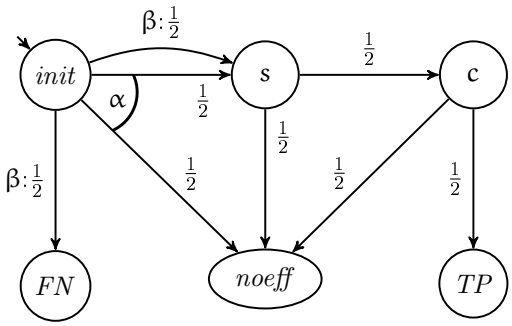

Figure 4.10: MDP $\mathcal{M}$ from Example 65

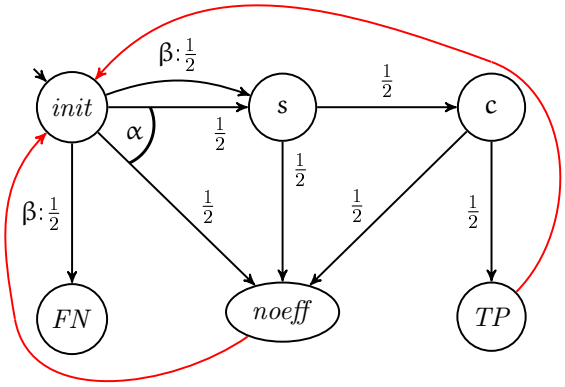

Figure 4.11: MDP $\mathcal{N}$ with reset transitions for $ratio^{\min}_{\mathcal{M}}(TP, FN)$

Before we show this claim, let us take a look at an example for the construction of $\mathcal{N}$.

Example 65 (Ratio Transformation). Consider the MDP $\mathcal{M}$ from Figure 4.10 and the effect set $Eff = \{FN, TP\}$ and suppose the task is to compute $covrat(c) = ratio^{\min}_{\mathcal{M}}(TP, FN)$. The construction on $\mathcal{N}$ is depicted in Figure 4.11 resulting in $\mathcal{N}$, where reset transitions for *noeff* and TP have been added (red edges) and FN is the only terminal state. The weight function now assigns 1 to TP and 0 to all others. By applying Theorem 64 we have $covrat(c) = ratio^{\min}_{\mathcal{M}}(TP, FN) = E^{\min}_{\mathcal{N}}(\boxplus FN)$. A MD scheduler realizing this value would always chooses β in *init*. The corresponding value is $E^{\min}_{\mathcal{N}}(\boxplus FN) = \frac{1}{4}$. ◁

Proof of Theorem 64. $\mathcal{M}$ has a scheduler $\mathfrak{S}$ with $Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond U) > 0$ and $Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond V) = 0$ if and only if the transformed MDP $\mathcal{N}$ from above has an EC containing at least one U-state. Therefore, we then have $E^{\max}_{\mathcal{N}}(\boxplus V) = +\infty$. Otherwise, $E^{\max}_{\mathcal{N}}(\boxplus V)$ is finite.

For the following we only consider the minimal ratio $ratio^{\min}_{\mathcal{M}}(U, V) = E^{\min}_{\mathcal{N}}(\boxplus V)$ since the arguments for the maximum are similar. First, we show $ratio^{\min}_{\mathcal{M}}(U, V) \geqslant E^{\min}_{\mathcal{N}}(\boxplus V)$. For this, we consider an arbitrary scheduler $\mathfrak{S}$ for $\mathcal{M}$. Let

$$x = Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond U), \qquad p = Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond V), \qquad q = 1 - x - p.$$

So, in particular for $p > 0$ we have

$$ratio^{\mathfrak{S}}_{\mathcal{M}}(U, V) = \frac{Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond U)}{Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond V)} = \frac{x}{p}$$

Consider the scheduler $\mathfrak{T}$ of $\mathcal{N}$ behaving as $\mathfrak{S}$ in $\mathcal{M}$ every iteration. Then:

$$E^{\mathfrak{T}}_{\mathcal{N}}(\boxplus V) = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} n \cdot x^n \cdot \binom{n+k}{k} \cdot q^k \cdot p \overset{(4.9)}{=} \frac{x}{p}, \tag{4.8}$$

where (4.9) relies on some basic calculations (see Lemma 66).

This yields:

$$ratio^{\mathfrak{S}}_{\mathcal{M}}(U,V) \;=\; \frac{x}{p} \;=\; \mathrm{E}^{\mathfrak{T}}_{\mathcal{N}}(\boxplus V) \;\geqslant\; \mathrm{E}^{\min}_{\mathcal{N}}(\boxplus V).$$

Hence, $ratio^{\min}_{\mathcal{M}}(U,V) \geqslant \mathrm{E}^{\min}_{\mathcal{N}}(\boxplus V)$.

For the other direction $ratio^{\min}_{\mathcal{M}}(U,V) \leqslant \mathrm{E}^{\min}_{\mathcal{N}}(\boxplus V)$, we use the fact that there is an MD scheduler $\mathfrak{T}$ for $\mathcal{N}$ such that $\mathrm{E}^{\mathfrak{T}}_{\mathcal{N}}(\boxplus V) = \mathrm{E}^{\min}_{\mathcal{N}}(\boxplus V)$. We then interpret $\mathfrak{T}$ as an MD scheduler for the original MDP $\mathcal{M}$ choosing exactly the same actions in all states but the terminal ones. Using (4.8) we obtain

$$\mathrm{E}^{\mathfrak{T}}_{\mathcal{N}}(\boxplus V) \;=\; \frac{\Pr^{\mathfrak{T}}_{\mathcal{M}}(\Diamond U)}{\Pr^{\mathfrak{T}}_{\mathcal{M}}(\Diamond V)} \;=\; ratio^{\mathfrak{T}}_{\mathcal{M}}(U,V) \;\geqslant\; ratio^{\min}_{\mathcal{M}}(U,V),$$

which yields $ratio^{\min}_{\mathcal{M}}(U,V) \leqslant \mathrm{E}^{\min}_{\mathcal{N}}(\boxplus V)$ and concludes the proof for the minimal ratio. For the maximal value $\mathrm{E}^{\max}_{\mathcal{N}}(\boxplus V) \;=\; ratio^{\max}_{\mathcal{M}}(U,V)$ similar arguments lead to the claim.

Furthermore, we can now use known results [BT91; Alf99; Bai+18] to compute the extremal expectations $\mathrm{E}^{\min}_{\mathcal{N}}(\boxplus V)$ and $\mathrm{E}^{\max}_{\mathcal{N}}(\boxplus V)$ in polynomial time. □

Lemma 66. *For* $x, p, q \in [0,1]$ *such that* $x+q+p = 1$ *we have*

$$\sum_{n=0}^{\infty}\sum_{k=0}^{\infty} n \cdot x^n \cdot \binom{n+k}{k} \cdot q^k \cdot p \;=\; \frac{x}{p}. \tag{4.9}$$

Proof. We first show for $0 < q < 1$, $n \in \mathbb{N}$ and

$$\sum_{k=0}^{\infty} \binom{n+k}{k} \cdot q^k = \frac{1}{(1-q)^{n+1}}. \tag{4.10}$$

This is done by induction on n. The claim is clear for $n = 0$. For the step of induction let a_n denote the left hand side of (4.10). Note that

$$\binom{n+1+k}{k} \;=\; \binom{n+k}{k} + \binom{n+k}{k-1} \;=\; \binom{n+k}{k} + \binom{(n+1)+(k-1)}{k-1},$$

which yields $a_{n+1} = a_n + q \cdot a_{n+1}$ and hence $a_{n+1} = \frac{a_n}{1-q}$. The equation (4.10) then follows directly from the induction hypothesis.

The claim now follows by the following calculations and using the preliminary induction:

$$\begin{aligned}
\sum_{n=0}^{\infty}\sum_{k=0}^{\infty} n \cdot x^n \cdot \binom{n+k}{k} \cdot q^k \cdot p &= \sum_{n=0}^{\infty} n \cdot x^n \cdot \frac{1}{(1-q)^{n+1}} \cdot p \\
&= \frac{p}{1-q} \cdot \sum_{n=0}^{\infty} n \cdot \left(\frac{x}{1-q}\right)^n \\
&= \frac{p}{1-q} \cdot \frac{\frac{x}{1-q}}{\left(1-\frac{x}{1-q}\right)^2} \\
&= \frac{px}{(1-q-x)^2} \;=\; \frac{px}{p^2} \;=\; \frac{x}{p},
\end{aligned}$$

where we use $p = 1-q-x$. □

We are now in a position to apply this framework for $ratio^{\max}_{\mathcal{M}}(\mathbf{U}, \mathbf{V})$ to the f-score, which yields a polynomial time algorithm to compute $fscore(Cause)$.

Theorem 67 (Computing $fscore$). *Given an MDP $\mathcal{M}$ with a non-empty set of terminal states $Eff \subset S$ and a non-empty set $Cause \subseteq S \setminus Eff$, the value $fscore(Cause)$ and corresponding worst-case schedulers are computable in polynomial time.*

Proof. We use the polynomial time transformation from $\mathcal{M}$ to $\mathcal{M}^{\min}_{Cause}$ which does not change the value $fscore(Cause)$ due to Corollary 58 and Lemma 59. Using the calculations of Remark 62 we have

$$fscore(Cause) \;=\; \frac{2}{\mathsf{X}+2} \quad \text{where} \quad \mathsf{X} \;=\; \sup_{\mathfrak{S}} \frac{\Pr^{\mathfrak{S}}(\lozenge FP) + \Pr^{\mathfrak{S}}(\lozenge FN)}{\Pr^{\mathfrak{S}}(\lozenge TP)}$$

where $\mathfrak{S}$ ranges over all schedulers with $\Pr^{\mathfrak{S}}(\lozenge TP) > 0$. Thus, the task is to compute

$$\mathsf{X} = \sup_{\mathfrak{S}} \frac{2}{fscore^{\mathfrak{S}}(Cause)} - 2 = \sup_{\mathfrak{S}} \frac{\Pr^{\mathfrak{S}}_{\mathcal{M}}(\lozenge FP) + \Pr^{\mathfrak{S}}_{\mathcal{M}}(\lozenge FN)}{\Pr^{\mathfrak{S}}_{\mathcal{M}}(\lozenge TP)},$$

over the range of schedulers $\mathfrak{S}$ with $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\lozenge TP) > 0$. As X is a supremum in the form of Theorem 64, this yields the claim that the optimal value is computable in polynomial time.

In the case where $fscore(Cause) = 0$, we do not obtain an optimal scheduler via Theorem 64. However, Lemma 63 shows that in such a case there exists a scheduler $\mathfrak{S}$ with $\Pr^{\mathfrak{S}}(\lozenge TP) = 0$ and $\Pr^{\mathfrak{S}}(\lozenge FP) + \Pr^{\mathfrak{S}}(\lozenge FN) > 0$. Such a scheduler can be computed in polynomial time as any (memoryless) scheduler in the largest sub-MDP of $\mathcal{M}$ that does not contain TP. This sub-MDP can be constructed by successively removing states and state-action pairs. □

Remark 68 (Memory Requirement of Worst-Case Schedulers for $fscore$). Recall the action τ in $\mathcal{M}_{Cause}$ which was freshly introduced in the transformation $\mathcal{M} \to \mathcal{M}^{\min}_{Cause}$. Action τ is only enabled in states $s_{\mathcal{E}}$ representing MECs of the original MDP and which transitions to TN with probability 1. Since a worst-case scheduler $\mathfrak{U}$ for $fscore^{\mathfrak{U}}(Cause) = fscore(Cause)$ in $\mathcal{M}^{\min}_{Cause}$ is memoryless deterministic, it chooses τ with probability 0 or 1. Therefore, we can use Theorem 39 to lift $\mathfrak{U}$ to an MD scheduler $\mathfrak{T}$ in $\mathcal{M}_{<Cause>}$ and then to a finite-memory scheduler $\mathfrak{S}$ in $\mathcal{M}$ with two memory cells for "before" and "after" $Cause$.

For the special case that $fscore(Cause) = 0$ the witnessing scheduler $\mathfrak{S}$ must only satisfy $\Pr^{\mathfrak{S}}(\lozenge TP) = 0$ and $\Pr^{\mathfrak{S}}(\lozenge FP) + \Pr^{\mathfrak{S}}(\lozenge FN) > 0$ by Lemma 63. An MD scheduler of the largest sub-MDP of $\mathcal{M}$ that does not contain TP is sufficient for this and thus we can again use above analysis. ◁

4.3 Quality-Optimal Probability-Raising Causes

Up to this point we addressed the computation of quality measures for a given cause. This can be used to compare different PR causes in terms of their coverage or quality properties with respect to the effect. Moreover, the motivation for introducing quality measures comes from the need of a criteria for which optimal causes can be computed. For this section the goal is to present algorithms to determine "good" causes according to the quality measures defined in Section 3.3. Throughout this section we are given an MDP $\mathcal{M} = (\mathrm{S}, Act, \mathrm{P}, init)$ and a non-empty set of terminal effect states $\mathit{Eff} \subseteq \mathrm{S}$. We have seen so far that small (singleton) causes are easy to determine (see Section 4.1.2). Moreover, we can easily adapt proposed existence-checking algorithm Lemma 49 in such a way that the algorithm returns a singleton probability-raising cause $\{\mathrm{c}_0\}$ with maximal precision.

Remark 69 (Precision-Optimal PR Cause). Recall that the precision of a cause Cause is a minimal conditional probability

$$\inf_{\mathfrak{S}} \Pr^{\mathfrak{S}}(\Diamond \mathit{Eff} \mid \Diamond \mathit{Cause}) = \Pr^{\min}(\Diamond \mathit{Eff} \mid \Diamond \mathit{Cause}).$$

Given a scheduler $\mathfrak{S}$ the conditional probability is a weighted average over $\Pr^{\mathfrak{S}}_{\mathrm{c}}(\Diamond \mathit{Eff})$ by

$$\Pr^{\mathfrak{S}}(\Diamond \mathit{Eff} \mid \Diamond \mathit{Cause}) \;=\; \frac{1}{\Pr^{\mathfrak{S}}(\Diamond \mathit{Cause})} \sum_{\mathrm{c} \in \mathit{Cause}} \Pr^{\mathfrak{S}}((\neg \mathit{Cause}) \,\mathrm{U}\, \mathrm{c}) \cdot \Pr^{\mathfrak{S}}_{\mathrm{c}}(\Diamond \mathit{Eff}).$$

Thus, a precision optimal GPR and SPR cause is a singleton $\mathrm{c}_0 \in \mathrm{S} \setminus \mathit{Eff}$ such that the value $\Pr^{\min}_{\mathrm{c}_0}(\Diamond \mathit{Eff})$ is maximal. Note that such a singleton can be a *sufficient* cause for Eff if $\mathit{precision}(\mathrm{c}_0) = 1$. In such cases, reaching c_0 on any path immediately implies that Eff will be reached eventually. ◁

However, as already argued in Section 3.3, singleton or small cause sets tend to have poor coverage as the probability for false negative paths that reach an effect state without visiting a cause state before can be large. In order to derive "good" causes with respect to different needs, we will focus on the quality measures recall, covrat and fscore. In this section we thus address the computation of quality-optimal probability-raising causes.

For the computation of a given quality measure for a fixed PR cause there is no difference between global- and strict PR causes as the quality values are only defined in terms of the set. However, when finding optimal causes the distinction between GPR and SPR makes a difference as it changes the candidate sets in question. In particular, let

$$\mathcal{C}_{\mathrm{S}} = \{\mathrm{C} \subseteq \mathrm{S} \setminus \mathit{Eff} \mid \mathrm{C} \text{ is an SPR cause for } \mathit{Eff}\} \quad \text{and}$$
$$\mathcal{C}_{\mathrm{G}} = \{\mathrm{C} \subseteq \mathrm{S} \setminus \mathit{Eff} \mid \mathrm{C} \text{ is a GPR cause for } \mathit{Eff}\}.$$

(Note that $\mathcal{C}_{\mathrm{S}} \subseteq \mathcal{C}_{\mathrm{G}}$ by Proposition 15.) Then optimality with respect to a quality measure

means optimality among the respective type of causes. We say an SPR cause Cause is *recall-optimal* if its recall is maximal among all SPR causes, i.e.

$$\mathit{recall}(\mathit{Cause}) = \max_{\mathbf{C} \in \mathcal{C}_S} \mathit{recall}(\mathbf{C}).$$

Likewise, *ratio-optimality* and *f-score-optimality* of Cause means maximality of $\mathit{covrat}(\mathit{Cause})$ resp. $\mathit{fscore}(\mathit{Cause})$ among all SPR causes. Recall-, ratio- and f-score-optimality for GPR causes are defined accordingly with $\mathcal{C}_G$ instead of $\mathcal{C}_S$.

For this subsection we still consider an MDP $\mathcal{M} = (S, \mathit{Act}, P, \mathit{init})$ and a set of terminal states $\mathit{Eff} \subset S$ constituting the effect. For each quality measure $\mathit{quality} \in \{\mathit{recall}, \mathit{covrat}, \mathit{fscore}\}$ address the computation of respective sets $\mathit{Cause} \subseteq S \setminus \mathit{Eff}$ such that Cause is a $\mathit{quality}$-optimal GPR (resp. SPR) cause. Note that we do not assume a transformed version of $\mathcal{M}$ as the all of the given transformations need a fixed set Cause to work. We note that for both GPR and SPR causes Cause being recall-optimal or being ratio-optimal is equivalent by Remark 61.

4.3.1 Quality-Optimal SPR Causes

We first consider the strict case and look for quality-optimal SPR causes in the following paragraphs. The computation of a recall- and ratio-optimal SPR cause is particularly efficient as we can describe a canonical SPR cause $\mathit{CanCause}$ which achieves optimal recall and coverage ratio. For this we make use of the techniques of Section 4.1.2 which yield an algorithm for generating this canonical PR cause.

For this consider the set of all states which constitute a singleton SPR cause, denoted $\mathcal{S} = \{s \in S \mid \{s\} \text{ is an SPR cause}\}$. The canonical cause is defined as the set of states $c \in \mathcal{S}$ such that there is a scheduler $\mathfrak{S}$ with $\Pr^{\mathfrak{S}}(\neg \mathcal{S} \,\mathrm{U}\, c) > 0$, that is $\mathit{CanCause} = \{c \in \mathcal{S} \mid \Pr^{\max}(\neg \mathcal{S} \,\mathrm{U}\, c) > 0\}$. Intuitively, this ensures that $\mathit{CanCause}$ is satisfying the minimality condition as it can be seen as the "front" of $\mathcal{S}$.

Theorem 70 (Canonical SPR Cause). *Given an MDP $\mathcal{M} = (S, \mathit{Act}, P, \mathit{init})$, a non-empty set of terminal states $\mathit{Eff} \subset S$, if the set $\mathcal{S}$ as defined above is non-empty, then $\mathit{CanCause} = \{c \in \mathcal{S} \mid \Pr^{\max}(\neg \mathcal{S} \,\mathrm{U}\, c) > 0\}$ is a ratio- and recall-optimal SPR cause and computable in polynomial time.*

Proof. By definition of SPR causes any subset $C \subseteq \mathcal{S}$ satisfying $\Pr^{\max}((\neg c) \,\mathrm{U}\, c) > 0$ for each $c \in C$ constitutes an SPR cause and thus $\mathit{CanCause}$ is also an SPR cause. We show that $\mathit{CanCause}$ has path-wise optimal coverage in the following sense: If C is any SPR cause then $C \subseteq \mathcal{S}$ by definition and for each path π in $\mathcal{M}$ we have

$$\pi \models \neg\Diamond \mathit{CanCause} \wedge \Diamond \mathit{Eff} \implies \pi \models \neg\Diamond C \wedge \Diamond \mathit{Eff} \text{ and}$$
$$\pi \models \Diamond C \wedge \Diamond \mathit{Eff} \implies \pi \models \Diamond \mathit{CanCause} \wedge \Diamond \mathit{Eff}.$$

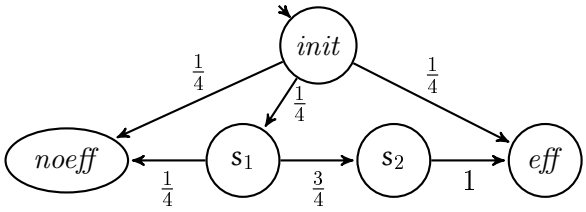

Figure 4.12: Markov chain where $CanCause$ is not fscore-optimal

This implies

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathbf{C} \wedge \Diamond \mathit{Eff}) \leqslant \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{CanCause} \wedge \Diamond \mathit{Eff}),$$
$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\neg\Diamond \mathbf{C} \wedge \Diamond \mathit{Eff}) \geqslant \Pr^{\mathfrak{S}}_{\mathcal{M}}(\neg\Diamond \mathit{CanCause} \wedge \Diamond \mathit{Eff})$$

for every scheduler $\mathfrak{S}$, which yields recall- and ratio-optimality of $CanCause$. Using Algorithm 41 to compute the set $\mathcal{S}$ and standard depth-first search algorithms we see that $CanCause$ is computable in polynomial time (cf. Theorem 44). □

Remark 71 (Screening off). Computing the canonical SPR cause $CanCause$ for an effect Eff in $\mathcal{M}$ also addresses problems with causal chains and screening off which are discussed in Example 22. To see this, recall that we say a PR cause $Cause_1$ is *screening off* another PR cause $Cause_0$ if $\Diamond Cause_0$ and $\Diamond \mathit{Eff}$ are probabilistically independent conditional upon $\Diamond Cause_1$ [Hit21]. As the canonical SPR cause is the front of all SPR causes, on a temporal scale it will always be the first probability-raising cause which is visited . Therefore, $CanCause$ will never screen off any other SPR cause and can not be causally between the effect and another cause [Rei56]. However, it can very well happen that the canonical SPR cause is screened off by other PR causes. ◁

Next, we address the computation of fscore-optimal SPR causes. First, note that in general the canonical SPR cause $CanCause$ is not f-score-optimal. To see this, consider the Markov chain from Figure 4.12. There, we have $CanCause = \{s_1\}$, which has $\mathit{precision}(CanCause) = \frac{3}{4}$ and $\mathit{recall}(CanCause) = \frac{3}{8}/(\frac{1}{4} + \frac{3}{8}) = \frac{3}{5}$. But the SPR cause $\{s_2\}$ has better f-score as its precision is 1 and it has the same recall as $CanCause$.

Remark 72 (Computing fscore-Optimal SPR Causes). Since the value $\mathit{fscore}(Cause)$ of a given set $Cause$ can be computed in polynomial time by Theorem 67 this yields a polynomial space algorithm for computing an f-score-optimal SPR cause in $\mathcal{M}$. For this we compute the value $\mathit{fscore}(Cause)$ for all possible SPR causes $Cause$ one by one. For each candidate set $Cause \subseteq \mathbf{S} \setminus \mathit{Eff}$ the SPR check (Algorithm 41 for all $\mathbf{c} \in Cause$) runs in polynomial time as well. As the space used for $Cause$ can be reused after each computation, this results in polynomial space. ◁

For Markov chains, we can do better and compute an f-score-optimal SPR cause in polynomial time by using a polynomial reduction to the stochastic shortest path problem:

Theorem 73 ($fscore$-optimal in DTMCs). *Let $\mathcal{M} = (S, P, init)$ be a Markov chain, $\mathit{Eff} \subset S$ a non-empty set of terminal states and let $\mathcal{S} \neq \emptyset$. Then, an f-score-optimal SPR cause for Eff in $\mathcal{M}$ can be computed in polynomial time.*

Proof. We regard the given Markov chain $\mathcal{M}$ as an MDP with a singleton action set $Act = \{\alpha\}$. As $\mathcal{M}$ has SPR causes, the set $\mathcal{S}$ of states which constitute singleton SPR causes is nonempty. Furthermore, we may assume that $\mathcal{M}$ has no non-trivial (i.e., cyclic) bottom strongly connected components as we may collapse them. Let $p_c = \Pr_{\mathcal{M},c}(\Diamond \mathit{Eff})$.

We transform $\mathcal{M}$ to a new MDP $\mathcal{K} = (S_{\mathcal{K}}, Act_{\mathcal{K}}, P_{\mathcal{K}}, init)$ with

$$S_{\mathcal{K}} = S \cup \{TP, FP\} \qquad \text{and} \qquad Act_{\mathcal{K}} = \{\alpha, \gamma\}$$

where FP and TP are fresh states. For $s, s' \in S$ the function $P_{\mathcal{K}}$ works as in $\mathcal{M}$, i.e.,

$$P_{\mathcal{K}}(s, \alpha, s') = P(s, \alpha, s')$$

Moreover, we add the following transitions:

(i) In each SPR state $c \in \mathcal{S}$ the action γ is enabled such that $P_{\mathcal{K}}(c, \gamma, TP) = p_c$ and $P_{\mathcal{K}}(c, \gamma, FP) = 1 - p_c$.

(ii) In all terminal states of $\mathcal{M}$ and in the fresh state FP reset transitions to $init$ with action label α are added, i.e., $P_{\mathcal{K}}(FP, \alpha, init) = 1$ and $P_{\mathcal{K}}(s, \alpha, init) = 1$ for $s \in \mathit{Eff}$ or if s is a terminal non-effect state of $\mathcal{M}$.

So, TP is the only terminal state in $\mathcal{K}$, and $Act_{\mathcal{K}}(c) = \{\alpha, \gamma\}$ for $c \in \mathcal{S}$, while $Act_{\mathcal{K}}(s) = \{\alpha\}$ for all other states $s \in S \setminus \mathcal{S}$. Intuitively, taking action γ in state $c \in \mathcal{S}$ selects c to be a cause state. The states in Eff represent uncovered effects in $\mathcal{K}$, while TP stands for covered effects.

We now equip $\mathcal{K}$ with a weight function that assigns weight 1 to all states in $U = \mathit{Eff} \cup \{FP\}$ and weight 0 to all other states of $\mathcal{K}$. For $V = \{TP\}$ the expectation $f = E^{\min}_{\mathcal{K}}(\boxplus V)$ and a corresponding MD scheduler $\mathfrak{S}$ for $\mathcal{K}$ such that $E^{\mathfrak{S}}_{\mathcal{K}}(\boxplus V) = f$ are computable in polynomial time by [BT91; Alf99]. Let $\mathcal{S}_\gamma$ denote the set of states $c \in \mathcal{S}$ in which the computed scheduler $\mathfrak{S}$ chooses γ, so $\mathfrak{S}(c) = \gamma$. We then define the set Cause as the set of those states $c \in \mathcal{S}_\gamma$ such that $\mathcal{M}$ has a path satisfying $\neg \mathcal{S}_\gamma \,\mathrm{U}\, c$. Cause is an SPR cause of Eff in $\mathcal{M}$ as we only compose it of states $c \in \mathcal{S}$ and satisfies $\Pr^{\max}_{\mathcal{M}}((\neg \mathit{Cause})\,\mathrm{U}\, c) > 0$ is satisfied for all states $c \in \mathit{Cause}$. With arguments as in the proof of Theorems 67 and 64 we obtain $\mathit{fscore}(\mathit{Cause}) = 2/(f + 2)$.

It remains to show that Cause is f-score-optimal. To this end, let C be an arbitrary SPR cause of Eff in $\mathcal{M}$. By definition we have $C \subseteq \mathcal{S}$. Consider the MD scheduler $\mathfrak{T}$ for $\mathcal{K}$ that schedules γ in C and α for all other states of $\mathcal{K}$. Then, $\mathit{fscore}(C) = 2/(f^{\mathfrak{T}} + 2)$ where $f^{\mathfrak{T}} = E^{\mathfrak{T}}_{\mathcal{K}}(\boxplus V)$ and thus $f \leqslant f^{\mathfrak{T}}$, as f is minimal. This yields $\mathit{fscore}(\mathit{Cause}) \geqslant \mathit{fscore}(C)$ which proves that Cause is f-score-optimal. □

The naive adaption of the construction presented in the proof of Theorem 73 for proper MDPs yields a stochastic game structure where the objective of one player is to minimize the expected accumulated weight until reaching a target state. Although algorithms for *stochastic shortest path (SSP) games* are known [PB99], they rely on assumptions on the game structure which are not satisfied here. However, for the threshold problem *SPR-f-score* where inputs are an MDP $\mathcal{M}$, Eff and $\vartheta \in \mathbb{Q}_{\geqslant 0}$ and the task is to decide the existence of an SPR cause Cause such that $\mathit{fscore}(\mathit{Cause}) > \vartheta$, we can establish a polynomial reduction to SSP games, which yields an $\mathrm{NP} \cap \mathrm{coNP}$ upper bound. This comes from the fact that for SSP games it is known that optimal memoryless deterministic strategies exist for both players [PB99]. The following considerations in Lemma 74 and Theorem 75 are dedicated to prove these claims.

Recall that for a given SPR cause C and scheduler $\mathfrak{S}$ we have

$$\mathit{fscore}^{\mathfrak{S}}(\mathrm{C}) > \vartheta \quad \text{iff} \quad \frac{2\mathit{tp}^{\mathfrak{S}}}{2\mathit{tp}^{\mathfrak{S}} + \mathit{fp}^{\mathfrak{S}} + \mathit{fn}^{\mathfrak{S}}} > \vartheta.$$

To prove the upper bound of SPR-f-score we reformulate the condition of SPR-f-score.

Lemma 74 (Alternative fscore Threshold Condition). *Given an MDP $\mathcal{M}$ with a set of terminal states $\mathit{Eff} \subset \mathrm{S}$ and a non-negative rational number $\vartheta \in \mathbb{Q}_{\geqslant 0}$, let C be an SPR cause for Eff in $\mathcal{M}$. We have, $\mathit{fscore}(\mathrm{C}) > \vartheta$ if and only if*

$$2(1-\vartheta)\mathit{tp}^{\mathfrak{S}} - \vartheta\mathit{fp}^{\mathfrak{S}} - \vartheta\mathit{fn}^{\mathfrak{S}} > 0 \tag{4.11}$$

for all schedulers $\mathfrak{S}$ for $\mathcal{M}$ with $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff}) > 0$.

Proof. Assume $\mathit{fscore}(\mathrm{C}) > \vartheta$ and let $\mathfrak{S}$ be a scheduler with $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff}) > 0$. If $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathrm{C}) = 0$, then $\mathit{fscore}(\mathrm{C})$ would be 0 which contradicts the assumption. Thus, $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathrm{C}) > 0$. Then,

$$\mathit{fscore}^{\mathfrak{S}}(\mathrm{C}) = \frac{2\mathit{tp}^{\mathfrak{S}}}{2\mathit{tp}^{\mathfrak{S}} + \mathit{fp}^{\mathfrak{S}} + \mathit{fn}^{\mathfrak{S}}} > \vartheta \qquad \Longrightarrow \qquad 2(1-\vartheta)\mathit{tp}^{\mathfrak{S}} - \vartheta\mathit{fp}^{\mathfrak{S}} - \vartheta\mathit{fn}^{\mathfrak{S}} > 0.$$

Now, suppose that (4.11) holds for a scheduler $\mathfrak{S}$ with $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff}) > 0$. Without loss of generality we can assume $\mathfrak{S}$ to be a minimizing scheduler for $\mathit{fscore}^{\mathfrak{S}}(\mathrm{C})$, which exists by Theorem 67. From (4.11), we conclude

$$2(1-\vartheta)\mathit{tp}^{\mathfrak{S}} - \vartheta\mathit{fp}^{\mathfrak{S}} - \vartheta\mathit{fn}^{\mathfrak{S}} > 0$$

and hence that $\mathit{fscore}^{\mathfrak{S}}(\mathrm{C}) > \vartheta$ as claimed. □

Theorem 75 (Complexity for fscore-Optimal SPR Causes). *Given an MDP $\mathcal{M} = (\mathrm{S}, \mathit{Act}, \mathrm{P}, \mathit{init})$, a non-empty set of terminal states $\mathit{Eff} \subset \mathrm{S}$ and a non-negative rational number $\vartheta \in \mathbb{Q}_{\geqslant 0}$, deciding whether there exists an SPR cause $\mathit{Cause} \subseteq \mathrm{S} \setminus \mathit{Eff}$ for Eff such that $\mathit{fscore}(\mathit{Cause}) > \vartheta$ can be done in $\mathrm{NP} \cap \mathrm{coNP}$.*

Proof. Consider the set $\mathfrak{S}$ of states $\mathbf{c} \in S \setminus \mathit{Eff}$ where the singleton $\{\mathbf{c}\}$ is an SPR cause. If $\mathfrak{S}$ is empty then the threshold problem is trivial as there is no SPR cause at all. Thus, we suppose that $\mathfrak{S}$ is nonempty. Note that $\Pr^{\min}_{\mathcal{M},\mathbf{c}}(\Diamond \mathit{Eff}) > 0$ for all $\mathbf{c} \in \mathfrak{S}$.

Consider the MEC-quotient $\mathcal{N} = (S_{\mathcal{N}}, \mathit{Act}_{\mathcal{N}}, P_{\mathcal{N}}, \mathit{init}_{\mathcal{N}})$ of $\mathcal{M}$ with the fresh terminal state $\bot$ (cf. 9). As the states in Eff are not part of any end-component of $\mathcal{M}$, no state $\mathbf{c} \in \mathfrak{S}$ is contained in an end-component of $\mathcal{M}$ either. Therefore, the MEC-quotient $\mathcal{N}$ contains the states from Eff and $\mathfrak{C}$.

Claim 1: There is an SPR cause C for Eff in $\mathcal{M}$ with $\mathit{fscore}(C) > \vartheta$ if and only if there is an SPR cause C' for Eff in $\mathcal{N}$ with $\mathit{fscore}(C') > \vartheta$.

Proof of Claim 1. We first observe that all reachability probabilities involved in the claim do not depend on the behavior during the traversal of MECs. Furthermore, staying inside an MEC in $\mathcal{M}$ can be mimicked in $\mathcal{N}$ by moving to $\bot$, and vice versa. More precisely, let $C \subseteq \mathfrak{S}$. Then, analogously to Lemma 9, for each scheduler $\mathfrak{S}$ for $\mathcal{M}$, there is a scheduler $\mathfrak{T}$ for $\mathcal{N}$, and vice versa, such that

- $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff} \mid (\neg \mathbf{c}) \,\mathrm{U}\, \mathbf{c}) = \Pr^{\mathfrak{T}}_{\mathcal{N}}(\Diamond \mathit{Eff} \mid (\neg \mathbf{c}) \,\mathrm{U}\, \mathbf{c})$ for all $\mathbf{c} \in C$ for which these values are defined,
- $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff}) = \Pr^{\mathfrak{T}}_{\mathcal{N}}(\Diamond \mathit{Eff})$,
- $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff} \mid \Diamond C) = \Pr^{\mathfrak{T}}_{\mathcal{N}}(\Diamond \mathit{Eff} \mid \Diamond C)$ if the values are defined, and
- $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond C \mid \Diamond \mathit{Eff}) = \Pr^{\mathfrak{T}}_{\mathcal{N}}(\Diamond C \mid \Diamond \mathit{Eff})$ if the values are defined.

Hence, C is an SPR cause for Eff in $\mathcal{M}$ if and only if it is in an SPR cause in $\mathcal{N}$. Furthermore, the f-score of C in $\mathcal{M}$ and in $\mathcal{N}$ agree, as the input values of fscore agree in both MDPs. This finishes the proof of Claim 1.

Model Transformation for Ensuring Positive Effect Probabilities. Recall that the f-score is only defined for schedulers reaching Eff with positive probability. Now, we apply a further model transformation which ensures that Eff is reached with positive probability under all schedulers. If this is already the case, there is nothing to do. Therefore, we assume that $\Pr^{\min}_{\mathcal{N}}(\Diamond \mathit{Eff}) = 0$.

We define the subset of states from which Eff can be avoided as $D \subseteq S_{\mathcal{N}}$ by

$$D = \{\mathbf{s} \in S_{\mathcal{N}} \mid \Pr^{\min}_{\mathcal{N},\mathbf{s}}(\Diamond \mathit{Eff}) = 0\}.$$

By assumption D is nonempty as $\mathit{init}_{\mathcal{N}} \in D$. For each $\mathbf{s} \in D$, we further define the set of enabled actions in $\mathbf{s}$ which minimize the reachability probability of Eff from $\mathbf{s}$ by

$$\mathit{Act}^{\min}(\mathbf{s}) = \{\alpha \in \mathit{Act}_{\mathcal{N}}(\mathbf{s}) \mid P_{\mathcal{N}}(\mathbf{s}, \alpha, D) = 1\}.$$

Finally, let $\mathrm{E} \subseteq \mathrm{D}$ be the set of states which are reachable from $init_{\mathcal{N}}$ when only choosing actions from $Act^{\min} = \bigcap_{s \in \mathrm{D}} Act^{\min}(s)$. Note that E does not contain states from $\mathcal{S}$, as no state in E constitutes a singleton SPR cause.

All schedulers that reach *Eff* with positive probability in $\mathcal{N}$ have to leave the sub-MDP consisting of E and the actions in $Act^{\min}$ at some point. Let us denote this sub-MDP by $\mathcal{N}_{\mathrm{E}}^{\min}$. We define the set of state-action pairs Π which leave the sub-MDP $\mathcal{N}_{\mathrm{E}}^{\min}$:

$$\Pi = \{(s, \alpha) \mid s \in \mathrm{E} \text{ and } \alpha \in Act_{\mathcal{N}}(s) \setminus Act^{\min}(s)\}.$$

Starting from $\mathcal{N}$ we now construct a further MDP $\mathcal{K}$ for which the idea is that $\mathcal{K}$ behaves like the end-component free MDP $\mathcal{N}$ after initially a scheduler is forced to choose a probability distribution over state-action pairs from Π. This way, *Eff* is reached with positive probability under all schedulers. Given an SPR cause C, we will observe that for the $fscore^{\mathfrak{S}}(\mathrm{C})$ under a scheduler $\mathfrak{S}$ it is only important how large the probabilities with which state action pairs from Π are chosen are relative to each other while the absolute values are not important. Due to this observation, for each SPR cause C and for each scheduler $\mathfrak{S}$ for $\mathcal{N}$ that reaches *Eff* with positive probability, we can then construct a scheduler $\mathfrak{T}$ for $\mathcal{K}$ that leads to the same recall and precision of C.

Formally, $\mathcal{K}$ is defined as follows: The state space is $\mathrm{S}_{\mathcal{N}} \cup \{init_{\mathcal{K}}\}$ where $init_{\mathcal{K}}$ is a fresh initial state. For all states in $\mathrm{S}_{\mathcal{N}}$, the same actions as in $\mathcal{N}$ are available with the same transition probabilities. i.e., for all $s, t \in \mathrm{S}_{\mathcal{N}}$

$$Act_{\mathcal{K}}(s) = Act_{\mathcal{N}}(s) \quad \text{and} \quad \mathrm{P}_{\mathcal{K}}(s, \alpha, t) = \mathrm{P}_{\mathcal{N}}(s, \alpha, t) \text{ for all } \alpha \in Act_{\mathcal{K}}(s).$$

For each state-action pair (s, α) from Π, we now add a fresh action $\beta_{(s,\alpha)}$ which is enabled only in $init_{\mathcal{K}}$. There are no further actions enabled in $init_{\mathcal{K}}$, i.e.,

$$Act_{\mathcal{K}}(init_{\mathcal{K}}) = \{\beta_{(s,\alpha)} \mid (s, \alpha) \in \Pi\}.$$

For each state $t \in \mathrm{S}_{\mathcal{N}}$, we define the transition probabilities under $\beta_{(s,\alpha)}$ by

$$\mathrm{P}_{\mathcal{K}}(init_{\mathcal{K}}, \beta_{(s,\alpha)}, t) = \mathrm{P}_{\mathcal{N}}(s, \alpha, t).$$

Claim 2: A subset $\mathrm{C} \subseteq \mathcal{S}$ such that for all $c \in \mathrm{C} : \Pr_{\mathcal{N}}^{\max}((\neg c)\,\mathrm{U}\,c) > 0$ is an SPR cause for *Eff* in $\mathcal{N}$ with $fscore(\mathrm{C}) > \vartheta$ if and only if for all schedulers $\mathfrak{T}$ for $\mathcal{K}$, we have

$$2(1-\vartheta)tp_{\mathcal{K}}^{\mathfrak{T}} - \vartheta fp_{\mathcal{K}}^{\mathfrak{T}} - \vartheta fn_{\mathcal{K}}^{\mathfrak{T}} > 0. \tag{4.12}$$

Proof of Claim 2. "⇒": Let C be an SPR cause for *Eff* in the end-component free MDP $\mathcal{N}$ with $fscore(\mathrm{C}) > \vartheta$.

We first observe that it suffices to consider schedulers $\mathfrak{T}$ that start with a deterministic choice for state $init_{\mathcal{K}}$ and then behave in an arbitrary way in order to prove (4.12) for all

schedulers $\mathfrak{T}$ for $\mathcal{K}$. To see this, we consider the MDP $\mathcal{K}_{\mathbf{C}}$ which consists of two copies of $\mathcal{K}$: "before $\mathbf{C}$" and "after $\mathbf{C}$". When $\mathcal{K}_{\mathbf{C}}$ enters a $\mathbf{C}$-state in the first copy ("before $\mathbf{C}$"), it switches to the second copy ("after $\mathbf{C}$") and stays there forever. We denote the corresponding state to $\mathfrak{s}$ in the first copy by $(\mathfrak{s},1)$ and respectively in the second copy by $(\mathfrak{s},2)$. Thus, in $\mathcal{K}_{\mathbf{C}}$ the event $\Diamond\mathbf{C} \wedge \Diamond \mathit{Eff}$ is equivalent to reaching a state $(\mathit{eff},2)$ where $\mathit{eff} \in \mathit{Eff}$. Similarly, $\Diamond\mathbf{C} \wedge \neg\Diamond \mathit{Eff}$ is equivalent to reaching a terminal non-effect state in the second copy, while $\neg\Diamond\mathbf{C} \wedge \Diamond \mathit{Eff}$ corresponds to the event of reaching an effect state in the first copy.

Obviously, there is a one-to-one-correspondence of the schedulers of $\mathcal{K}$ and $\mathcal{K}_{\mathbf{C}}$. As $\mathcal{K}$ has no end-components so does $\mathcal{K}_{\mathbf{C}}$. Therefore, a terminal state will be reached almost surely under every scheduler. Furthermore, we equip $\mathcal{K}_{\mathbf{C}}$ with a weight function on states which assigns

- weight $2(1-\vartheta)$ to the states $(\mathit{eff},2)$ where $\mathit{eff} \in \mathit{Eff}$,
- weight $-\vartheta$ to the states $(\mathit{eff},1)$ where $\mathit{eff} \in \mathit{Eff}$ and to the states $(\mathfrak{s},2)$ where $\mathfrak{s}$ is a terminal non-effect state in $\mathcal{K}$ (and $\mathcal{K}_{\mathbf{C}}$), and
- weight 0 to all other states.

Let $\mathbf{V}$ denote the set of all terminal states in $\mathcal{K}_{\mathbf{C}}$. Then, the expression on the left hand side of (4.12) is equal to $\mathrm{E}^{\mathfrak{T}}_{\mathcal{K}_{\mathbf{C}}}(\boxplus \mathbf{V})$, which is the expected accumulated weight until reaching a terminal state under scheduler $\mathfrak{T}$. Hence, (4.12) holds for all schedulers $\mathfrak{T}$ in $\mathcal{K}$ if and only if $\mathrm{E}^{\min}_{\mathcal{K}_{\mathbf{C}}}(\boxplus \mathbf{V}) > 0$.

It is well-known that the minimal expected accumulated weight in end-component-free MDPs is achieved by an MD scheduler [BK08]. Thus, there is an MD scheduler $\mathfrak{T}$ of $\mathcal{K}_{\mathbf{C}}$ with $\mathrm{E}^{\min}_{\mathcal{K}_{\mathbf{C}}}(\boxplus \mathbf{V}) = \mathrm{E}^{\mathfrak{T}}_{\mathcal{K}_{\mathbf{C}}}(\boxplus \mathbf{V})$. When viewed as a scheduler of $\mathcal{K}$, $\mathfrak{T}$ behaves memoryless deterministic before reaching $\mathbf{C}$. In particular, $\mathfrak{T}$'s initial choice in $\mathit{init}_{\mathcal{K}}$ is deterministic.

Recall the set $\mathit{Act}^{\min}(\mathfrak{s})$ of actions minimizing the reachability probability of Eff from $\mathfrak{s} \in \mathrm{D}$ and the set of state action pairs Π which leave the sub-MDP $\mathcal{N}^{\min}_{\mathrm{E}}$. Consider a scheduler $\mathfrak{T}$ for $\mathcal{K}$ with a deterministic choice $\mathfrak{T}(\mathit{init}_{\mathcal{K}}, \beta_{(\mathfrak{s},\alpha)}) = 1$ where $(\mathfrak{s},\alpha) \in \Pi$. To construct an analogous scheduler $\mathfrak{S}$ of $\mathcal{N}$, we pick an MD scheduler $\mathfrak{U}$ of the sub-MDP $\mathcal{N}^{\min}_{\mathrm{E}}$ of $\mathcal{N}$ induced by the state-action pairs $(\mathfrak{u},\beta)$ where $\mathfrak{u} \in \mathrm{E}$ and $\beta \in \mathit{Act}^{\min}(\mathfrak{u})$ such that there is a $\mathfrak{U}$-path from $\mathit{init}_{\mathcal{N}}$ to state $\mathfrak{s}$.

The general scheduler $\mathfrak{S}$ of $\mathcal{N}$ then operates with the mode $\mathfrak{m}_1$ and the modes $\mathfrak{m}_{2,\mathfrak{t}}$ for $\mathfrak{t} \in \mathrm{S}_{\mathcal{N}}$. In its initial mode $\mathfrak{m}_1$, scheduler $\mathfrak{S}$ behaves as the scheduler $\mathfrak{U}$ of the sub-MDP $\mathcal{N}^{\min}_{\mathrm{E}}$ as long as state $\mathfrak{s}$ has not been visited. Whenever state $\mathfrak{s}$ is reached in mode $\mathfrak{m}_1$ the general scheduler $\mathfrak{S}$ chooses action α with probability 1. Let $\mathfrak{t} \in \mathrm{S}_{\mathcal{N}}$ be a state that $\mathfrak{S}$ reaches via an α-transition from $\mathfrak{s}$. Then, $\mathfrak{S}$ switches to mode $\mathfrak{m}_{2,\mathfrak{t}}$ and then behaves as the residual scheduler $\mathit{res}(\mathfrak{T},\pi)$ of the deterministic scheduler $\mathfrak{T}$ of $\mathcal{K}$ for the $\mathfrak{T}$-path $\pi = \mathit{init}_{\mathcal{K}}\, \beta_{(\mathfrak{s},\alpha)}\, \mathfrak{t}$ in $\mathcal{K}$. That is, after having scheduled the action $\beta_{(\mathfrak{s},\alpha)}$, scheduler $\mathfrak{S}$ behaves exactly as $\mathfrak{T}$.

Let λ denote $\mathfrak{S}$'s probability to leave mode $\mathfrak{m}_1$, which equals $\mathfrak{U}$'s probability to reach $\mathfrak{s}$ from $\mathit{init}_{\mathcal{N}}$. Thus, $\lambda = \Pr^{\mathfrak{U}}_{\mathcal{N}}(\Diamond \mathfrak{s})$ when $\mathfrak{U}$ is viewed as a scheduler of $\mathcal{N}$. Recall the set of states E

of the sub-MDP $\mathcal{N}_{\mathrm{E}}^{\min}$ which ensures that Eff is not reached. As argued before, E is disjoint from C and Eff, scheduler $\mathfrak{S}$ stays forever in mode $\mathfrak{m}_1$ and never reaches a state in $\mathrm{C} \cup \mathit{Eff}$ with probability $1-\lambda$.

$\mathfrak{S}$ and $\mathfrak{T}$ behave identically after choosing the state-action pair $(s, \alpha) \in \Pi$ or the corresponding action $\beta_{(s,\alpha)}$, respectively, which implies that

- $\Pr_{\mathcal{N}}^{\mathfrak{S}}(\Diamond \mathrm{C} \wedge \Diamond \mathit{Eff}) = \lambda \cdot \Pr_{\mathcal{K}}^{\mathfrak{T}}(\Diamond \mathrm{C} \wedge \Diamond \mathit{Eff})$,
- $\Pr_{\mathcal{N}}^{\mathfrak{S}}(\Diamond \mathit{Eff}) = \lambda \cdot \Pr_{\mathcal{K}}^{\mathfrak{T}}(\Diamond \mathit{Eff})$ and
- $\Pr_{\mathcal{N}}^{\mathfrak{S}}(\Diamond \mathrm{C} \wedge \neg\Diamond \mathit{Eff}) = \lambda \cdot \Pr_{\mathcal{K}}^{\mathfrak{T}}(\Diamond \mathrm{C} \wedge \neg\Diamond \mathit{Eff})$.

As $\mathfrak{S}$ leaves the sub-MDP $\mathcal{N}_{\mathrm{E}}^{\min}$ with probability $\lambda > 0$, we have $\Pr_{\mathcal{N}}^{\mathfrak{S}}(\Diamond \mathit{Eff}) > 0$. By Lemma 74, we can conclude that

$$2(1-\vartheta)\mathit{tp}_{\mathcal{N}}^{\mathfrak{S}} - \vartheta \mathit{fp}_{\mathcal{N}}^{\mathfrak{S}} - \vartheta \mathit{fn}_{\mathcal{N}}^{\mathfrak{S}} > 0$$

By the equations above, this in turn implies that

$$2(1-\vartheta)\mathit{tp}_{\mathcal{K}}^{\mathfrak{T}} - \vartheta \mathit{fp}_{\mathcal{K}}^{\mathfrak{T}} - \vartheta \mathit{fn}_{\mathcal{K}}^{\mathfrak{T}} > 0$$

"$\Leftarrow$": First recall that any subset $\mathrm{C} \subseteq \mathcal{S}$ satisfying $\Pr_{\mathcal{N}}^{\max}((\neg \mathrm{c}) \,\mathrm{U}\, \mathrm{c})$ for each $\mathrm{c} \in \mathrm{C}$ is an SPR cause for Eff in $\mathcal{N}$ by definition of $\mathcal{C}$. Consider a scheduler $\mathfrak{S}$ of $\mathcal{N}$ with $\Pr_{\mathcal{N}}^{\mathfrak{S}}(\Diamond \mathit{Eff}) > 0$. Let Φ be the set of finite $\mathfrak{S}$-paths ϕ in the sub-MDP $\mathcal{N}_{\mathrm{E}}^{\min}$ such that $\mathfrak{S}$ chooses an action in $\mathit{Act}_{\mathcal{N}}(\mathit{last}(\phi)) \setminus \mathit{Act}^{\min}(\mathit{last}(\phi))$ with positive probability after ϕ where $\mathit{last}(\phi)$ denotes the last state of ϕ. For $A(s) = \mathit{Act}_{\mathcal{N}}(s) \setminus \mathit{Act}^{\min}(s)$ let

$$q = \sum_{\phi \in \Phi} \sum_{\alpha \in A(\mathit{last}(\phi))} \Pr_{\mathcal{N}}(\phi) \cdot \mathfrak{S}(\phi)(\alpha).$$

Thus, q is the overall probability that a state-action pair from Π is chosen under $\mathfrak{S}$. Consider a scheduler $\mathfrak{T}$ for $\mathcal{K}$ such that for each $\phi \in \Phi$ ending in a state s and each $\alpha \in A(s)$ an action $\beta_{(s,\alpha)}$ in $\mathit{init}_{\mathcal{K}}$ with probability $\dfrac{\Pr_{\mathcal{N}}(\gamma) \cdot \mathfrak{S}(\gamma)(\alpha)}{q}$ is chosen. Furthermore, when reaching a state t afterwards, $\mathfrak{T}$ behaves like $\mathit{res}(\mathfrak{S}, \gamma\, \alpha\, t)$ afterwards. Note that by definition this defines a probability distribution over the actions in the initial state $\mathit{init}_{\mathcal{K}}$.

By assumption, we know that now

$$(1-\vartheta)\mathit{tp}_{\mathcal{K}}^{\mathfrak{T}} - \vartheta \mathit{fp}_{\mathcal{K}}^{\mathfrak{T}} - \vartheta \mathit{fn}_{\mathcal{K}}^{\mathfrak{T}} > 0$$

As the probability with which an action $\beta_{(s,\alpha)}$ is chosen by $\mathfrak{T}$ for a $(s, \alpha) \in \Pi$ is $1/q$ times the probability that α is chosen in s to leave the sub-MDP $\mathcal{N}_{\mathrm{E}}^{\min}$ under $\mathfrak{S}$ in $\mathcal{N}$ and as the residual behavior is identical, we conclude that

$$2(1-\vartheta)\mathit{tp}_{\mathcal{N}}^{\mathfrak{S}} - \vartheta \mathit{fp}_{\mathcal{N}}^{\mathfrak{S}} - \vartheta \mathit{fn}_{\mathcal{N}}^{\mathfrak{S}} = q \cdot (2(1-\vartheta)\mathit{tp}_{\mathcal{K}}^{\mathfrak{T}} - \vartheta \mathit{fp}_{\mathcal{K}}^{\mathfrak{T}} - \vartheta \mathit{fn}_{\mathcal{K}}^{\mathfrak{T}}) > 0.$$

By Lemma 74, this shows that $\mathit{fscore}(\mathrm{C}) > \vartheta$ in $\mathcal{N}$ and finishes the proof of Claim 2.

Construction of a Game Structure. Recall the set of singleton SPR causes $\mathcal{S}$. We now construct a *stochastic shortest path game* (see [PB99]) to check whether there is a subset $\mathbf{C} \subseteq \mathcal{S}$ such that (4.12) holds in the EC-free MDP $\mathcal{K}$ in which visiting effect states always has a non-zero probability. Such a game is played on an MDP-like structure with the only difference that the set of states is partitioned into two sets indicating which player controls which states.

The game $\mathcal{G}$ has states $(S_{\mathcal{K}} \times \{yes, no\}) \cup \mathcal{S} \times \{choice\}$. On the subset $S_{\mathcal{K}} \times \{yes\}$, all actions and transition probabilities are just as in $\mathcal{K}$ and this copy of $\mathcal{K}$ cannot be left. Formally, for all $\mathbf{s}, \mathbf{t} \in S_{\mathcal{K}}$ and $\alpha \in Act_{\mathcal{K}}(\mathbf{s})$, we define $Act_{\mathcal{G}}((\mathbf{s}, yes)) = Act_{\mathcal{K}}(\mathbf{s})$ and then define $P_{\mathcal{G}}((\mathbf{s}, yes), \alpha, (\mathbf{t}, yes)) = P_{\mathcal{K}}(\mathbf{s}, \alpha, \mathbf{t})$.

In the "no"-copy, the game also behaves like $\mathcal{K}$ but when a state in $\mathcal{S}$ would be entered, the game moves to a state in $\mathcal{S} \times \{choice\}$ instead. In a state of the form $(\mathbf{c}, choice)$ with $\mathbf{c} \in \mathcal{S}$, two actions α and β are available. Choosing α leads to the state $(\mathbf{c}, yes)$ while choosing β leads to $(\mathbf{c}, no)$ both with probability 1.

Formally, this means that for all states $\mathbf{s} \in S_{\mathcal{K}}$, we define $Act_{\mathcal{G}}((\mathbf{s}, no)) = Act_{\mathcal{K}}(\mathbf{s})$ and for all actions $\alpha \in Act_{\mathcal{K}}(\mathbf{s})$:

- $P_{\mathcal{G}}((\mathbf{s}, no), \alpha, (\mathbf{t}, no)) = P_{\mathcal{K}}(\mathbf{s}, \alpha, \mathbf{t})$ for all states $\mathbf{t} \in S_{\mathcal{K}} \setminus \mathcal{S}$
- $P_{\mathcal{G}}((\mathbf{s}, no), \alpha, (\mathbf{c}, choice)) = P_{\mathcal{K}}(\mathbf{s}, \alpha, \mathbf{c})$ for all states $\mathbf{c} \in \mathcal{S}$

For states $\mathbf{s} \in S_{\mathcal{K}}$, and $\mathbf{c} \in \mathcal{S}$, we furthermore define $Act((\mathbf{c}, choice)) = \{\alpha, \beta\}$ and:

$$P_{\mathcal{G}}((\mathbf{c}, choice), \alpha, (\mathbf{c}, yes)) = P_{\mathcal{G}}((\mathbf{c}, choice), \beta, (\mathbf{c}, no)) = 1.$$

Intuitively speaking, whether a state $\mathbf{c} \in \mathcal{S}$ should belong to the cause set can be decided in the state $(\mathbf{c}, choice)$. The "yes"-copy encodes that an effect state has been selected. More concretely, the "yes-copy" is entered as soon as α has been chosen in some state $(\mathbf{c}, choice)$ and will never be left from then on. The "no"-copy of $\mathcal{K}$ then encodes that no state $\mathbf{c} \in \mathcal{S}$ visited so far has been selected to become a cause state. That is, if the current state of a play in $\mathcal{G}$ belongs to the no-copy then in all previous decisions in the states $(\mathbf{c}, choice)$, action β has been chosen.

Finally, we equip the game with a weight structure. All states in $\mathit{Eff} \times \{yes\}$ have weight $2(1-\vartheta)$, while all *terminal* states in $S_{\mathcal{K}} \times \{yes\}$ and in $\mathit{Eff} \times \{no\}$ get weight $-\vartheta$. All remaining states have weight 0.

The game is played between two players 0 and 1. Player 0 controls all states in $\mathcal{S} \times \{choice\}$ while player 1 controls the remaining states. The goal of player 0 is to ensure that the expected accumulated weight is > 0.

Claim 3: In $\mathcal{G}$ the player 0 has a winning strategy ensuring that the expected accumulated weight is > 0 if and only if there is a subset $\mathbf{C} \subseteq \mathcal{S}$ in $\mathcal{K}$ which satisfies for all $\mathbf{c} \in \mathbf{C}$: $\Pr^{\max}_{\mathcal{K}}((\neg \mathbf{c}) \,\mathrm{U}\, \mathbf{c}) > 0$ and for all schedulers $\mathfrak{T}$ for $\mathcal{K}$ we have

$$2(1-\vartheta)\, tp^{\mathfrak{T}}_{\mathcal{K}} - \vartheta\, fp^{\mathfrak{T}}_{\mathcal{K}} - \vartheta\, fn^{\mathfrak{T}}_{\mathcal{K}} > 0. \tag{4.13}$$

Proof of Claim 3. As $\mathcal{K}$ has no end-components, in the game $\mathcal{G}$ a terminal state is reached almost surely under any pair of strategies. Hence, we can rely on the results of [PB99] which state that both players have an optimal memoryless deterministic strategy.

"⇒" of Claim 3: Let ζ be a memoryless deterministic winning strategy for player 0. I.e., ζ deterministically assigns to each state in $\mathcal{S} \times \{choice\}$ an action from $\{\alpha, \beta\}$. We define

$$\mathcal{S}_\alpha = \{\mathbf{c} \in \mathcal{S} \mid \zeta((\mathbf{c}, choice)) = \alpha\}.$$

Note that $\mathcal{S}_\alpha$ is not empty as otherwise a positive expected accumulated weight in the game is not possible. (Here we use the fact that only the effect states in the yes-copy have positive weight and that the yes-copy can only be entered by taking α in one of the states $(\mathbf{c}, choice)$.) To ensure for all $\mathbf{c} \in \mathbf{C} : \Pr^{\max}_{\mathcal{K}}((\neg \mathbf{c}) \,\mathrm{U}\, \mathbf{c}) > 0$, we remove states that cannot be visited as the first state of this set:

$$\mathbf{C} = \{\mathbf{c} \in \mathcal{S}_\alpha \mid \mathcal{K}, \mathbf{c} \models \exists(\neg \mathcal{C}_\alpha) \,\mathrm{U}\, \mathbf{c}\}.$$

Note that all memoryless deterministic strategies for player 0 in $\mathcal{G}$ which correspond to the sets $\mathcal{S}_\alpha$ and $\mathbf{C}$ lead to exactly the same plays.

Let $\mathfrak{T}$ be a scheduler for $\mathcal{K}$. This scheduler can be used as a strategy for player 1 in $\mathcal{G}$. We denote the expected accumulated weight when player 0 plays according to ζ and player 1 plays according to $\mathfrak{T}$ by $w(\zeta, \mathfrak{T})$. As ζ is winning for player 0 we have $w(\zeta, \mathfrak{T}) > 0$. By the construction of the game, it follows directly that

$$w(\zeta, \mathfrak{T}) = 2(1-\vartheta) tp^{\mathfrak{T}}_{\mathcal{K}} - \vartheta fp^{\mathfrak{T}}_{\mathcal{K}} - \vartheta fn^{\mathfrak{T}}_{\mathcal{K}}.$$

Putting things together yields:

$$2(1-\vartheta) tp^{\mathfrak{T}}_{\mathcal{K}} - \vartheta fp^{\mathfrak{T}}_{\mathcal{K}} - \vartheta fn^{\mathfrak{T}}_{\mathcal{K}} > 0. \tag{4.14}$$

"⇐": Suppose there is a set $\mathbf{C} \subseteq \mathcal{S}$ which satisfies $\Pr^{\max}_{\mathcal{K}}((\neg \mathbf{c}) \,\mathrm{U}\, \mathbf{c}) > 0$ for all $\mathbf{c} \in \mathbf{C}$ and (4.13) for all schedulers $\mathfrak{T}$ for $\mathcal{K}$. We define the MD-strategy ζ from $\mathbf{C}$ by letting $\zeta((\mathbf{c}, choice)) = \alpha$ if and only if $\mathbf{c} \in \mathbf{C}$. For any strategy $\mathfrak{T}$ for player 1, we can again view $\mathfrak{T}$ also as a scheduler for $\mathcal{K}$. Equation (4.14) holds again and shows that the expected accumulated weight in $\mathcal{G}$ is positive if player 0 plays according to ζ against any strategy for player 1. This finishes the proof of Claim 3.

Putting together Claims 1-3. We conclude that there is an SPR cause $\mathbf{C}$ in the original MDP $\mathcal{M}$ with $fscore(\mathbf{C}) > \vartheta$ if and only if player 1 has a winning strategy in the constructed game $\mathcal{G}$. As both players have optimal MD-strategies in $\mathcal{G}$ [PB99], the decision problem is in $\mathrm{NP} \cap \mathrm{coNP}$: We can guess the MD-strategy for player 0 and solve the resulting stochastic shortest path problem in polynomial time [BT91] to obtain an NP-upper bound. Likewise, we can guess the MD-strategy for player 1 and solve the resulting stochastic shortest path problem to obtain the coNP-upper bound. □

4.3.2 Quality-Optimal GPR Causes

We now turn to the problem of computing quality-optimal GPR causes. Similar to checking the conditions (Section 4.1) it turns out that computing quality-optimal GPR causes is harder than for SPR causes. In particular, both the recall-optimal SPR cause (Thm. 70) and the method for computing f-score-optimal SPR causes (Thm. 73 and Thm. 75) are reliant on the fact that the (SPR) condition **(S)** is stable under union. This is not the case for GPR causes and thus there is no analogous canonical GPR cause. So, the techniques for computing quality-optimal SPR causes are not applicable here. Therefore, the majority of this section will focus on the corresponding threshold problems for each quality measure.

Remark 76 (Computing Quality-Optimal GPR Causes). Computing optimal GPR causes for any quality measure can be done by considering all cause candidates, checking **(G)** in coNP and computing the corresponding quality measure in polynomial time (Section 4.2). As the space can be reused after each computation, this results in polynomial space. ◁

However, we show that no polynomial-time algorithms can be expected as the corresponding threshold problems are NP-hard. For this, we let GPR-recall (resp. GPR-covratio and GPR-f-score) denote the decision problems:

GPR-recall Input: MDP $\mathcal{M} = (\mathrm{S}, Act, \mathrm{P}, init)$ with set of terminal states $\mathit{Eff} \subset \mathrm{S}$ and $\vartheta \in \mathbb{Q}$
Question: does there exists a GPR cause Cause with $\mathit{recall}(\mathit{Cause}) \geqslant \vartheta$?

GPR-covratio Input: MDP $\mathcal{M} = (\mathrm{S}, Act, \mathrm{P}, init)$ with set of terminal states $\mathit{Eff} \subset \mathrm{S}$ and $\vartheta \in \mathbb{Q}$
Question: does there exists a GPR cause Cause with $\mathit{covrat}(\mathit{Cause}) \geqslant \vartheta$?

GPR-f-score Input: MDP $\mathcal{M} = (\mathrm{S}, Act, \mathrm{P}, init)$ with set of terminal states $\mathit{Eff} \subset \mathrm{S}$ and $\vartheta \in \mathbb{Q}$
Question: does there exists a GPR cause Cause with $\mathit{fscore}(\mathit{Cause}) \geqslant \vartheta$?

Theorem 77 (Complexity of Quality-Optimal GPR Causes). *The problems GPR-recall, GPR-covratio and GPR-f-score are NP-hard and belong to Σ_2^{P}. For Markov chains, all three problems are NP-complete. NP-hardness even holds for tree-like Markov chains.*

Proof. *Σ_2^{P}-membership.* The algorithms for GPR-covratio, GPR-recall and GPR-f-score rely on the guess-and-check principle. First a set $\mathit{Cause} \subseteq \mathrm{S}$ is guessed non-deterministically. Then the GPR check for Cause can be done coNP by Theorem 53. Lastly $\mathit{quality}(\mathit{Cause}) \geqslant \vartheta$ can be checked in polynomial time using Lemma 60 and Theorem 67. The alternation between the existential quantification for guessing Cause and the universal quantification for the coNP check of the GPR condition results in the complexity Σ_2^{P} of the polynomial-time hierarchy.

NP-membership for Markov chains. If $\mathcal{M}$ is a Markov chain we can non-deterministically guess a set Cause and then check whether Cause is a GPR cause in polynomial time by computing the corresponding conditional probabilities. The computation of the quality measure $\mathit{quality}(\mathit{Cause})$ can also be done in polynomial time by using Remark 27. Therefore, the threshold problem can be decided in NP.

NP-hardness of GPR-recall and GPR-covratio. With arguments as in the proof of Lemma 60, the problems GPR-recall and GPR-covratio are polynomially interreducible for Markov chains. Thus, it suffices to prove NP-hardness of GPR-recall. For this, we provide a polynomial reduction from the knapsack problem which is known to be NP-complete [GJ90]. The input of the knapsack problem are sequences $A_1, \ldots, A_n, A$ and $B_1, \ldots, B_n, B$ of positive natural numbers and the task is to decide whether there exists a subset I of $\{1, \ldots, n\}$ such that

$$\sum_{i \in I} A_i \; < \; A \qquad \text{and} \qquad \sum_{i \ in I} B_i \; \geqslant \; B. \tag{4.15}$$

We start by transforming this problem into an equivalent form which we will then encode into a Markov chain. For this, let K be the maximum of $A, A_1, \ldots, A_n, B, B_1, \ldots, B_n$ and let $N = 8(n+1) \cdot (K+1)$. We define

$$a_i = \frac{A_i}{N}, \qquad a = \frac{A}{N}, \qquad b_i = \frac{B_i}{N}, \qquad b = \frac{B}{N}.$$

Therefore, $a, a_1, \ldots, a_n, b, b_1, \ldots, b_n$ are positive rational numbers strictly smaller than $\frac{1}{8(n+1)}$ and the knapsack problem (4.15) can be rewritten as:

$$\sum_{i \in I} a_i \; < \; a \qquad \text{and} \qquad \sum_{i \in I} b_i \; \geqslant \; b \tag{4.16}$$

For $i \in \{1, \ldots, n\}$, let

$$p_i = 2(a_i + b_i) \quad \text{and} \quad w_i = \frac{b_i}{p_i} = \frac{1}{2} \cdot \frac{b_i}{a_i + b_i},$$

which establishes, $0 < p_i < \frac{1}{2(n+1)}$ and $0 < w_i < \frac{1}{2}$. Moreover, $p_i\big(\frac{1}{2} - w_i\big) = a_i$ and $p_i \cdot w_i = b_i$. Hence, we can rewrite the knapsack problem (4.16) further as:

$$\sum_{i \in I} p_i\big(\frac{1}{2} - w_i\big) \; < \; a \qquad \text{and} \qquad \sum_{i \in I} p_i w_i \; \geqslant \; b$$

which again is equivalent to:

$$\frac{\sum\limits_{i \in I_0} p_i w_i}{\sum\limits_{i \ in I_0} p_i} \; > \; \frac{1}{2} \qquad \text{and} \qquad \sum_{i \in I_0} p_i w_i \; \geqslant \; p_0 + b, \tag{4.17}$$

where $p_0 = 2a$, $w_0 = 1$ and $I_0 = I \cup \{0\}$. Note that $a < \frac{1}{8(n+1)}$ and hence $p_0 < \frac{1}{4(n+1)}$.

Define a tree-shape Markov chain $\mathcal{M}$ with non-terminal states init, $s_0, s_1, \ldots, s_n$, and terminal states $\mathit{eff}_0, \ldots, \mathit{eff}_n$, FN and $\mathit{noeff}, \mathit{noeff}_1, \ldots, \mathit{noeff}_n$. Transition probabilities are:

- $P(\mathit{init}, s_i) = p_i$ for $i = 0, \ldots, n$,
- $P(\mathit{init}, \mathit{FN}) = \frac{1}{2} - \sum\limits_{i=0}^{n} p_i w_i$,
- $P(\mathit{init}, \mathit{noeff}) = 1 - \sum\limits_{i=0}^{n} p_i - P(\mathit{init}, \mathit{FN})$,

- $P(s_i, \mathit{eff}_i) = w_i$, $P(s_i, \mathit{noeff}_i) = 1 - w_i$ for $i = 1, \ldots, n$,
- $P(s_0, \mathit{eff}_0) = 1 = w_0$.

Note that $p_0 + p_1 + \ldots + p_n < \frac{1}{2}$ as all p_i's are strictly smaller than $\frac{1}{2(n+1)}$. As the w_i's are bounded by 1, this yields $0 < P(\mathit{init}, \mathit{FN}) < \frac{1}{2}$ and $0 < P(\mathit{init}, \mathit{noeff}) < 1$.

The graph structure of $\mathcal{M}$ is a tree and $\mathcal{M}$ can be constructed from the values $A, A_1, \ldots, A_n$ and $B, B_1, \ldots, B_n$ in polynomial time. For $\mathit{Eff} = \{\mathit{eff}_i : i = 0, 1, \ldots, n\} \cup \{\mathit{FN}\}$ we have:

$$\Pr_{\mathcal{M}}(\Diamond \mathit{Eff}) \;=\; \sum_{i=0}^{n} p_i w_i + P(\mathit{init}, \mathit{FN}) \;=\; \frac{1}{2}.$$

As the values $w_1, \ldots, w_n$ are strictly smaller than $\frac{1}{2}$, we have $\Pr_{\mathcal{M}}(\Diamond \mathit{Eff} \mid \Diamond C) < \frac{1}{2}$ for each nonempty subset $C \subseteq \{s_1, \ldots, s_n\}$. Thus, the only candidates for GPR causes are the sets $C_I = \{s_i : i \in I_0\}$ where $I \subseteq \{1, \ldots, n\}$ where as before $I_0 = I \cup \{0\}$. For all states $s \in C_I$ there is a path satisfying $((\neg c)_I) \,\mathrm{U}\, s$. Thus, C_I is a GPR cause iff C_I satisfies (G) . We have:

$$\Pr_{\mathcal{M}}(\Diamond \mathit{Eff} \mid \Diamond C_I) = \frac{\sum_{i \in I_0} p_i w_i}{\sum_{i \in I_0} p_i} \quad \text{and}$$

$$\mathit{recall}(C_I) = \Pr_{\mathcal{M}}(\Diamond C_I \wedge \Diamond \mathit{Eff} \mid \Diamond \mathit{Eff}) = 2 \cdot \sum_{i \in I_0} p_i w_i.$$

Thus, C_I is a GPR cause with recall at least $2(p_0 + b)$ if and only if the two conditions in (4.17) hold, which is equivalent to the satisfaction of the conditions in (4.15). But this yields that $\mathcal{M}$ has a GPR cause with recall at least $2(p_0 + b)$ if and only if the knapsack problem is solvable for the input $A, A_1, \ldots, A_n, B, B_1, \ldots, B_n$.

NP-hardness of GPR-f-score. Using similar ideas, we also provide a polynomial reduction from the knapsack problem. Let $A, A_1, \ldots, A_n, B, B_1, \ldots, B_n$ be an input for the knapsack problem. We replace the A-sequence with $a, a_1, \ldots, a_n$ where $a = \frac{A}{N}$ and $a_i = \frac{A_i}{N}$ where N is as before. The topological structure of the Markov chain that we are going to construct is the same as in the NP-hardness proof for GPR-recall.

We will define polynomial-time computable values $p_0, p_1, \ldots, p_n \in (0, 1)$ ($p_i = P(\mathit{init}, s_i)$ and for $w_1, \ldots, w_n \in (0, 1)$, $w_i = P(s_i, \mathit{eff}_i)$) and auxiliary variables $\delta \in (0, 1)$ and $\lambda > 1$ such that:

(1) $p_0 + p_1 + \ldots + p_n < \frac{1}{2}$

(2) $\lambda = \frac{p_0 + \frac{1}{2} - \delta}{p_0}$

(3) for all $i \in \{1, \ldots, n\}$:

  (3.1) $a_i = p_i\big(\frac{1}{2} - w_i\big)$ (in particular $w_i < \frac{1}{2}$)

  (3.2) $B_i = \frac{1}{\delta} B p_i \big(\lambda w_i - 1\big)$ (in particular $w_i > \frac{1}{\lambda}$)

Assuming such values have been defined, we obtain:

$$\begin{aligned}\sum_{i\in I} B_i \geqslant B &\iff \frac{1}{\delta}B\sum_{i\in I} p_i(\lambda w_i - 1) \geqslant B\\ &\iff \sum_{i\in I} p_i(\lambda w_i - 1) \geqslant \delta\\ &\iff \lambda\sum_{i\in I} p_i w_i \geqslant \delta + \sum_{i\in I} p_i\end{aligned}$$

Hence:

$$\sum_{i\in I} B_i \geqslant B \iff \frac{\sum_{i\in I} p_i w_i}{\delta + \sum_{i\in I} p_i} \geqslant \frac{1}{\lambda}.$$

For all positive real numbers x, y, u, v with $\frac{x}{y} = \frac{1}{\lambda}$ we have:

$$\frac{x+u}{y+v} \geqslant \frac{1}{\lambda} \qquad \iff \qquad \frac{u}{v} \geqslant \frac{1}{\lambda}$$

By the constraints for λ (see (2)), we have $\frac{p_0}{p_0+\frac{1}{2}-\delta} = \frac{1}{\lambda}$. Therefore:

$$\frac{\sum_{i\in I} p_i w_i}{\delta + \sum_{i\in I} p_i} \geqslant \frac{1}{\lambda} \quad \iff \quad \frac{p_0 + \sum_{i\ in I} p_i w_i}{(p_0 + \frac{1}{2} - \delta) + \delta + \sum_{i\in I} p_i} = \frac{p_0 + \sum_{i\in I} p_i w_i}{p_0 + \frac{1}{2} + \sum_{i\ in I} p_i} \geqslant \frac{1}{\lambda}.$$

As before let $w_0 = 1$ and $I_0 = I \cup \{0\}$. Then, the above yields:

$$\sum_{i\in I} B_i \geqslant B \quad \iff \quad \frac{\sum_{i\in I_0} p_i w_i}{\frac{1}{2} + \sum_{i\in I_0} p_i} \geqslant \frac{1}{\lambda}.$$

As in the NP-hardness proof for GPR-recall and using (3.1):

$$Pr_{\mathcal{M}}(\Diamond \mathit{Eff}) = \frac{1}{2} > w_i \qquad \text{for } i = 1, \ldots, n$$

Thus, each GPR cause must have the form $C_I = \{s_i : i \in I_0\}$ for some subset I of $\{1, \ldots, n\}$. Moreover:

$$Pr_{\mathcal{M}}(\Diamond C_I) = \sum_{i\in I_0} p_i \qquad \text{and} \qquad Pr_{\mathcal{M}}(\Diamond C_I \wedge \Diamond \mathit{Eff}) = \sum_{i\in I_0} p_i w_i$$

So, the f-score of C_I is:

$$\mathit{fscore}(C_I) = 2 \cdot \frac{Pr_{\mathcal{M}}(\Diamond C_I \wedge \Diamond \mathit{Eff})}{Pr_{\mathcal{M}}(\Diamond \mathit{Eff}) + Pr_{\mathcal{M}}(\Diamond C_I)} = 2 \cdot \frac{\sum_{i\in I_0} p_i w_i}{\frac{1}{2} + \sum_{i\in I_0} p_i}$$

This implies:

$$\sum_{i\in I} B_i \geqslant B \quad \iff \quad fscore(C_I) \geqslant \frac{2}{\lambda}$$

With $p_0 = 2a$ and using (3.1) and arguments as in the NP-hardness proof for GPR-recall, we obtain:

$$\sum_{i\in I} A_i < A \quad \iff \quad C_I \text{ is a GPR cause}$$

Thus, the constructed Markov chain has a GPR cause $Cause$ with $fscore(Cause) \geqslant \frac{2}{\lambda}$ if and only if the knapsack problem is solvable for the input $A, A_1, \ldots, A_n, B, B_1, \ldots, B_n$.

It remains to define the values $p_1, \ldots, p_n, w_1, \ldots, w_n$ and δ. (The value of λ is then obtained by (2).) (3.1) and (3.2) can be rephrased as equations for w_i:

(3.1′) $w_i = \frac{1}{2} - \frac{a_i}{p_i}$

(3.2′) $w_i = \frac{1}{\lambda}\left(\delta \frac{B_i}{B p_i} + 1\right)$

This yields an equation for p_i:

$$\frac{1}{2} - \frac{a_i}{p_i} = \frac{1}{\lambda}\left(\delta \frac{B_i}{B p_i} + 1\right)$$

and leads to:

$$p_i = \frac{2\lambda}{\lambda - 2} a_i + \frac{2\delta}{\lambda - 2} \cdot \frac{B_i}{B}. \tag{4.18}$$

We now substitute λ by (2) and arrive at

$$p_i = \frac{p_0}{\frac{1}{2} - \delta} a_i + a_i + \frac{\delta p_0}{\frac{1}{2} - \delta} \frac{B_i}{B}.$$

By choice of N, all a_i's and a are smaller than $\frac{1}{8(n+1)}$. Using this together with $p_0 = 2a$, we get:

$$p_i \;<\; \frac{1}{4(n+1)(\frac{1}{2} - \delta)} \frac{1}{8(n+1)} \;+\; \frac{1}{8(n+1)} \;+\; \frac{\delta}{4(n+1)(\frac{1}{2} - \delta)} \frac{B_i}{B} \tag{4.19}$$

Let now $\delta = \frac{1}{8K}$ (where K is as above, i.e., the maximum of the values $A, A_1, \ldots, A_n$, B, $B_1, \ldots, B_n$). Then, $p_1, \ldots, p_n$ are computable in polynomial time, and so are the values $w_1, \ldots, w_n$ (by (3.1′)). As $\frac{2\lambda}{\lambda - 2} > 2$ and using (4.18), we obtain $p_i > 2a_i$. So, by (3.1′) we get $0 < w_i < \frac{1}{2}$.

It remains to prove (1). Using $\delta = \frac{1}{8K}$, we obtain from (4.19):

$$p_i \;<\; \frac{1}{4(n+1)(\frac{1}{2} - \frac{1}{8K})} \frac{1}{8(n+1)} + \frac{1}{8(n+1)} + \frac{1}{32(n+1)(\frac{1}{2} - \frac{1}{8K})K} \frac{B_i}{B} \;=\; x$$

As $\frac{1}{2} - \frac{1}{8K} \geqslant \frac{1}{4}$ and $\frac{B_i}{B} < K$, this yields:

$$p_i \;<\; x \;<\; \frac{1}{8(n+1)^2} + \frac{1}{8(n+1)} + \frac{1}{8(n+1)} \;<\; \frac{1}{2(n+1)}.$$

But then condition (1) holds which finishes the reduction. □

Similarly to the complexity of checking the PR conditions we again notice that computing quality-optimal PR causes is harder if the global PR condition is considered in comparison to the strict PR condition. However, this time it is already quite involved to compute quality-optimal SPR causes for a more complex quality measure like the fscore. It is already in $\mathrm{NP} \cap \mathrm{coNP}$ to decide the threshold problem for f-score optimal SPR causes. Moreover, for quality-optimal GPR causes their complexity lower-bound falls together. Here, the complexity for the computation does not seem to lie in the considered quality-measure but rather in the probability-raising condition which governs the candidate sets for quality-optimality. Overall, we can conclude that considering the GPR condition (G) as a basis for probabilistic causality in MDPs is algorithmically more involved or outright harder than the SPR condition (S) . However, in the following chapter we will see that for extensions to more general settings (where the effect can be an ω-regular property) the complexity of GPR and SPR will coincide again.

Arbitrary Quality Measures.

As a conclusion of this chapter, we like to briefly consider more general quality measures. An example of this has already been discussed in Section 3.3 with Matthews correlation coefficient (MCC). For this let $\mathfrak{f}(\mathit{tp}, \mathit{tn}, \mathit{fp}, \mathit{fn})$ be an algebraic function. That is, $\mathfrak{f}$ satisfies some polynomial equation where the coefficients are polynomials in $\mathit{tp}, \mathit{tn}, \mathit{fp}$ and fn. Almost every quality measure for binary classifiers (see [Pow11]) is such a function. Taking the worst case scheduler for such a function we define

$$\mathfrak{f}(\mathit{Cause}) = \inf_{\mathfrak{S}} \mathfrak{f}^{\mathfrak{S}}(\mathit{tp}^{\mathfrak{S}}_{\mathit{Cause}}, \mathit{tn}^{\mathfrak{S}}_{\mathit{Cause}}, \mathit{fp}^{\mathfrak{S}}_{\mathit{Cause}}, \mathit{fn}^{\mathfrak{S}}_{\mathit{Cause}}),$$

where $\mathfrak{S}$ ranges over all schedulers such that $\mathfrak{f}^{\mathfrak{S}}$ is well defined. Given a PR cause Cause and a rational $\vartheta \in \mathbb{Q}$, deciding whether $\mathfrak{f}(\mathit{Cause}) \geqslant \vartheta$ can be done in PSPACE as a satisfiability problem in the existential theory of the reals [Can88]. This gives an upper bound for the computation of any general quality measure.

Moreover, since we can decide for a given cause candidate Cause whether it is a SPR cause in P or a GPR cause in coNP, this also yields an algorithm for finding optimal causes for $\mathfrak{f}$. Given an MDP $\mathcal{M}$ with terminal effect set Eff and quality measure $\mathfrak{f}$ as an algebraic function we consider each cause candidate Cause, check whether it is a PR cause (SPR or GPR) and consider the decision problem $\mathfrak{f}(\mathit{Cause}) \geqslant \vartheta$. As all of these steps have a complexity upper bound of PSPACE and we only need to save the best cause candidate so far with its value $\mathfrak{f}(\mathit{Cause})$. This results in a polynomial space algorithm for the computation of a quality-optimal cause for a general quality measure which coincides with our computational upper bound for recall- or fscore-optimal GPR causes .

5 Variants of Probability-Raising Causality in MDPs

By now we have established a notion of probability-raising causes in MDPs in which both the effect and the cause are reachability properties of sets of states in the MDP (Def. 14). We extensively studied this notion in both meaning (Sec. 3.1,3.2) and algorithmic properties (Ch. 4). In this chapter we discuss variants and extensions of this state-based notion of probability-raising cause. This includes algorithmic studies similar to Chapter 4. So, for each variant of PR causality we address the complexities of checking the resulting PR conditions as well as computing quality-optimal causes. The variants discussed in this chapter modify the notion of PR cause in two different aspects, which we will outline here.

The first aspect in consideration is the quantification over the scheduler of the MDP. We will decouple the scheduler from the condition **(G)** notion and from condition **(S)** for SPR causes and thus define *probability-raising scheduler*. That is, we define schedulers for which the (GPR) or (SPR) conditions hold. This gives rise to a variant of PR causality in which only the existence of a probability-raising scheduler is required. This way, the PR conditions can hold *potentially*, as there are scenarios in which the cause is indeed raising the probability of the effect.

The second aspect in consideration is the type of the effect and cause events. We have already shown in Remark 13 that more general types of an effect event can be reduced to a reachability property. This includes a setting in which the effect is a regular language given by a deterministic finite automaton. In a more general approach, we consider the effect to be an ω-regular language over the state space S given by a deterministic Rabin automaton (Def. 10). While a reduction from ω-regular effect settings to the state-based setting is still possible, we show that this has a considerable impact on the complexity of checking and deriving causes. Furthermore, we also generalize the description of the cause event to be ω-regular co-safety properties over the state space S. Again, this extension

comes natural as the reachability property of the a non-terminal set Cause in an MDP is a special case of a co-safety property. As temporal priority still needs to be ensured by the resulting notion of PR causality, the choice of co-safety properties supports this as any word of a co-safety language can be fully characterized by some finite prefix. While this allows us to express much more involved cause-effect relationships, we outline that attempts of checking co-safety SPR causality or of finding good causes for a given effect encounter major new difficulties.

5.1 Probability-Raising Scheduler and Potential PR Causes

In this section we consider the established notion of probability-raising cause (Def. 14) with state-based events but switch the quantification over scheduler. So, instead of requiring probability-raising causality under all schedulers of an MDP, we only require the existence of a scheduler for which probability-raising holds. This consideration makes sense, since intuitively the non-determinism of the system described by the MDP is sometimes under control, for example, by a user. The contents of this section are not part of any current publication and build on the definitions of [BPZ24].

The GPR- and SPR causes described by Definition 14 are *universal* in the sense that the probability-raising is independent of schedulers. Here, we define a causal link which can hold *potentially* as it may also be desirable to have a notion of causality which takes the perspective of scheduler. Thus, we call PR causes by Definition 14 *universal* within this section to distinguish these causes from the notion introduced here. Before we go into the resulting definition of potential PR causes, let us establish the notion of a *probability-raising scheduler*, which is, in simple terms, a scheduler which witnesses the corresponding PR condition.

Throughout this section we are given an MDP $\mathcal{M} = (\mathbf{S}, \mathit{Act}, \mathbf{P}, \mathit{init})$ and a non-empty set of terminal states $\mathit{Eff} \subset \mathbf{S} \setminus \{\mathit{init}\}$.

Definition 78 (Probability-Raising Scheduler). Given $\mathcal{M}$ and Eff as above and a non-empty set $\mathit{Cause} \subseteq \mathbf{S} \setminus \mathit{Eff}$. A scheduler $\mathfrak{S}$ of $\mathcal{M}$ is a *global probability-raising scheduler* (short GPR scheduler) for Cause and Eff in $\mathcal{M}$ if $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Cause}) > 0$ and (GPR) holds, i.e.

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\, \Diamond \mathit{Eff} \mid \Diamond \mathit{Cause} \,) \;>\; \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff}). \tag{GPR}$$

The scheduler $\mathfrak{S}$ is a *strict probability-raising scheduler* (short SPR scheduler) for Cause and Eff in $\mathcal{M}$ if $\Pr^{\mathfrak{S}}(\Diamond \mathit{Cause}) > 0$ and for each state $\mathbf{c} \in \mathit{Cause}$ with $\Pr^{\mathfrak{S}}_{\mathcal{M}}((\neg \mathit{Cause}) \,\mathrm{U}\, \mathbf{c}) > 0$ and (SPR) holds, i.e.

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\, \Diamond \mathit{Eff} \mid (\neg \mathit{Cause}) \,\mathrm{U}\, \mathbf{c} \,) \;>\; \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff}). \tag{SPR}$$

◁

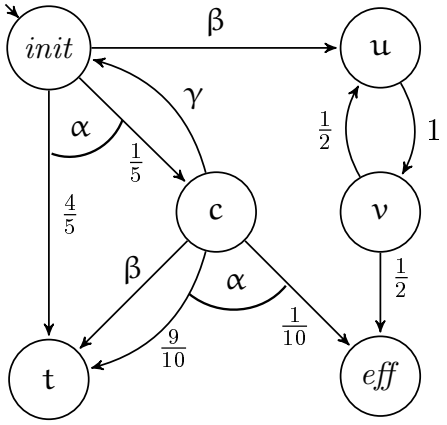

Figure 5.1: MDP $\mathfrak{M}$ from Example 80

A set of states Cause with $\Pr^{\max}((\neg \mathit{Cause})\,\mathrm{U}\,\mathbf{c}) > 0$ for each $\mathbf{c} \in \mathit{Cause}$ is a universal SPR cause if and only if every scheduler $\mathfrak{S}$ of $\mathfrak{M}$ is an SPR scheduler for Cause and Eff in $\mathfrak{M}$. Similarly, Cause is a universal GPR cause for Eff in $\mathfrak{M}$ if and only if every scheduler $\mathfrak{S}$ of $\mathfrak{M}$ with $\Pr^{\mathfrak{S}}(\Diamond \mathit{Cause}) > 0$ is a GPR scheduler for Eff in $\mathfrak{M}$.

Definition 79 (Potential PR Cause). We define a set $\mathit{Cause} \subseteq \mathrm{S} \setminus \mathit{Eff}$ to be a *potential GPR cause* for Eff in $\mathfrak{M}$ if there exists a GPR scheduler for Cause and Eff in $\mathfrak{M}$. Analogously, Cause is a *potential SPR cause* for Eff if there exists an SPR scheduler for Cause and Eff in $\mathfrak{M}$. ◁

For the remainder of the section, we discuss this notion with respect to the already established notion of universal PR causes in MDPs.

First, we note that we do not consider the minimality condition $\Pr^{\max}((\neg \mathit{Cause})\,\mathrm{U}\,\mathbf{c}) > 0$ for each $\mathbf{c} \in \mathit{Cause}$ here, as for potential PR causes it is not relevant whether all states can be reached independently. Thus, the relevant states $\mathbf{c} \in \mathit{Cause}$ for the (PR) condition depend on the scheduler and it might happen that for a potential PR cause there are states $\mathbf{c} \in \mathit{Cause}$ such that there is no PR scheduler $\mathfrak{S}$ satisfying $\Pr^{\mathfrak{S}}((\neg \mathit{Cause})\,\mathrm{U}\,\mathbf{c}) > 0$. Furthermore, it is obvious that Cause being a potential PR cause for Eff is a strictly weaker condition than when Cause is a universal PR cause satisfying (G) or (S) .

Example 80 (PR scheduler). For an example we consider the MDP $\mathfrak{M}$ from Figure 5.1 with $\mathit{Cause} = \{\mathbf{c}\}$ and $\mathit{Eff} = \{\mathit{eff}\}$. For the MD scheduler $\mathfrak{S}$ of $\mathfrak{M}$ which chooses $\boldsymbol{\alpha}$ in init and in $\mathbf{c}$ we have

$$\Pr^{\mathfrak{S}}(\Diamond \mathit{eff} \mid \Diamond \mathbf{c}) = \frac{1}{10} > \frac{1}{5} \cdot \frac{1}{10} = \Pr^{\mathfrak{S}}(\Diamond \mathit{eff})$$

and thus $\mathfrak{S}$ satisfies (GPR) (and (SPR)) and $\mathit{Cause} = \{\mathbf{c}\}$ is a potential GPR cause. However, there are other schedulers (like the MD scheduler $\mathfrak{U}$ choosing $\boldsymbol{\alpha}$ in init and $\boldsymbol{\beta}$ in $\mathbf{c}$) which do not satisfy (SPR). ◁

The difference in the quantification of scheduler imposed by potential PR causes does not make a difference on the relation between SPR and GPR. As the argumentation of Proposition 15 is considering the individual scheduler, it also holds that an SPR scheduler $\mathfrak{S}$ for Cause and Eff in $\mathcal{M}$ is also a GPR scheduler. Similarly, for singletons $\{\mathbf{c}\}$ the equivalence of $\Diamond\mathbf{c}$ and $(\neg\mathbf{c})\,\mathrm{U}\,\mathbf{c}$ also means that SPR scheduler for $\{\mathbf{c}\}$ and Eff in $\mathcal{M}$ are GPR scheduler and vice versa.

Furthermore, if $\mathcal{M}$ is a DTMC, i.e. in each state $\mathbf{s} \in \mathbf{S}$ there is exactly one enabled action $\alpha \in \mathit{Act}(\mathbf{s})$, then Definitions 14 and 79 are equivalent modulo the minimality condition. As there is only one scheduler $\mathfrak{S}$ for $\mathcal{M}$ a set Cause with $\Pr((\neg \mathit{Cause})\,\mathrm{U}\,\mathbf{c}) > 0$ is a potential SPR cause for Eff if and only if it is a universal SPR cause for Eff. In both cases we have $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Cause}) > 0$ and (SPR) holds for $\mathfrak{S}$. Analogously, such a set Cause is a potential GPR cause for Eff if and only if it is a universal GPR cause for Eff. By Example 16 this means that GPR scheduler (resp. potential GPR causes) are more general than SPR scheduler (resp. potential SPR causes).

Potential PR Causes in Racetrack

We will now return to the Racetrack MDP (Section 3.2.2) as an example of a system where considering potential PR causality makes more sense as universal PR causality. For this, we look at the same track as before (Figure 3.10). Therefore, recall from Figure 3.12 that in $\mathcal{M}$ each state has four coordinates $(\mathbf{x}, \mathbf{y}, \mathbf{v}_\mathbf{x}, \mathbf{v}_\mathbf{y})$ specifying the location and speed of the car on the track and there is one state specifying a crash into the wall. Each non-crashed state has some enabled actions $(\mathbf{a}_\mathbf{x}, \mathbf{a}_\mathbf{y})$ for some $\mathbf{a}_\mathbf{x}, \mathbf{a}_\mathbf{y} \in \{(-1,\ 0,\ 1)\}$ specifying an acceleration of the car. With a probability of $1 - \mathbf{p}$ the system updates both location and velocity according to the action and with probability $\mathbf{p}$ it does not change the current velocity and just updates the location according to the current velocity.

We again consider the effect $\mathit{Eff} = \{\mathit{goal}\} = \{(3, 2, \mathbf{v}_\mathbf{x}, \mathbf{v}_\mathbf{y}) \mid \mathbf{v}_\mathbf{x}, \mathbf{v}_\mathbf{y} \in \{-2,\ -1,\ 0,\ 1,\ 2\}\}$. As a cause candidate we again consider the set $\mathbf{C}_1$ specified visually in Figure 3.13 and in formula by

$$\mathbf{C}_1 = \{(1, 2, \mathbf{v}_\mathbf{x}, \mathbf{v}_\mathbf{y}) \mid \mathbf{v}_\mathbf{x}, \mathbf{v}_\mathbf{y} \in \{-2, -1, 0, 1, 2\}\}.$$

We already established that $\mathbf{C}_1$ is not a (universal) GPR or SPR cause. However, consider the following MD scheduler $\mathfrak{S}$ of $\mathcal{M}$:

- In start, $\mathfrak{S}$ chooses action $(1, 0)$ and reaches $\mathbf{s}_1 = (1, 0, 1, 0)$ with probability 1.
- In $\mathbf{s}_1$, $\mathfrak{S}$ chooses $(-1, 1)$ resulting in a probability of $(1 - \mathbf{p})$ to reach $\mathbf{s}_2 = (1, 1, 0, 1)$ and a probability of $\mathbf{p}$ to reach crash.
- In $\mathbf{s}_2$, $\mathfrak{S}$ chooses $(0, 0)$ which transitions to $\mathbf{c}_1 = (1, 2, 0, 1) \in \mathbf{C}_1$ with probability 1.
- Starting from $\mathbf{c}_1$, $\mathfrak{S}$ maximizes the probability to reach goal.

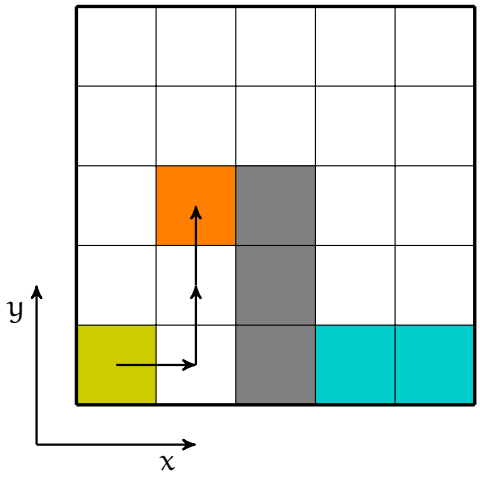

Figure 5.2: The path in the Racetrack which the scheduler $\mathfrak{S}$ tries to realize

This fully specifies $\mathfrak{S}$ and is visualized in Figure 5.2. We thus have

$$\Pr^{\mathfrak{S}}(\Diamond \mathit{Eff} \mid \Diamond \mathrm{C}_1) = \frac{\Pr^{\mathfrak{S}}(\Diamond \mathit{goal} \wedge \Diamond \mathrm{C}_1)}{\Pr^{\mathfrak{S}}(\Diamond \mathrm{C}_1)} = \frac{\Pr^{\mathfrak{S}}(\Diamond \mathit{goal})}{1-\mathrm{p}} > \Pr^{\mathfrak{S}}(\Diamond \mathit{Eff})$$

for $\mathrm{p} > 0$, where the second equality comes from the fact that all $\mathfrak{S}$-paths reaching goal satisfy $\Diamond \mathrm{C}_1$. Therefore, $\mathfrak{S}$ is a GPR scheduler for C_1 and Eff in $\mathcal{M}$.

With this example we observe an important difference between potential and universal probability-raising causes. The scheduler $\mathfrak{S}$ describes a scenario in which it is still beneficial to reach the coordinate C_1 on the Racetrack in order to reach the goal. Therefore, there is a probabilistic cause-effect relation between reaching C_1 and reaching goal in this specific scenario. However, as this relation does not hold in all possible scenarios, C_1 is not a universal probability-raising cause. So, in cases where no (universal) PR causes for an effect exist in the MDP it may still make sense to consider potential PR causes.

Similar to universal PR causality it is not clear when a set Cause constitutes a potential SPR or GPR cause. Thus, we address algorithms for checking whether there exists a strict or global probability-raising scheduler in the following subsections.

5.1.1 Checking for SPR Scheduler

Here, we consider the following problem: Given an MDP $\mathcal{M}$ with a non-empty set of terminal states $\mathit{Eff} \subset \mathrm{S}$ and a non-empty set of states $\mathit{Cause} \subseteq \mathrm{S} \setminus \mathit{Eff}$, is there a scheduler $\mathfrak{S}$ of $\mathcal{M}$ such that (SPR) holds for $\mathfrak{S}$? This is equivalent to asking whether Cause is a potential SPR cause for Eff in $\mathcal{M}$.

Since the transformation to the canonical causal MDP (Definition 31) does not affect the satisfaction of (SPR) or (GPR) by Lemma 32 and Corollary 33, we can assume the canonical MDP $\mathcal{M} = \mathcal{M}_{\mathit{Cause}} = (\mathrm{S}, \mathit{Act}, \mathrm{P}, \mathit{init})$ which satisfies assumptions (A1)-(A3).

Thus, $\mathcal{M}$ is end-component free and a terminal state of $\mathcal{M}$ is reached almost surely under any scheduler after finitely many steps in expectation. There are only four terminal states in $\mathcal{M}$, namely TP, FP, FN and TN. For each state $\mathfrak{c} \in \mathit{Cause}$ there are only the two actions $\alpha_{\min}$ and $\alpha_{\max}$ available which transition to TP with $\mathfrak{p}_{\mathfrak{c},\min}$ or $\mathfrak{p}_{\mathfrak{c},\max}$ respectively and with the remaining probability they transition to FP.

The first intuitive solution for the problem would now be to only enable the maximizing action $\alpha_{\max}$ in all states $\mathfrak{c} \in \mathit{Cause}$, since it contrasts the minimizing MDP $\mathcal{M}^{\min}_{\mathit{Cause}}$ for universal SPR causes. However, this may give a too high probability for reaching Eff such that for some states $\mathfrak{c} \in \mathit{Cause}$ the maximal probability to reach Eff from $\mathfrak{c}$ given by $\mathfrak{p}_{\mathfrak{c},\max}$ can be too low. Thus, there needs to be a balancing between states $\mathfrak{c} \in \mathit{Cause}$ with a high value $\mathfrak{p}_{\mathfrak{c},\max}$ and states $\mathfrak{c}' \in \mathit{Cause}$ with a low value $\mathfrak{p}_{\mathfrak{c}',\max}$.

We will provide a characterization of SPR scheduler using this idea. For this, we consider the maximal probability to reach TP among all cause states:

$$\mathfrak{p}^\star = \max_{\mathfrak{c} \in \mathit{Cause}} \mathfrak{p}_{\mathfrak{c},\max}$$

Then, for any $\mathfrak{p} \in (0, \mathfrak{p}^\star]$ we define the transformed MDP $\mathcal{M}_\mathfrak{p} = (S, \mathit{Act} \cup \{\delta\}, P', \mathit{init})$ by

- for $\mathfrak{c} \in \mathit{Cause}$ the enabled actions are $\mathit{Act}(\mathfrak{c}) = \{\delta\}$ with $P'(\mathfrak{c}, \delta, TP) = \max\{\mathfrak{p}, \mathfrak{p}_{\mathfrak{c},\min}\}$ and $P'(\mathfrak{c}, \delta, FP) = 1 - P'(\mathfrak{c}, \delta, TP)$,
- for $\mathfrak{s} \in S \setminus \mathit{Cause}$, $\mathfrak{t} \in S$ and $\alpha \in \mathit{Act}$ we have $P'(\mathfrak{s}, \alpha, \mathfrak{t}) = P(\mathfrak{s}, \alpha, \mathfrak{t})$.

So, $\mathcal{M}_\mathfrak{p}$ behaves as $\mathcal{M}$ except when a state $\mathfrak{c} \in \mathit{Cause}$ is reached, in which case the single enabled action leads to TP with $\mathfrak{p}$ as long as $\mathfrak{p} > \mathfrak{p}_{\mathfrak{c},\min}$. We abbreviate $\mathcal{M}_{\mathfrak{p}^\star}$ by $\mathcal{M}^\star$. By construction, any scheduler $\mathfrak{U}$ of $\mathcal{M}_\mathfrak{p}$ corresponds to a scheduler $\mathfrak{S}$ of $\mathcal{M}$.

Lemma 81 (Characterization of SPR Scheduler). *For an MDP $\mathcal{M} = (S, \mathit{Act}, P, \mathit{init})$ with a non-empty set of terminal states $\mathit{Eff} \subset S$ and a non-empty set $\mathit{Cause} \subseteq S \setminus \mathit{Eff}$ the following statements are equivalent:*

(1) There is an SPR scheduler $\mathfrak{S}$ for Cause and Eff in $\mathcal{M}$.

(2) There exists a scheduler $\mathfrak{S}$ for $\mathcal{M}$ such that

- $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff}) < \mathfrak{p}^\star$ *and*
- *if $\mathfrak{c} \in \mathit{Cause}$ with $\Pr^{\mathfrak{S}}((\neg \mathit{Cause}) \,\mathrm{U}\, \mathfrak{c}) > 0$ then* (SPR) *holds for $\mathfrak{c}$ under $\mathfrak{S}$.*

(3) There exists $\mathfrak{p} \in (0, 1]$ with $\mathfrak{p} \leqslant \mathfrak{p}^\star$ such that $\Pr^{\min}_{\mathcal{M}_\mathfrak{p}}(\Diamond \mathit{Eff}) < \mathfrak{p}$.

(4) $\Pr^{\min}_{\mathcal{M}^\star}(\Diamond \mathit{Eff}) < \mathfrak{p}^\star$.

Proof. Obviously, (1) $\implies$ (2). For (2) $\implies$ (1) consider the following MR scheduler $\mathfrak{U}$ of $\mathcal{M}$:

- For $\mathfrak{s} \in S \setminus \mathit{Cause}$ $\mathfrak{U}$ define $\mathfrak{U}(\mathfrak{s}, \cdot)$ to be a uniform distribution over $\mathit{Act}(\mathfrak{s})$.
- For $\mathfrak{c} \in \mathit{Cause}$ define $\mathfrak{U}(\mathfrak{c}, \alpha_{\max}) = 1$.

By definition, $\mathcal{M}$ and $\mathcal{M}_{\mathfrak{U}}$ have the same underlying graph. Let $\mathfrak{S}$ be the SPR scheduler for Cause and Eff in $\mathcal{M}$ from (2). Since $\mathcal{M}$ has no end-components, we can assume $\mathfrak{S}$ to be memoryless (cf. Lemma 6). Moreover, for a value $\varepsilon \in (0,1]$ we can consider the scheduler $\mathfrak{S}_\varepsilon = \varepsilon\mathfrak{U} \oplus (1-\varepsilon)\mathfrak{S}$. For all states $s \in S$ we then get $\Pr^{\mathfrak{S}_\varepsilon}_{\mathcal{M}}(\Diamond s) > 0$. Moreover, let

$$\mathfrak{p}^{\mathfrak{S}}_{\mathit{Cause}} = \min_{c \in \mathit{Cause}} \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff} \mid (\neg \mathit{Cause}) \,\mathrm{U}\, c).$$

By condition (2) we have $\mathfrak{p}^{\mathfrak{S}}_{\mathit{Cause}} > \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff})$ and we can consider $\varepsilon > 0$ with

$$\frac{\varepsilon}{1-\varepsilon} < \mathfrak{p}^{\mathfrak{S}}_{\mathit{Cause}} - \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff}).$$

For such ε we have

$$\begin{aligned}
\Pr^{\mathfrak{S}_\varepsilon}_{\mathcal{M}}(\Diamond \mathit{Eff}) &\leqslant (1-\varepsilon) \cdot \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff}) + \varepsilon \cdot \Pr^{\mathfrak{U}}_{\mathcal{M}}(\Diamond \mathit{Eff}) \\
&< (1-\varepsilon) \cdot \mathfrak{p}^{\mathfrak{S}}_{\mathit{Cause}} \\
&\leqslant (1-\varepsilon) \cdot \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff} \mid (\neg \mathit{Cause}) \,\mathrm{U}\, c) \\
&< \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff} \mid (\neg \mathit{Cause}) \,\mathrm{U}\, c) \\
&\leqslant \Pr^{\mathfrak{S}_\varepsilon}_{\mathcal{M}}(\Diamond \mathit{Eff} \mid (\neg \mathit{Cause}) \,\mathrm{U}\, c),
\end{aligned}$$

which shows that $\mathfrak{S}_\varepsilon$ is an SPR scheduler for Cause and Eff in $\mathcal{M}$.

We next show (1) $\implies$ (3). Here, we are given the SPR scheduler $\mathfrak{S}$ for Cause and Eff in $\mathcal{M}$. For $\mathfrak{p} = \min_{c \in \mathit{Cause}} \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff} \mid (\neg \mathit{Cause})$, consider the scheduler $\mathfrak{T}$ of $\mathcal{M}_{\mathfrak{p}}$ that behaves identically to $\mathfrak{S}$ for all paths ending in non-terminal states $s \in S \setminus \mathit{Cause}$. Furthermore, for $c \in \mathit{Cause}$ consider the following sets of paths:

$$\begin{aligned}
\Pi_c &= \{\pi = \mathit{init} \cdots c \;\; \mathit{TP} \mid \pi \text{ is } \mathfrak{S}-\text{path}\}, \\
\Pi_{\mathit{FN}} &= \{\pi = \mathit{init} \cdots \mathit{FN} \mid \pi \text{ is } \mathfrak{S}-\text{path}\} \quad \text{and} \\
\Pi_{\mathit{TP}} &= \bigcup_{c \in \mathit{Cause}} \Pi_c.
\end{aligned}$$

This allows us to express the reachability probabilities under $\mathfrak{S}$ and $\mathfrak{T}$ by

$$\begin{aligned}
\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff}) &= \sum_{\pi \in \Pi_{\mathit{FN}}} \Pr^{\mathfrak{S}}_{\mathcal{M}}(\pi) + \sum_{\pi \in \Pi_{\mathit{TP}}} \Pr^{\mathfrak{S}}_{\mathcal{M}}(\pi), \\
\Pr^{\mathfrak{T}}_{\mathcal{M}_{\mathfrak{p}}}(\Diamond \mathit{Eff}) &= \sum_{\pi \in \Pi_{\mathit{FN}}} \Pr^{\mathfrak{S}}_{\mathcal{M}}(\pi) + \sum_{\pi \in \Pi_{\mathit{TP}}} \Pr^{\mathfrak{T}}_{\mathcal{M}_{\mathfrak{p}}}(\pi).
\end{aligned}$$

However, by choice of $\mathfrak{p}$ and definition of $\mathcal{M}_{\mathfrak{p}}$ we have $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\pi) \geqslant \Pr^{\mathfrak{T}}_{\mathcal{M}_{\mathfrak{p}}}(\pi)$ for every $\pi \in \Pi_{\mathit{TP}}$ which yields:

$$\Pr^{\min}_{\mathcal{M}_{\mathfrak{p}}}(\Diamond \mathit{Eff}) \leqslant \Pr^{\mathfrak{T}}_{\mathcal{M}_{\mathfrak{p}}}(\Diamond \mathit{Eff}) \leqslant \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond \mathit{Eff}) < \mathfrak{p}.$$

We now turn to (3) $\implies$ (4). So, let $\mathfrak{p} \in (0, \mathfrak{p}^\star)$ and $\Pr^{\min}_{\mathcal{M}_\mathfrak{p}}(\Diamond \mathit{Eff}) < \mathfrak{p}$. For $\mathcal{M}^\star = \mathcal{M}_{\mathfrak{p}^\star}$ we need to show $\Pr^{\min}_{\mathcal{M}^\star}(\Diamond \mathit{Eff}) < \mathfrak{p}^\star$. Let $\mathfrak{T}$ be a scheduler of $\mathcal{M}_\mathfrak{p}$ which minimizes the probability of reaching Eff in $\mathcal{M}_\mathfrak{p}$. That is, $\Pr^{\mathfrak{T}}_{\mathcal{M}_\mathfrak{p}}(\Diamond \mathit{Eff}) = \Pr^{\min}_{\mathcal{M}_\mathfrak{p}}(\Diamond \mathit{Eff}) < \mathfrak{p}$. By using the abbreviation $\mathfrak{p}^{\min}_\mathfrak{c} = \Pr^{\min}_{\mathcal{M}}(\Diamond \mathit{Eff} \mid (\neg \mathit{Cause}) \,\mathrm{U}\, \mathfrak{c})$, we proceed to partition Cause into the sets

$$\begin{aligned}
\mathit{Cause}_{\leqslant} &= \{\mathfrak{c} \in \mathit{Cause} \mid \mathfrak{p}^{\min}_\mathfrak{c} \leqslant \mathfrak{p}\},\\
\mathit{Cause}_{\mathrm{in}} &= \{\mathfrak{c} \in \mathit{Cause} \mid \mathfrak{p} < \mathfrak{p}^{\min}_\mathfrak{c} \leqslant \mathfrak{p}^\star\},\\
\mathit{Cause}_{>} &= \{\mathfrak{c} \in \mathit{Cause} \mid \mathfrak{p}^{\min}_\mathfrak{c} > \mathfrak{p}^\star\}.
\end{aligned}$$

We also use the following abbreviations:

$$\rho = \Pr^{\mathfrak{T}}_{\mathcal{M}_\mathfrak{p}}(\Diamond \mathit{Cause}_{\leqslant}), \qquad \text{for } \mathfrak{c} \in \mathit{Cause}_{\mathrm{in}}:\ \rho_\mathfrak{c} = \Pr^{\mathfrak{T}}_{\mathcal{M}_\mathfrak{p}}(\Diamond \mathfrak{c}),$$
$$R = \sum_{\mathfrak{c} \in \mathit{Cause}_{>}} \Pr^{\mathfrak{T}}_{\mathcal{M}_\mathfrak{p}}(\Diamond \mathfrak{c}) \cdot \mathfrak{p}^{\min}_\mathfrak{c} + \Pr^{\mathfrak{T}}_{\mathcal{M}_\mathfrak{p}}(\Diamond \mathit{FN}).$$

We can thus express the effect probability in $\mathcal{M}_\mathfrak{p}$ under $\mathfrak{T}$ by

$$\Pr^{\mathfrak{T}}_{\mathcal{M}_\mathfrak{p}}(\Diamond \mathit{Eff}) = \rho \cdot \mathfrak{p} + \sum_{\mathfrak{c} \in \mathit{Cause}_{\mathrm{in}}} \rho_\mathfrak{c} \cdot \mathfrak{p}^{\min}_\mathfrak{c} + R$$

When viewing $\mathfrak{T}$ as a scheduler for $\mathcal{M}^\star$ we get

$$\Pr^{\mathfrak{T}}_{\mathcal{M}_\mathfrak{p}}(\Diamond \mathit{Eff}) = \rho \cdot \mathfrak{p}^\star + \sum_{\mathfrak{c} \in \mathit{Cause}_{\mathrm{in}}} \rho_\mathfrak{c} \cdot \mathfrak{p}^\star + R = r \cdot \mathfrak{p}^\star + R,$$

where $r = \rho + \sum_{\mathfrak{c} \in \mathit{Cause}_{\mathrm{in}}} \rho_\mathfrak{c}$. We can now use $0 \leqslant \mathfrak{p}^\star - \mathfrak{p}^{\min}_\mathfrak{c} < \mathfrak{p}^\star - \mathfrak{p}$ to get

$$\Pr^{\mathfrak{T}}_{\mathcal{M}^\star}(\Diamond \mathit{Eff}) - \Pr^{\mathfrak{T}}_{\mathcal{M}_\mathfrak{p}}(\Diamond \mathit{Eff}) = \rho(\mathfrak{p}^\star - \mathfrak{p}) + \sum_{\mathfrak{c} \in \mathit{Cause}_{\mathrm{in}}} \rho_\mathfrak{c}(\mathfrak{p}^\star - \mathfrak{p}^{\min}_\mathfrak{c}) \leqslant r(\mathfrak{p}^\star - \mathfrak{p}).$$

We obtain:

$$\begin{aligned}
\Pr^{\min}_{\mathcal{M}^\star}(\Diamond \mathit{Eff}) &\leqslant \Pr^{\mathfrak{T}}_{\mathcal{M}^\star} \leqslant \Pr^{\mathfrak{T}}_{\mathcal{M}_\mathfrak{p}}(\Diamond \mathit{Eff}) + r(\mathfrak{p}^\star - \mathfrak{p})\\
&\leqslant \Pr^{\mathfrak{T}}_{\mathcal{M}_\mathfrak{p}}(\Diamond \mathit{Eff}) + \mathfrak{p}^\star - \mathfrak{p} = \mathfrak{p}^\star(\mathfrak{p} - \Pr^{\mathfrak{T}}_{\mathcal{M}_\mathfrak{p}}(\Diamond \mathit{Eff})) < \mathfrak{p}^\star,
\end{aligned}$$

where we use $r \leqslant 1$ and $\Pr^{\mathfrak{T}}_{\mathcal{M}_\mathfrak{p}}(\Diamond \mathit{Eff}) < \mathfrak{p}$ for the second line.

We are left to proof (4) $\implies$ (2). Here, we consider an MD scheduler $\mathfrak{T}$ of $\mathcal{M}^\star$ which realizes $\Pr^{\mathfrak{T}}_{\mathcal{M}^\star}(\Diamond \mathit{Eff}) = \Pr^{\min}_{\mathcal{M}^\star}(\Diamond \mathit{Eff})$. The corresponding scheduler $\mathfrak{S}$ of $\mathcal{M}$ to $\mathfrak{T}$ is a witness for condition (2). □

We can use the characterization of SPR scheduler given by Lemma 81 to obtain a polynomial time algorithm to check the existence of a SPR scheduler.

Theorem 82 (Complexity for Checking SPR Scheduler). *Given an MDP $\mathcal{M} = (S, \mathit{Act}, P, \mathit{init})$ with a non-empty set of terminal states $\mathit{Eff} \subset S$ and a non-empty set of states $\mathit{Cause} \subseteq S \setminus \mathit{Eff}$ the existence of an SPR scheduler for Cause and Eff in $\mathcal{M}$ can be decided in P. Moreover, a corresponding finite-memory (randomized) scheduler $\mathfrak{S}$ can be computed in polynomial time.*

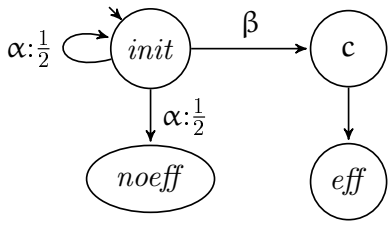

Figure 5.3: MDP $\mathcal{M}$ from Remark 83

Proof. We first note that we can transform $\mathcal{M}$ to $\mathcal{M}_{Cause}$ in polynomial time. We can now rely on the equivalences of (1) and (4) from Lemma 81. So, by further transforming $\mathcal{M}_{Cause}$ to $\mathcal{M}^{\star}_{Cause}$ we can check whether $\Pr^{\min}_{\mathcal{M}^{\star}_{Cause}} < \mathfrak{p}^{\star}$ in polynomial time [BK08].

If this existence check is positive, an MD scheduler $\mathfrak{T}$ of $\mathcal{M}^{\star}_{Cause}$ with $\Pr^{\mathfrak{T}}_{\mathcal{M}^{\star}_{Cause}} = \Pr^{\min}_{\mathcal{M}^{\star}_{Cause}}$ can be derived. By Remark 40 we can further derive the corresponding finite-memory (randomized) scheduler $\mathfrak{S}$ for $\mathcal{M}$ in polynomial time. □

Similar to Remarks 47 and 46 the usage of finite-memory and randomization is needed.

Remark 83 (PR Schedulers may need Finite Memory or Randomization). For the MDP $\mathcal{M}$ in Figure 5.3, with $\mathit{Eff} = \{\mathit{eff}\}$ and $\mathit{Cause} = \{\mathbf{c}\}$, there is no memoryless deterministic scheduler for which (GPR) or (SPR) hold. However, if we consider the finite memory deterministic scheduler $\mathfrak{S}_{\mathrm{fm}}$ which chooses $\boldsymbol{\alpha}$ for its first visit of init and $\boldsymbol{\beta}$ for his second, then we have

$$\Pr^{\mathfrak{S}_{\mathrm{fm}}}(\Diamond \mathit{eff} \mid \Diamond \mathbf{c}) = 1 > \frac{1}{2} = \Pr^{\mathfrak{S}_{\mathrm{fm}}}(\Diamond \mathit{eff}).$$

Similar, the memoryless randomized scheduler $\mathfrak{S}_{\mathrm{MR}}$ which chooses $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ with probability $1/2$ satisfies

$$\Pr^{\mathfrak{S}_{\mathrm{MR}}}(\Diamond \mathit{eff} \mid \Diamond \mathbf{c}) = 1 > \frac{1}{2} = \Pr^{\mathfrak{S}_{\mathrm{MR}}}(\Diamond \mathit{eff}).$$

Thus, for an potential GPR- or SPR cause it might be needed to have finite-memory or randomized scheduler as a GPR- or SPR scheduler. ◁

Remark 84 (Complexity for Singleton Potential PR Causes). The case where $\mathit{Cause} = \{\mathbf{c}\}$ is a singleton can be solved more efficient. To see this, note that there does not need to be a balancing between different states of Cause. Therefore, it is sufficient to only enable action $\boldsymbol{\alpha}_{\max}$ in $\mathbf{c}$. Let us call the resulting MDP $\mathcal{N}$. We then only have to check, whether $\mathfrak{p}_{\max,\mathbf{c}} > \Pr^{\min}_{\mathcal{N}}(\Diamond \mathit{Eff})$ as this corresponds to the best case SPR scheduler. If this scheduler does not reach Cause we can use the same argumentation as for Theorem 82, to see that Cause is still a potential SPR cause. Furthermore, a check for the existence of a potential PR cause can also be solved by using this method as a set $\mathit{Cause} \subseteq \mathrm{S} \setminus \mathit{Eff}$ and a scheduler $\mathfrak{S}$ satisfying (SPR) exists if and only if there exists a singleton $\{\mathbf{c}\} = \mathit{Cause}$ with these properties. Therefore, checking $\mathfrak{p}_{\max,\mathbf{s}} > \Pr^{\min}_{\mathcal{N}}(\Diamond \mathit{Eff})$ for all states $\mathbf{s} \in \mathrm{S} \setminus \mathit{Eff}$ gives an answer to the existence query for potential PR causes. ◁

5.1.2 Checking for GPR Scheduler

We now turn to the global probability-raising condition and address the question whether for a given an MDP $\mathcal{M}$ with non-empty set of terminal states $\mathit{Eff} \subset \mathrm{S}$ and non-empty set of states $\mathit{Cause} \subseteq \mathrm{S} \setminus \mathit{Eff}$ there is a GPR scheduler $\mathfrak{S}$ for Cause and Eff in $\mathcal{M}$. This is the same query as checking whether Cause is a potential GPR cause for Eff in $\mathcal{M}$. For this we use a very similar approach to checking universal GPR causality. We use Lemma 32 to assume an end-component free MDP $\mathcal{M} = \mathcal{M}_{\mathit{Cause}}$ and then encode the desired property (GPR) in terms of frequency variables. Therefore, we consider the equations (S1) - (S4) (cf. Sections 2.1.4 and 4.1.3) and the corresponding frequency inequality resulting from (GPR-1) given by:

$$\mathrm{x}_{FN} \cdot \mathrm{x}_{FP} - \mathrm{x}_{TN} \cdot \mathrm{x}_{TP} < 0. \tag{freq-GPR}$$

This leads to the following result:

Lemma 85 (Quadratic Program for GPR Scheduler). *For an MDP $\mathcal{M} = \mathcal{M}_{\mathit{Cause}}$, the set Cause is a potential GPR cause for Eff if and only if the system of inequalities given by* (S1) - *(S3) and (freq-GPR) has a solution.*

Proof. Let $\mathfrak{S}$ be a GPR scheduler for Cause and Eff 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 $\mathrm{x} \in \mathbb{R}^{\mathit{StAct}}$ is a solution to (S1)-(S3) and (freq-GPR) and let $\mathfrak{S}$ be an MR scheduler corresponding to x. By Lemma 50 the inequality (freq-GPR) is equivalent to (GPR) in $\mathcal{M} = \mathcal{M}_{\mathit{Cause}}$. As (freq-GPR) and (S1) hold, we also have $\mathrm{x}_{TN} \cdot \mathrm{x}_{TP} > 0$ and thus in particular $\mathrm{x}_{TP} > 0$. Since $\Pr^{\mathfrak{S}}(\Diamond \mathit{Cause}) = \mathrm{x}_{TP} + \mathrm{x}_{FP}$ we also get $\Pr^{\mathfrak{S}}(\Diamond \mathit{Cause}) > 0$. Therefore, $\mathfrak{S}$ is a GPR scheduler for Cause and Eff in $\mathcal{M}$. □

However, since (freq-GPR) is a strict inequality we can not use quadratic programming in a straight forward way to find a solution to our system of inequalities. We can still show an NP upper bound for deciding the existence of a GPR scheduler for Cause and Eff in $\mathcal{M}$.

Theorem 86 (Complexity of GPR-Scheduler Check). *Given an MDP $\mathcal{M} = (\mathrm{S}, \mathit{Act}, \mathrm{P}, \mathit{init})$, a non-empty set of terminal states $\mathit{Eff} \subset \mathrm{S}$ and a non-empty set $\mathit{Cause} \subseteq \mathrm{S} \setminus \mathit{Eff}$, deciding whether Cause is a potential GPR cause can be done in* NP.

Proof. The transformations from $\mathcal{M}$ to $\mathcal{M}_{\mathit{Cause}}$ take polynomial time. As such, we will assume $\mathcal{M} = \mathcal{M}_{\mathit{Cause}}$ throughout the proof. Define the polytope P of feasible state-action frequencies for a scheduler $\mathfrak{S}$ of $\mathcal{M}$ defined by inequalities (S1)-(S3) by

$$\mathrm{P} = \{\mathrm{x} \in \mathbb{R}^{\mathit{StAct}} \mid \mathrm{x} \text{ satisfies (S1) - (S3)}\}$$

and the function defined by the left hand side of (freq-GPR) by $\mathrm{f} : \mathrm{P} \to \mathbb{R}$ with

$$\mathrm{f}(\mathrm{x}) = \mathrm{x}_{FN} \cdot \mathrm{x}_{FP} - \mathrm{x}_{TN} \cdot \mathrm{x}_{TP}.$$

By definition P is an $\mathfrak{m} = |StAct| - |\mathrm{S}|$ dimensional subspace of $\mathbb{R}^{|StAct|}$ since each equation given by (S2) and (S3) takes away one dimension. For the remaining proof, however, we will consider P as a subspace of $\mathbb{R}^{\mathfrak{m}}$. We note that $\mathfrak{f}$ is continuous as it is quadratic and that P is closed and convex as it is defined solely by linear equations and inequalities. If there are both $\mathfrak{x}_0 \in \mathrm{P}$ with $\mathfrak{f}(\mathfrak{x}_0) < 0$ and $\mathfrak{x}_1 \in \mathrm{P}$ with $\mathfrak{f}(\mathfrak{x}_1) > 0$ then there is also $\mathfrak{x} \in \mathrm{P}$ with $\mathfrak{f}(\mathfrak{x}) = 0$ by the intermediate value theorem. We thus get three disjoint cases

Case 1. For all $\mathfrak{x} \in \mathrm{P}$ we have $\mathfrak{f}(\mathfrak{x}) > 0$.

Case 2. For all $\mathfrak{x} \in \mathrm{P}$ we have $\mathfrak{f}(\mathfrak{x}) < 0$.

Case 3. There is $\mathfrak{x} \in \mathrm{P}$ with $\mathfrak{f}(\mathfrak{x}) = 0$.

In order to decide in which case we are, we define the set P_0 of zero points of $\mathfrak{f}$ in P, that is, $\mathrm{P}_0 = \{\mathfrak{x} \in \mathrm{P} \mid \mathfrak{f}(\mathfrak{x}) = 0\}$. Deciding whether this set is empty can be done by quadratic programming in NP by [Vav90]. If the set is empty then we take any point $\mathfrak{x} \in \mathrm{P}$ and evaluate $\mathfrak{f}(\mathfrak{x})$ in polynomial time to decide whether we are in case 1 or case 2. In case 1 there is no MR scheduler satisfying (GPR) and thus Cause is not a potential GPR cause. In case 2 all MR schedulers $\mathfrak{S}$ of $\mathcal{M}$ satisfy (GPR) and thus Cause is even a universal GPR cause. If the set P_0 is non-empty then we have to do further analysis.

Recall that the *interior* $\overset{\circ}{\mathrm{S}}$ of a set $\mathrm{S} \subseteq \mathbb{R}^{\mathfrak{m}}$ is the set of all $\mathfrak{x} \in \mathrm{S}$ for which there is an $\varepsilon > 0$ such that $\{\mathfrak{y} \in \mathbb{R}^{\mathfrak{m}} \mid \|\mathfrak{y} - \mathfrak{x}\| \leqslant \varepsilon\} \subseteq \mathrm{S}$. Furthermore, if S is a closed set then its *boundary* is $\partial \mathrm{S} = \mathrm{S} \setminus \overset{\circ}{\mathrm{S}}$. Since P_0 is non-empty and closed there are two further disjoint cases

Case 3.1 All $\mathfrak{x} \in \mathrm{P}_0$ are on the boundary $\partial\mathrm{P}$ of P.

Case 3.2 There is some $\mathfrak{x} \in \mathrm{P}_0$ that is in the interior $\overset{\circ}{\mathrm{P}}$ of P.

If we are in case 3.2 we can non-deterministically guess a point $\mathfrak{x} \in \mathrm{P}_0$ and confirm that $\mathfrak{x} \notin \partial\mathrm{P}$ in polynomial time, since $\partial\mathrm{P}$ can be computed in polynomial-time as the boundary of a convex polytope given by a linear program. Note, that the encoding of $\mathfrak{x} \in \mathrm{P}_0$ is also bounded by [Vav90] as it is a solution of a quadratic program. Therefore, we can decide whether we are in case 3.2 in NP if not, then we are in case 3.1.

In case 3.1 either all interior points $\mathfrak{x} \in \overset{\circ}{\mathrm{P}}$ of P satisfy $\mathfrak{f}(\mathfrak{x}) > 0$ or all such points satisfy $\mathfrak{f}(\mathfrak{x}) < 0$ as otherwise there would be a point $\mathfrak{x} \in \overset{\circ}{\mathrm{P}}$ with $\mathfrak{f}(\mathfrak{x}) = 0$ by the intermediate value theorem as P is convex and $\mathfrak{f}$ is continuous. Therefore, we can take any point $\mathfrak{x} \in \overset{\circ}{\mathrm{P}}$ and check whether $\mathfrak{f}(\mathfrak{x}) < 0$. If yes, then there is a scheduler $\mathfrak{S}$ corresponding to $\mathfrak{x}$ which is also a GPR scheduler for Cause and Eff in $\mathcal{M}$. If not, then there is no GPR scheduler for Cause and Eff in $\mathcal{M}$.

We are left with case 3.2. So, let $\mathfrak{x} \in \mathrm{P}_0 \cap \overset{\circ}{\mathrm{P}}$. Intuitively, we will move from $\mathfrak{x}$ in the direction of $-\nabla\mathfrak{f}(\mathfrak{x})$ while staying inside of P. This way, we arrive at a point $\mathfrak{x}' \in \mathrm{P}$ for which $\mathfrak{f}(\mathfrak{x}') < 0$. Note, that there are no outgoing actions from the states TP, TN, FP and FN. Thus, the values $\mathfrak{x}_{TP}, \mathfrak{x}_{TN}, \mathfrak{x}_{FP}$ and $\mathfrak{x}_{FN}$ directly appear in $\mathfrak{x}$, w.l.o.g these are the last four coordinates

of $\mathfrak{x}$ in the order as listed. Moreover, since they are the only four terminal states and $\mathfrak{M}_{Cause}$ has no end-components we also have $\mathfrak{x}_{TP} + \mathfrak{x}_{TN} + \mathfrak{x}_{FP} + \mathfrak{x}_{FN} = 1$. We now consider the gradient $\nabla \mathfrak{f}(\mathfrak{x})$ of $\mathfrak{f}$ at $\mathfrak{x}$. Recall, that the gradient of a function $\mathfrak{f}$ is representing the direction of the maximal increase of $\mathfrak{f}$ and its norm is describing the rate of the maximal increase. Since $\mathfrak{f}$ only uses the values $\mathfrak{x}_{TP}, \mathfrak{x}_{TN}, \mathfrak{x}_{FP}$ and $\mathfrak{x}_{FN}$, we can ignore all other coordinates for the notation of $\nabla \mathfrak{f}(\mathfrak{x})$. So for $\nabla \mathfrak{f}(\mathfrak{x})$ we we have

$$\nabla \mathfrak{f}(\mathfrak{x}) = \begin{bmatrix} -\mathfrak{x}_{TN} \\ -\mathfrak{x}_{TP} \\ \mathfrak{x}_{FN} \\ \mathfrak{x}_{FP} \end{bmatrix} \quad \text{and} \quad \|\nabla \mathfrak{f}(\mathfrak{x})\| = \sqrt{\mathfrak{x}_{TN}^2 + \mathfrak{x}_{TP}^2 + \mathfrak{x}_{FN}^2 + \mathfrak{x}_{FP}^2}.$$

Since $\mathfrak{x}_{TP} + \mathfrak{x}_{TN} + \mathfrak{x}_{FP} + \mathfrak{x}_{FN} = 1$ we also get $\|\nabla \mathfrak{f}(\mathfrak{x})\| > 0$ for all $\mathfrak{x} \in \mathfrak{P}$. Thus, for all $\mathfrak{x} \in \mathfrak{P}$ the function $\mathfrak{f}$ is always increasing in direction $\nabla \mathfrak{f}(\mathfrak{x})$ and decreasing in direction $-\nabla \mathfrak{f}(\mathfrak{x})$. Futhermore, it has no local extreme points inside of $\mathfrak{P}$.

Recall, that we consider case 3.2 and $\mathfrak{x} \in \mathfrak{P}_0 \cap \overset{\circ}{\mathfrak{P}}$. Since $\mathfrak{x}$ is in the interior of $\mathfrak{P}$, there exists an $\varepsilon > 0$ such that $\mathfrak{x}' = \mathfrak{x} - \varepsilon \cdot \nabla \mathfrak{f}(\mathfrak{x}) \in \mathfrak{P}$. As the value of $\mathfrak{f}(\mathfrak{x})$ is decreasing in direction $-\nabla \mathfrak{f}(\mathfrak{x})$ we then get $\mathfrak{f}(\mathfrak{x}') < 0$. Since additionally $\mathfrak{x}' \in \mathfrak{P}$ is indeed a solution to (S1)-(S3) and (freq-GPR), an MR scheduler $\mathfrak{S}$ which corresponds to $\mathfrak{x}'$ constitutes a GPR scheduler for Cause and Eff in $\mathfrak{M}$. □

Corollary 87 (Memory Requirements of GPR Scheduler). *Given an MDP $\mathfrak{M}$, a non-empty set of terminal states $\mathit{Eff} \subset S$ and a non-empty set $\mathit{Cause} \subseteq S \setminus \mathit{Eff}$, if there exists a GPR scheduler for Cause and Eff in $\mathfrak{M}$, then there also exists a finite-memory randomized GPR scheduler $\mathfrak{S}$ with exactly two memory cells.*

Proof. We will use the established notations from the proof of Theorem 86. In the proof of Theorem 86 in case 2 all scheduler $\mathfrak{U}$ of $\mathfrak{M}_{Cause}$ are GPR scheduler. Then, we can pick any MR scheduler $\mathfrak{U}$ of $\mathfrak{M}_{Cause}$ as the GPR scheduler. In case 3.1 where all interior points $\mathfrak{x} \in \overset{\circ}{\mathfrak{P}}$ satisfy $\mathfrak{f}(\mathfrak{x}) < 0$, we can pick any point from the interior of $\mathfrak{P}$ and the corresponding MR scheduler is a GPR scheduler. In case 3.2 we take a point $\mathfrak{x} \in \overset{\circ}{\mathfrak{P}} \cap \mathfrak{P}_0$ that satisfies $\mathfrak{f}(\mathfrak{x}) = 0$ and consider $\nabla \mathfrak{f}(\mathfrak{x})$. We then move from $\mathfrak{x}$ in the direction of $-\nabla \mathfrak{f}(\mathfrak{x})$ until we reach the boundary of $\mathfrak{P}$ at some point $\mathfrak{y} \in \partial \mathfrak{P}$. An MR scheduler $\mathfrak{U}$ of $\mathfrak{M}_{Cause}$ corresponding to $\mathfrak{y}$ then satisfies (GPR).

The GPR scheduler $\mathfrak{U}$ can be chosen in a way such that the choices for the τ-action, which leaves MEC-states $\mathfrak{s}_{\mathcal{E}}$ with probability 1, are deterministic by using the same arguments as in the proof of Theorem 56. Thus, for $\mathfrak{U}$ a corresponding scheduler $\mathfrak{T}$ of the intermediate MDP $\mathfrak{M}_{<Cause>}$ can be computed in a way such that the traversal of MECs does not require additional memory due to Theorem 39. As the transformation $\mathfrak{M} \to \mathfrak{M}_{<Cause>}$ introduces two modes for "before" and "after" Cause, a corresponding scheduler $\mathfrak{S}$ of $\mathfrak{M}$ to the GPR scheduler $\mathfrak{T}$ in $\mathfrak{M}_{<Cause>}$ needs two memory cells. □

With this, we have established the complexities for checking the existence of probability-raising scheduler for a given set $Cause$. For continuing along the lines of Chapter 4 the next algorithmic question would be to derive a non-empty set $Cause \subseteq S \setminus Eff$ for which a probability-raising scheduler exists. For universal PR causes we use quality measures as criteria to which optimal universal PR causes are tailored. Although possible, this approach is not as sensible here, since the resulting quality-optimal potential PR cause may only have a few PR scheduler. Thus, it makes more sense to consider different quality criteria which take the amount of scheduler which are probability-raising for $Cause$ and Eff into account. This consideration is current ongoing work and will be addressed in more detail in the conclusion of the thesis (Chapter 6).

5.2 ω-Regular Effects - Reachability Causes

This section, discusses a modified variant of probability-raising causes in which the specification of the effect is changed. Instead of a reachability property of a set of terminal states we consider the effect to be a non-empty ω-regular language $rEff \subseteq S^\omega$ over the state space S. This extension is natural since reachability properties are special cases of ω-regular languages. We assume that the ω-regular language $rEff$ is given by a deterministic Rabin automaton $\mathcal{A}_{rEff}$ (Def. 10). In this section we keep the specification of the cause event to be a reachability property and thus call the resulting notion *reachability causes*. For reachability GPR causes, the techniques from the previous section are mostly still applicable, as we can reduce the setting from ω-regular effects to state-based effects using the product MDP $\mathcal{M} \otimes \mathcal{A}_{rEff}$ For reachability SPR causes, on the other hand, we observe that they take on the flavor of state-based GPR causes when using this product construction.

We start by stating the modified variant of PR causes tailored to ω-regular effects. Let $\mathcal{M} = (S, Act, P, init)$ be an MDP and $rEff \subseteq S^\omega$ be a non-empty ω-regular language over S.

Definition 88 (Reachability Probability-Raising Causes). Given an MDP $\mathcal{M}$ and a non-empty ω-regular language $rEff \subseteq S^\omega$, a non-empty set $Cause \subseteq S$ which satisfies the condition $\Pr^{\max}_{\mathcal{M}}((\neg Cause)\,\mathrm{U}\,c) > 0$ for each $c \in Cause$ is a *reachability GPR cause* for $rEff$ in $\mathcal{M}$ if the following condition **(rG)** holds:

(rG) For each scheduler $\mathfrak{S}$ with $\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond Cause) > 0$ we have

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(rEff \mid \Diamond Cause) \;>\; \Pr^{\mathfrak{S}}_{\mathcal{M}}(rEff). \tag{rGPR}$$

$Cause$ is a *reachability SPR cause* for $rEff$ in $\mathcal{M}$ if the following condition **(rS)** holds.

(rS) For each state $c \in Cause$ and each scheduler $\mathfrak{S}$ with $\Pr^{\mathfrak{S}}_{\mathcal{M}}((\neg Cause)\,\mathrm{U}\,c) > 0$ we have

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(rEff \mid (\neg Cause)\,\mathrm{U}\,c) \;>\; \Pr^{\mathfrak{S}}_{\mathcal{M}}(rEff). \tag{rSPR}$$

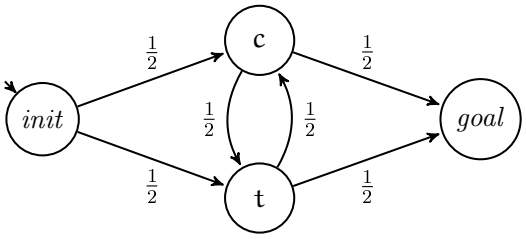

Figure 5.4: Markov chain where $\mathbf{c}$ is a reachability cause

There is one caveat which we need to mention here. Suppose the effect $rEff$ is a reachability property $\Diamond Eff$ for a set of states $Eff \subset \mathrm{S}$, then this notion of causality allows for PR causes $Cause$ to not be disjoint from the set of states Eff. However, if both such sets $Cause, Eff \subset \mathrm{S}$ are disjoint then $Cause$ is a state based PR cause for Eff according to Definition 14 if and only if $Cause$ is a reachability PR cause for the ω-regular event $\Diamond Eff$ according to Definition 88. In particular, this means that simple properties collected in Section 3.1 (Prop. 15, Ex. 16, Rem. 17) still hold for reachability PR causes. As we now view the effect as the ω-regular property on infinite executions, we argue that the *temporal priority* (C2) is captured by the new definition, since the cause will be reached after finitely many steps, if it is reached. We will address problems with this interpretation and a stronger notion of temporal priority in Section 5.2.2. Moreover, all of the motivation for considering state-based PR causality in MDPs still holds in this setting. For more details on general motivation and simple properties of PR causes see Section 3.1 and for motivating examples see Section 3.2.

The extension to reachability PR causes now allows us to consider more general properties as effects. Consider the DTMC $\mathcal{M}$ given in Figure 5.4 as an example. Let $rEff$ be described by the LTL formula $\varphi = \neg \mathbf{t}\, \mathrm{U}\, goal$ and consider the singleton $Cause = \{\mathbf{c}\}$. Then we have

$$\Pr(rEff \mid \Diamond Cause) = \frac{\Pr(\varphi \wedge \Diamond \mathbf{c})}{\Pr(\Diamond \mathbf{c})} = \frac{1}{4} / \frac{3}{4} = \frac{1}{3} > \frac{1}{2} = \Pr(rEff),$$

which shows that $Cause$ is indeed a reachability GPR and SPR cause for $rEff$ in $\mathcal{M}$. However, a model transformation of $\mathcal{M}$ such that $rEff$ is described by some reachability property is possible and is indeed what we propose algorithmically in Section 5.2.1. Our algorithmic results are summarized in Table 5.1.

5.2.1 Checking Reachability Probability-Raising Causes

To explore how this change of definition influences the previously established results for GPR and SPR causes, we take a closer look at how ω-regular effects are represented. As stated in the beginning of this section, we use deterministic Rabin automata (DRAs) since they are expressive enough to capture all ω-regular languages and they are deterministic, which will allow us to form well-behaved products of the automata with MDPs. Let $\mathcal{M}$ be an MDP, $rEff$ an effect given by the DRA $\mathcal{A}_{rEff}$ and $Cause \subseteq \mathrm{S}$ a cause candidate.

reachability PR-causality	Check PR condition $Cause$ and $rEff$ are given	Compute PR cause $Cause$ for given effect $rEff$ $recall$-optimal	$fscore$-optimal
GPR	$\in \mathrm{coNP}$ (Cor.92)	poly-space NP-hard (Cor.98)	poly-space NP-hard (Cor.98)
SPR	$\in \mathrm{coNP}$ (Cor.92)	poly-time(Prop.97)	poly-space NP-hard (Cor.98)

Table 5.1: Complexity results for reachability PR with fixed regular effect $rEff$

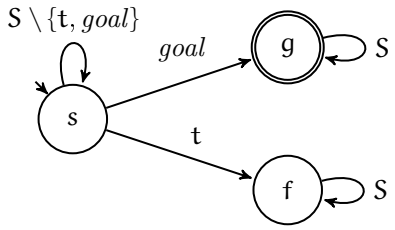

Figure 5.5: DRA for $\varphi = \neg \mathrm{t}\, \mathrm{U}\, goal$

For an example again consider the MDP $\mathcal{M}$ from Figure 5.4 with the ω-regular effect $rEff$ given by the LTL formula $\varphi = \neg \mathrm{t}\, \mathrm{U}\, goal$. A simple corresponding DRA $\mathcal{A}_{rEff}$ is given in Figure 5.5 where f stands for $fail$ and g stands for $goal$. The DRA is equipped with the acceptance set $Acc = \{(\{\mathrm{s}\}, \{\mathrm{g}\})\}$.

We now address the algorithmic question, whether $Cause$ is a reachability SPR/GPR cause. For, Markov chains the conditions **(rS)** and **(rG)** can again easily be checked by computing the corresponding (conditional) probabilities in polynomial time by [BK08; Bai+14]. For general MDPs we will present the cause checking algorithms by reducing the setting to the state-based case and then relying on previously established results (Thm. 53). For this, we will provide a model transformation of $\mathcal{M}$ using the DRA $\mathcal{A}_{rEff}$ such that the resulting MDP has no end-components and the effect is a reachability property again.

Notation 89 (Removing End-Components for Regular Effects). Let $\mathcal{M}$ and $\mathcal{A}_{rEff}$ be as above and consider the product MDP $\mathcal{N} = \mathcal{M} \otimes \mathcal{A}_{rEff}$. This product is an MDP equipped with a Rabin acceptance condition found in the second component of each state in the product.

We now take two copies of $\mathcal{N}$ representing a mode before $Cause$ has been reached and one mode after $Cause$ has been reached. So, each state $\mathfrak{s}$ is equipped with one extra bit 0 or 1. Initially, the MDP starts in the copy labeled with 0 and behaves like $\mathcal{N}$ until a state with its first component in $Cause$ is reached. From there, the process moves to the corresponding successor states in the second copy labeled with 1. We call the resulting MDP $\mathcal{N}'$ and denote

the set of states with their first component in $Cause$ in the first copy that are reachable in $\mathcal{N}'$ by $Cause_{\mathcal{A}_{rEff}}$, to express that these states are enriched with states of the automaton $\mathcal{A}_{rEff}$.

Next, we consider the MECs $\mathcal{E}_1, \ldots, \mathcal{E}_k$ of $\mathcal{N}'$. Note that the states in $Cause_{\mathcal{A}_{rEff}}$ are not contained in any MEC as they can only move to the second copy of $\mathcal{N}$. Furthermore, any MECs is either completely contained in the first copy and thus its states labeled 0, or it is completely contained in the second copy and its states labeled with 1. For each MEC $\mathcal{E}_i$ we determine whether there are scheduler making this MEC accepting of refuting. This means we check whether the following holds:

(i) For $\mathcal{E}_i$ there is a scheduler which ensures the event $Acc(\mathcal{A}_{rEff})$ (the acceptance condition of $\mathcal{A}_{rEff}$) is met with probability 1.

(ii) For $\mathcal{E}_i$ there is a scheduler which ensures this event is met with probability 0.

With the techniques of [Alf97] and [BGC09] both checks can be done in polynomial time.

We then add four new terminal states TP, FP, FN, and TN which will have the same meaning as in Definition 31. Next, we construct the MEC-quotient of $\mathcal{N}'$ while enabling a fresh actions in the states $s_{\mathcal{E}_i}$ representing MECs $\mathcal{E}_i$. For each $i \leqslant k$ we add a fresh actions in $s_{\mathcal{E}_i}$ which lead with probability 1 to

- FN if (i) holds and $\mathcal{E}_i$ is contained in copy 0,
- TN if (ii) holds and $\mathcal{E}_i$ is contained in copy 0,
- TP if (i) holds and $\mathcal{E}_i$ is contained in copy 1,
- FP if (ii) holds and $\mathcal{E}_i$ is contained in copy 1.

Finally, we remove all states which are not reachable from the initial state. We call the resulting MDP $\mathcal{M}_{rEff,Cause}$ and emphasize that this MDP contains all states in $Cause_{\mathcal{A}_{rEff}}$, has no end-components and has the four terminal states TP, FP, FN and TN. Furthermore, for each $c \in Cause$, we denote the subset of states in $Cause_{\mathcal{A}_{rEff}}$ in the first copy 0 with c in their first component by $c_{\mathcal{A}_{rEff}}$. ◁

Intuitively, this transformed product MDP works the same way as the canonical MDP for $Cause$ (Def. 31) in the state-based setting does. It consists of two modes for "before" and "after" $Cause$ and depending on that has four terminal states representing the confusion matrix (Fig. 3.14). This way, all relevant events can be expressed as (conjuncts of) reachability properties of those four states.

Example 90 (Product Construction). As an example for the transformation described in Notation 89 consider again the DTMC $\mathcal{M}$ from Figure 5.4 for the ω-regular LTL property $\varphi = \neg t \,\mathrm{U}\, goal$ given by the DRA $\mathcal{A}$ from Figure 5.5 with acceptance set $Acc = \{(\{s\}, \{g\})\}$. For $Cause = \{c\}$ the product DTMC $\mathcal{M} \otimes \mathcal{A}$ is depicted in Figure 5.6. Here, transition probabilities are omitted, as they are always equally distributed over all outgoing edges of a state.

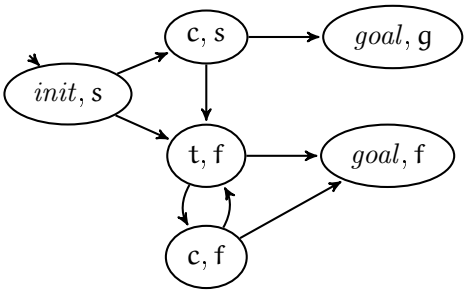

Figure 5.6: Product DTMC $\mathcal{M} \otimes \mathcal{A}$ from Example 90

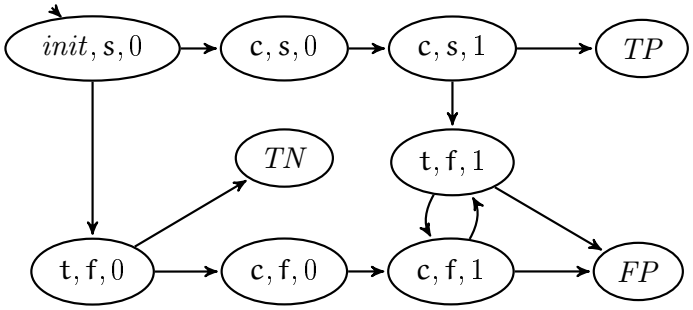

Figure 5.7: $\mathcal{M}_{rEff,\mathbf{c}}$ from Example 90

By further transforming $\mathcal{M} \otimes \mathcal{A}$ to $\mathcal{M}_{rEff,\mathbf{c}}$ we arrive at the DTMC depicted in Figure 5.7. Here, we omitted the intermediate (previously terminal) states $(goal, \mathrm{g}, 1), (goal, \mathrm{f}, 1)$ and $(goal, \mathrm{f}, 0)$ and immediately transitioned to the corresponding states TP, FP and TN respectively. In $\mathcal{M}_{rEff,\mathbf{c}}$ we have $\mathbf{c}_{\mathcal{A}_{rEff}} = \{(\mathrm{c}, \mathrm{s}, 0),\ (\mathrm{c}, \mathrm{f}, 0)\}$. ◁

The soundness of the transformation with regard to reachability PR causes is stated in the following Lemma 91.

Lemma 91 (Characterization of Reachability PR Causes). *Let $\mathcal{M}$, $Cause$, and $\mathcal{A}_{rEff}$ be as above and let $\mathcal{M}_{rEff,Cause}$ be the constructed MDP in Notation 89 which contains the set $Cause_{\mathcal{A}_{rEff}}$. Then,*

- *$Cause$ is a reachability GPR cause for $rEff$ in $\mathcal{M}$ if and only if $Cause_{\mathcal{A}_{rEff}}$ is a GPR cause for $\{TP, FN\}$ in $\mathcal{M}_{rEff,Cause}$ and*
- *$Cause$ is a reachability SPR cause for $rEff$ in $\mathcal{M}$ if and only if, for each $\mathbf{c} \in Cause$, the set of states $\mathbf{c}_{\mathcal{A}_{rEff}}$ is a GPR cause for $\{TP, FN\}$ in $\mathcal{M}_{rEff,Cause}$.*

Proof. Let us denote $\mathcal{K} = \mathcal{M}_{rEff,rCause}$ and $Cause_{\mathcal{A}_{rEff}} = Cause_{\mathcal{K}}$. For the set $Cause_{\mathcal{K}}$ in $\mathcal{K}$ we have $\Pr^{\max}_{\mathcal{K}}(((\neg Cause)_{\mathcal{K}}) \,\mathrm{U}\, \mathrm{d}) > 0$ for each state $\mathrm{d} \in Cause_{\mathcal{A}_{rEff}}$ by construction. This comes from the fact that all states in $Cause_{\mathcal{K}}$ are reachable and no run can visit two different states in $Cause_{\mathcal{K}}$. Therefore, the minimality condition is satisfied in any case and the sets $\{TP, FN\}$ and $Cause_{\mathcal{K}}$ are disjoint.

For the comparison of **(rG)** in $\mathcal{M}$ and **(G)** in $\mathcal{K}$ (resp. **(rS)** in $\mathcal{M}$ and **(S)** in $\mathcal{K}$) we establish pairs of corresponding scheduler between the two MDPs. For this, consider a scheduler $\mathfrak{T}$ of $\mathcal{K}$. We construct a corresponding scheduler $\mathfrak{S}$ of $\mathcal{M}$ mimicking $\mathfrak{T}$. As long as $\mathfrak{T}$ moves through the MEC-quotient of $\mathcal{N}'$ (Not. 89), the scheduler $\mathfrak{S}$ follows this behavior by leaving MECs through the corresponding actions in $\mathcal{M}$. Whenever $\mathfrak{T}$ moves to one of the states TP and FN, the last transition starts in a state $s_{\mathcal{E}}$ representing an MEC $\mathcal{E}$ of $\mathcal{M}$. In this case, $\mathfrak{S}$ will stay in $\mathcal{E}$ while ensuring with probability 1 that the resulting run is accepted by $\mathcal{A}_{rEff}$. This is possible by construction since then condition (i) from Notation 89 holds. Similarly, if $\mathfrak{T}$ moves

to one of the states TN or FP, the scheduler $\mathfrak{S}$ of $\mathcal{M}$ stays in the corresponding MEC and realizes the acceptance condition of $\mathcal{A}_{rEff}$ with probability 0 (condition (ii) of Notation 89 is satisfied).

Vice versa, consider a scheduler $\mathfrak{S}$ of $\mathcal{M}$. Then, we construct a scheduler $\mathfrak{T}$ of $\mathcal{K}$ analogously. As long as $\mathfrak{S}$ does not visit an end-component $\mathfrak{T}$ mimics this behavior. When an end-component $\mathcal{E}$ in $\mathcal{M}$ is by $\mathfrak{S}$ reached there are three events with respective probabilities

- $\mathfrak{S}$ stays in $\mathcal{E}$ and realizes the acceptance condition of $\mathcal{A}_{rEff}$ with probability $\mathfrak{p}_1$,
- $\mathfrak{S}$ stays in $\mathcal{E}$ and ensures the run is not accepted by $\mathcal{A}_{rEff}$ with probability $\mathfrak{p}_2$ and
- $\mathfrak{S}$ leaves the end-component $\mathcal{E}$ with probability $\mathfrak{p}_3$.

The scheduler $\mathfrak{T}$ of $\mathcal{K}$ can realizes the probability $\mathfrak{p}_1$ by choosing to move to the corresponding state TP of FN with $\mathfrak{p}_1$. Similarly, $\mathfrak{T}$ realizes $\mathfrak{p}_2$ by moving to the corresponding states TN or FP with $\mathfrak{p}_2$. For $\mathfrak{p}_3$ scheduler $\mathfrak{T}$ takes actions leaving $\mathcal{E}$ with respective probability $\mathfrak{T}$ will move to the corresponding state TP or FN with probability $\mathfrak{p}_1$, to TN or FP with probability $\mathfrak{p}_2$, and leave take actions leaving $s_{\mathcal{E}}$ to other states with the same probability distribution with which $\mathfrak{S}$ takes the leaving actions of $\mathcal{E}$.

This establishes a corresponding pair of schedulers $\mathfrak{S}$ and $\mathfrak{T}$. For these we observe

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\lozenge \mathit{Cause}) = \Pr^{\mathfrak{T}}_{\mathcal{K}}(\lozenge \mathit{Cause}_{\mathcal{K}}), \tag{5.1}$$

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\mathit{rEff}) = \Pr^{\mathfrak{T}}_{\mathcal{K}}(\lozenge\{TP, FN\}), \text{ and} \tag{5.2}$$

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\lozenge \mathit{Cause} \wedge \mathit{rEff}) = \Pr^{\mathfrak{T}}_{\mathcal{K}}(\lozenge TP) = \Pr^{\mathfrak{T}}_{\mathcal{K}}(\lozenge \mathit{Cause}_{\mathcal{K}} \wedge \lozenge\{TP, FN\}). \tag{5.3}$$

This establishes the first claim that $\mathit{Cause}_{\mathcal{A}_{rEff}}$ is a GPR cause for $\{TP, FN\}$ in $\mathcal{M}_{\mathit{rEff}, \mathit{Cause}}$ if and only if Cause is a reachability GPR cause for rEff in $\mathcal{M}$.

For the correspondence of (rS) in $\mathcal{M}$ to (S) in $\mathcal{K}$ we consider each element $\mathbf{c} \in \mathit{Cause}$ individually. We denote the set of states $\mathbf{c}_{\mathcal{A}_{rEff}}$ by $\mathbf{c}_{\mathcal{K}}$. For the above established pair of scheduler $\mathfrak{S}$ of $\mathcal{M}$ and $\mathfrak{T}$ of $\mathcal{K}$ we also have

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}((\neg \mathit{Cause})\,\mathrm{U}\,\mathbf{c}) = \Pr^{\mathfrak{T}}_{\mathcal{K}}(\lozenge \mathbf{c}_{\mathcal{K}}), \text{ and}$$
$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\mathit{rEff} \mid (\neg \mathit{Cause})\,\mathrm{U}\,\mathbf{c}) = \Pr^{\mathfrak{T}}_{\mathcal{K}}(\lozenge TP \mid \lozenge \mathbf{c}_{\mathcal{K}}).$$

Therefore, Cause is a reachability SPR cause for rEff in $\mathcal{M}$ if and only if $\mathbf{c}_{\mathcal{A}_{rEff}}$ is a GPR cause for $\{TP, FN\}$ in $\mathcal{M}_{\mathit{rEff}, \mathit{Cause}}$ for all $\mathbf{c} \in \mathit{Cause}$. □

Lemma 91 shows that reachability SPR causality shares similarities with GPR causality. This comes from the fact that in the transformation of Notation 89 a single state $\mathbf{c} \in \mathit{Cause}$ in $\mathcal{M}$ corresponds to a set of states $\mathbf{c}_{\mathcal{A}_{rEff}}$ in $\mathcal{M}_{\mathit{rEff}, \mathit{Cause}}$ which cab also be observed in Example 90. Here, checking whether $\mathbf{c}$ is a reachability SPR cause for $\varphi = \neg \mathbf{t}\,\mathrm{U}\,\mathit{goal}$ in $\mathcal{M}$ (Fig. 5.4) amounts to checking GPR causality of $\mathbf{c}_{\mathcal{A}_{rEff}} = \{(\mathbf{c}, \mathbf{s}, 0),\ (\mathbf{c}, \mathbf{f}, 0)\}$ for $\{TP\}$ in $\mathcal{M}_{\mathit{rEff}, \mathbf{c}}$ (Fig. 5.7).

Our following algorithmic results for reachability PR causes stem from the reduction provided by Notation 89 and Lemma 91. As an immediate consequence, we can check conditions (rG) and (rS) in $\mathcal{M}_{rEff,Cause}$ by using the already established algorithms for GPR causes (Theorem 53). This results in the following complexity upper bounds.

Corollary 92 (Complexity of Checking Reachability PR Causes). *For an MDP $\mathcal{M}$, a non-empty ω-regular language $rEff \subseteq S^\omega$ given as DRA $\mathcal{A}_{rEff}$ and a non-empty set $Cause \subseteq S$, we can decide whether $Cause$ is a reachability SPR/GPR cause for $rEff$ in coNP.*

Proof. By Lemma 91, we can use the construction of $\mathcal{M}_{rEff,Cause}$, which takes polynomial time. Then, Theorem 53 allows us to check whether $Cause$ is a reachability GPR cause in coNP directly. For SPR causality, we can apply the GPR check to each set $\mathbf{c}_{\mathcal{A}_{rEff}}$ and the effect set $\{TP, FN\}$ in $\mathcal{M}_{rEff,Cause}$ for $\mathbf{c} \in Cause$ in order to determine whether $Cause$ is a reachability SPR $Cause$ for $rEff$. □

The reduction from Notation 89 and Lemma 91 also shows that a characterization for the existence of reachability PR causes is analogous to Prop 48. Therefore, there is a reachability GPR cause iff there is a reachability SPR cause iff there is a singleton reachability SPR cause. Consequently, the existence of a reachability SPR/GPR cause can be checked by checking for each state $\mathbf{c}$ of the MDP whether $\{\mathbf{c}\}$ constitutes a reachability SPR cause. Thus, we conclude:

Corollary 93. *The existence of a reachability SPR/GPR cause can be decided in coNP.*

5.2.2 Quality and Optimality of Reachability PR Causes

As motivated in Chapter 3.3, we can view reachability PR causes as binary classifiers. This leads to the confusion matrix (Figure 3.14) with the difference that this time the path does not "hit" the effect set, but rather $rEff$ holds on the path. Hence, we define the following entries of the confusion matrix: Given $rEff$, $Cause$ and a scheduler $\mathfrak{S}$, we let

$$\mathsf{tp}^{\mathfrak{S}}_{\mathcal{M}} = \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond Cause \wedge rEff), \quad \mathsf{tn}^{\mathfrak{S}}_{\mathcal{M}} = \Pr^{\mathfrak{S}}_{\mathcal{M}}(\neg\Diamond Cause \wedge \neg rEff),$$
$$\mathsf{fp}^{\mathfrak{S}}_{\mathcal{M}} = \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond Cause \wedge \neg rEff), \quad \mathsf{fn}^{\mathfrak{S}}_{\mathcal{M}} = \Pr^{\mathfrak{S}}_{\mathcal{M}}(\neg\Diamond Cause \wedge rEff).$$

We omit the subscript $\mathcal{M}$ if the MDP is clear from context.

Lemma 94 (Product Construction preserves Confusion Matrix). *Let $\mathcal{M}$ be an MDP as above, $rEff$ an ω-regular effect given by the DRA $\mathcal{A}_{rEff}$ and let $Cause \subseteq S$ be a reachability SPR/GPR cause. Further, let $\mathcal{M}_{rEff,Cause}$ be as defined in Notation 89 with the set of cause states $Cause_{\mathcal{A}_{rEff}}$ and the set of effect states $\{TP, FN\}$. Then, for each scheduler $\mathfrak{S}$ for $\mathcal{M}$, there is a scheduler $\mathfrak{T}$ for $\mathcal{N} = \mathcal{M}_{rEff,Cause}$, and vice versa, such that*

$$\mathsf{tp}^{\mathfrak{S}}_{\mathcal{M}} = \mathsf{tp}^{\mathfrak{T}}_{\mathcal{N}}, \quad \mathsf{tn}^{\mathfrak{S}}_{\mathcal{M}} = \mathsf{tn}^{\mathfrak{T}}_{\mathcal{N}}, \quad \mathsf{fp}^{\mathfrak{S}}_{\mathcal{M}} = \mathsf{fp}^{\mathfrak{T}}_{\mathcal{N}}, \quad \mathsf{fn}^{\mathfrak{S}}_{\mathcal{M}} = \mathsf{fn}^{\mathfrak{T}}_{\mathcal{N}}.$$

Proof. In the proof of Lemma 91, we have seen that for each scheduler $\mathfrak{S}$ for $\mathcal{M}$, we can find a scheduler $\mathfrak{T}$ for $\mathcal{M}_{rEff,Cause}$ and vice versa such that Equations (5.1)-(5.3) hold. This implies the equalities claimed here. □

Analogous to Chapter 3.3, we now define $recall$, $covrat$, and $fscore$ of a reachability PR cause as the infimum of these values in terms of $\mathsf{tp}^{\mathfrak{S}}$, $\mathsf{fp}^{\mathfrak{S}}$, $\mathsf{tn}^{\mathfrak{S}}$, and $\mathsf{fn}^{\mathfrak{S}}$ over all schedulers $\mathfrak{S}$ for which the respective quality measures are defined. The computation of these values can then be done with the methods from the state-based setting:

Corollary 95 (Computing Quality Measures for Reachability PR Causes). *Let $\mathcal{M}$ be an MDP and $rEff \subseteq S^\omega$ an ω-regular language given as DRA $\mathcal{A}_{rEff}$. Given a reachability SPR/GPR cause $Cause \subseteq S$ we can compute $recall(Cause)$, $covrat(Cause)$ and $fscore(Cause)$ in polynomial time.*

Proof. Use Lemma 94, Lemma 60 and Theorem 67. □

Quality-Optimal Reachability PR Causes

When trying to find good causes for an ω-regular effect $rEff$, we cannot say that effects and causes should be disjoint as in the setting where effects and causes were sets of states. This leads to the possibility that causes might exist that do not capture the intuition behind temporal priority: For example, if the effect $rEff$ is a reachability property $\Diamond \mathrm{E}$ for a set of states E, the set of states E itself will be a reachability PR cause if for each state $\mathbf{c} \in \mathrm{E}$, $\Pr^{\max}(\neg \mathrm{E} \,\mathrm{U}\, \mathbf{c}) > 0$ holds. In turn, the set E would constitute a reachability PR cause with maximal quality values $precision(\mathrm{E}) = recall(\mathrm{E}) = fscore(\mathrm{E}) = 1$. Furthermore, there might be causes C that can only be reached after E has already been reached.

In order to account for the temporal priority of causes, we will include the following condition when trying to find good causes: We require for a cause candidate set $Cause \subseteq S$

$$\Pr^{\min}(rEff \mid (\neg Cause) \,\mathrm{U}\, \mathbf{c}) < 1 \text{ for all } \mathbf{c} \in Cause. \tag{TempPrio}$$

Intuitively, this condition requires that a reachability PR cause can not almost surely predict an effect. Thus it is never (probabilistically) certain that the execution will belong to $rEff$ after the cause has been reached. A variation of this criteria has also been proposed in [Bai+21b].

Remark 96 (Temporal Priority). The condition (TempPrio) can also be added to the definition of reachability PR causes. Algorithmically, a fter the product construction in Notation 89, the condition can be checked for a given cause candidate $Cause \subseteq S$ by checking whether for each $\mathbf{c} \in Cause$, there is at least one state $\mathfrak{s}$ in $\mathcal{N} = \mathcal{M}_{rEff,Cause}$ with $\mathbf{c}$ in the first component such that $\Pr^{\min}_{\mathcal{N},\mathfrak{s}}(\Diamond FP) > 0$.

Furthermore, this condition is stronger than the requirement that causes and effects are disjoint in the state-based setting. In the state-based setting, however, the analogue of condition (TempPrio) could also be included. Instead of having to be disjoint from a set of ef-

fect states Eff, a PR cause Cause would not be allowed to contain any state $\mathbf{s} \in \mathbf{S}$ with $\Pr_{\mathbf{s}}^{\min}(\Diamond \mathit{Eff}) = 1$ which implies disjointedness. However, this would restrict the notion of PR causes such that there can be no sufficient causes (PR causes with $\mathit{precision}(\mathit{Cause}) = 1$, also cf. Remark 69). ◁

We now turn to the complexity of computing recall- and coverage ratio-optimal reachability SPR causes. As in the state-based setting, we define the set $\mathfrak{S}$ of all possible singleton reachability SPR causes for rEff that in addition satisfy (TempPrio) as explained in Remark 96. By Corollary 92, we can check whether a state $\mathbf{c} \in \mathbf{S}$ is a singleton reachability SPR cause in coNP. The additional requirement (TempPrio) can be decided by using Remark 96. Thus, we can again define the set of singleton reachability SPR causes

$$\mathfrak{S} = \{\mathbf{c} \in \mathbf{S} \mid \mathbf{c} \text{ is reachability SPR cause with } \Pr^{\min}(\mathit{rEff} \mid (\neg \mathit{Cause}) \,\mathrm{U}\, \mathbf{c}) < 1\}.$$

We can now define a canonical reachability cause $\mathit{CanCause}$, by abuse of notation, by

$$\mathit{CanCause} = \{\mathbf{c} \in \mathfrak{S} \mid \Pr^{\max}(\neg \mathfrak{S} \,\mathrm{U}\, \mathbf{c}) > 0\}.$$

This set is now again the front of the set of all reachability SPR causes.

For the complexity of the computation of the recall- and coverage ratio-optimal canonical cause and its values, the observations above lead us to the complexity class $\mathrm{PF}^{\mathrm{NP}}$ as defined in [Sel94]. This class consists of all functions which can be computed in polynomial time with access to an NP-oracle, or equivalently a coNP-oracle.

Proposition 97 (Complexity for Quality-Optimal Reachability SPR Causes). *Given an MDP $\mathcal{M}$, a non-empty ω-regular effect rEff given by a DRA $\mathcal{A}_{\mathit{rEff}}$, if $\mathfrak{S} \neq \emptyset$ then $\mathit{CanCause}$ is a ratio- and recall optimal reachability SPR cause for rEff. The threshold problem for the coverage ratio and the recall can be decided in coNP. The optimal values $\mathit{recall}(\mathit{CanCause})$ and $\mathit{covrat}(\mathit{CanCause})$ can be computed in $\mathrm{PF}^{\mathrm{NP}}$.*

Proof. The optimality of $\mathit{CanCause}$ follows by the arguments used for Theorem 70. In order to compute $\mathit{CanCause}$ we check for each state $\mathbf{s} \in \mathbf{S}$ whether (rS) does not hold in NP and then take the remaining states $\mathbf{c} \in \mathbf{S}$ to define $\mathfrak{S}$ by deciding (TempPrio) for each $\mathbf{c}$ in $\mathbf{P}$. If (TempPrio) holds then $\mathbf{c} \in \mathfrak{S}$ and $\mathit{CanCause} = \{\mathbf{c} \in \mathfrak{S} \mid \Pr^{\max}(\neg \mathfrak{C} \,\mathrm{U}\, \mathbf{c}) > 0\}$. This allows to compute the recall- and coverage ratio-optimal cause $\mathit{CanCause}$ in $\mathrm{PF}^{\mathrm{NP}}$.

For the threshold problem whether there is a reachability SPR cause with recall at least a given $\vartheta \in \mathbb{Q}$, the coNP upper bound can be shown as follows: For each state $\mathbf{s}$ that does not belong to $\mathfrak{S}$, i.e., that is not a singleton reachability SPR cause, there is a polynomial size certificate, as it can be checked in coNP. The collection of all states that do not belong to $\mathfrak{S}$ together with these certificates can now serve as a certificate that the optimal recall is less than ϑ in this the case. Given such a collection of certificates, we can check in polynomial time that indeed all provided states do not belong to $\mathfrak{S}$. The complement of the provided states

forms a super set $\mathcal{D}$ of $\mathit{CanCause}$. Computing the recall of the set $\mathcal{D}$ can then be done in polynomial time as in Corollary 95. This value is an upper bound for the recall of $\mathit{CanCause}$. Note that if all states that do not belong to $\mathcal{S}$ are given in the certificate, the value even equals the recall of $\mathit{CanCause}$.

So, if the optimal value is less than ϑ, this is witnessed by the described certificate containing all states not in $\mathcal{S}$. Vice versa, if a certificate is given resulting in a super set $\mathcal{D}$ of $\mathit{CanCause}$ such that the recall of $\mathcal{D}$ is less than ϑ, then there is no reachability SPR cause with a recall of at least ϑ. So, the threshold problem lies in coNP. For the coverage ratio, the analogous argument works by using Lemma 60. □

For the threshold problems for f-score-optimal reachability SPR causes and $\mathit{quality}$-optimal reachability GPR causes (where $\mathit{quality} \in \{\mathit{recall}, \mathit{covrat}, \mathit{fscore}\}$) that satisfy (TempPrio), we rely on the guess-and-check approach used for optimal GPR causes in the state-based setting. We guess a subset Cause of states of the MDP, check whether we found a reachability SPR/GPR cause in coNP and compute the quality measure under consideration in polynomial time as explained in the previous section. For the computation of an optimal cause, we obtain a polynomial-space algorithm.

Corollary 98 (Complexity of Quality-Optimal Reachability GPR Causes). *Let $\mathcal{M}$ be an MDP and rEff be a non-empty ω-regular language given by a DRA $\mathcal{A}_{\mathit{rEff}}$. Given $\vartheta \in \mathbb{Q}$, deciding whether there exists a reachability GPR cause Cause with $\mathit{recall}(\mathit{Cause}) \geqslant \vartheta$ (resp. $\mathit{covrat}(\mathit{Cause}) \geqslant \vartheta$, $\mathit{fscore}(\mathit{Cause}) \geqslant \vartheta$) can be done in Σ_2^{P} and is NP-hard. NP-hardness holds for Markov chains.*

Deciding whether there is a reachability SPR cause Cause with $\mathit{fscore}(\mathit{Cause}) \geqslant \vartheta$ can be done in Σ_2^{P}. A recall-, covratio-, or f-score-optimal reachability GPR cause as well as an f-score-optimal reachability SPR cause can be computed in polynomial space.

Proof. Let us first note, that by the discussed embedding of state-based GPR causes into reachability GPR causes (cf. discussion after Definition 88) the established NP-hardness of finding quality-optimal GPR causes from Theorem 77 extend to quality-optimal reachability GPR causes. On the other hand, the upper bounds also extend by using the construction from Notation 89. Since for Theorem 77 we relied on guess-and-check algorithms to establish the upper bounds for the threshold problems, we can use analogous algorithms in setting of ω-regular effects. We guess a set $\mathit{Cause} \subseteq S$, check the reachability GPR causality in coNP (Corollary 92) and compute the value of the quality measure in polynomial time (Corollary 95). Again, the alternation between the existential quantification for guessing Cause and the universal quantification for the coNP check results in the complexity Σ_2^{P} of the polynomial-time hierarchy.

In order to show that the decision problem for $\mathit{fscore}(\mathit{Cause}) \geqslant \vartheta$ for SPR causes Cause is in Σ_2^{P} we resort to the constructed MDP $\mathcal{M}_{\mathit{rEff},\mathit{Cause}}$. By Lemma 91 a reachability SPR cause Cause in $\mathcal{M}$ corresponds to a set of GPR causes $\{\mathbf{c}_{\mathcal{A}_{\mathit{rEff}}} \mid \mathbf{c} \in \mathit{Cause}\}$ in $\mathcal{M}_{\mathit{rEff},\mathit{Cause}}$, which can

be interpreted as GPR cause $\mathbf{C} = \bigcup_{\mathbf{c} \in Cause} \mathbf{c}_{\mathcal{A}_{rEff}}$ for $\{TP, FP\}$ in $\mathcal{M}_{rEff,Cause}$. This way we can encode the property $fscore(Cause) \geqslant \vartheta$ in $\mathcal{M}$ by $fscore(\mathbf{C}) \geqslant \vartheta$ in $\mathcal{M}_{rEff,Cause}$. This results in a decision problem GPR-f-score for GPR causes which is in Σ_2^{P} by Theorem 77.

For computing optimal GPR causes as well as f-score-optimal SPR causes we try all cause candidates by computing the related value ($recall$, $covrat$ or $fscore$) and always store the best one so far. As the space for the cause can be reused, this results in a polynomial space algorithm. □

5.3 ω-Regular Co-Safety Causes

In this section we discuss a further extension to our probability-raising causality framework. For this, we also consider causes to be regular sets of executions. However, to account for the *temporal priority* of causes, the property of causes to occur before their effects, we restrict the causes to ω-regular co-safety properties $\mathcal{L}$. These have the feature that $\mathcal{L}$ is uniquely determined by the regular set of minimal *good* prefixes of words in $\mathcal{L}$. Recall that a good prefix π for $\mathcal{L}$ is a finite word such that all infinite extensions of π belong to $\mathcal{L}$ and that all infinite words in the co-safety language $\mathcal{L}$ have a good prefix. Hence, we can say that a cause $rCause$ occurred as soon as a good prefix for $rCause$ has been produced. Moreover, this extension comes natural as the reachability property of the non-terminal set $Cause$ in an MDP is a special case of co-safety property.

For this subsection we will denote regular effects and causes mainly by $rEff$ and $rCause$. Thus, we are given an MDP $\mathcal{M} = (S, Act, \mathrm{P}, init)$ together with a non-empty ω-regular language $rEff \subseteq S^\omega$. In the following formal definition, we use finite words $\sigma \in S^*$ to denote the event σS^ω.

Definition 99 (Co-Safety SPR/GPR Causes). Let $\mathcal{M}$ be an MDP with state space S and let $rEff \subseteq S^\omega$ be a non-empty ω-regular language. A non-empty ω-regular co-safety language $rCause \subseteq S^\omega$ is a *co-safety GPR cause* for $rEff$ in $\mathcal{M}$ if the following condition **(coG)** holds:

(coG) For each scheduler $\mathfrak{S}$ where $\Pr^{\mathfrak{S}}(rCause) > 0$ we have

$$\Pr^{\mathfrak{S}}(rEff \mid rCause) \;>\; \Pr^{\mathfrak{S}}(rEff). \tag{cosafeGPR}$$

The event $rCause$ is a *co-safety SPR cause* for $rEff$ in $\mathcal{M}$ if the following condition **(coS)** holds:

(coS) For each minimal good prefix σ for $rCause$ and each scheduler $\mathfrak{S}$ with $\Pr^{\mathfrak{S}}(\sigma) > 0$:

$$\Pr^{\mathfrak{S}}(rEff \mid \sigma) \;>\; \Pr^{\mathfrak{S}}(rEff). \tag{cosafeSPR}$$

◁

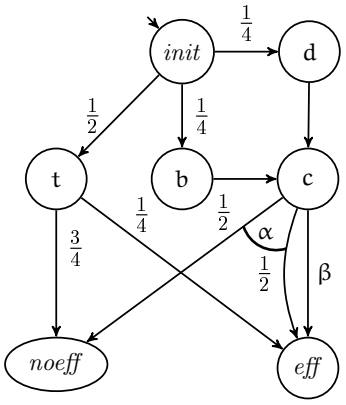

Figure 5.8: An MDP $\mathfrak{M}$ in which $\Diamond \mathbf{c}$ is a reachability SPR cause but no co-safety SPR cause

As mentioned beforehand, a reachability GPR cause is a special case of co-safety GPR cause as for a given set Cause the property $\Diamond \mathit{Cause}$ is co-safety. However, reachability SPR causes can not necessarily be considered as co-safety SPR causes. This comes from the fact, that the condition $(\neg \mathit{Cause})\,\mathrm{U}\,\mathbf{c}$ (for $\mathbf{c} \in \mathit{Cause}$) corresponds to a *set* of minimal good prefixes for $\Diamond \mathit{Cause}$ instead of only a single prefix.

Example 100 (Reachability SPR is not Co-Safety SPR). Consider the MDP $\mathfrak{M}$ from Figure 5.8 with $\mathit{rEff} = \Diamond \mathit{eff}$. For every scheduler $\mathfrak{S}$ with $\Pr^{\mathfrak{S}}_{\mathfrak{M}}(\Diamond \mathbf{c}) > 0$ we have

$$\Pr^{\mathfrak{S}}(\Diamond \mathit{eff} \mid \Diamond \mathbf{c}) > \Pr^{\mathfrak{S}}(\Diamond \mathit{eff})$$

and thus $\mathbf{c}$ is a reachability SPR cause for $\Diamond \mathit{eff}$. On the other hand for the scheduler $\mathfrak{T}$, which chooses $\boldsymbol{\alpha}$ after the path $\boldsymbol{\pi} = \mathbf{a}\ \mathbf{b}\ \mathbf{c}$ and $\boldsymbol{\beta}$ otherwise, we have

$$\Pr^{\mathfrak{T}}(\Diamond \mathit{eff} \mid \boldsymbol{\pi}) = \frac{1}{2} = \Pr^{\mathfrak{T}}(\Diamond \mathit{eff}).$$

However, the word $\boldsymbol{\pi}$ is also a minimal good prefix for $(\neg \mathbf{c})\,\mathrm{U}\,\mathbf{c} = \Diamond \mathbf{c}$ and thus $\mathbf{c}$ is not a co-safety SPR cause for $\Diamond \mathit{eff}$. ◁

Nonetheless, we can still use similar arguments as in Proposition 15 to establish that co-safety SPR causes are also co-safety GPR causes. In particular, the conditional probability $\Pr^{\mathfrak{S}}(\mathit{rEff} \mid \mathit{rCause})$ in condition (cosafeGPR) is a weighted sum of the conditional probabilities $\Pr^{\mathfrak{S}}(\mathit{rEff} \mid \boldsymbol{\sigma})$ in the condition (cosafeSPR). For checking co-safety GPR causality we again propose a polynomial-time model transformation of $\mathfrak{M}$ such that both rCause and rEff are described by reachability properties. Therefore, the results for state-based GPR causes transfer to co-safety GPR causes. However, the same reduction does not work for co-safety SPR causes. This leads to the algorithmic analysis summarized in Table 5.2.

co-safety PR-causality	Check PR condition $rCause$ and $rEff$ given	Compute PR cause $rCause$ for given effect $rEff$	
		$recall$-optimal	$fscore$-optimal
GPR	$\in$ coNP (Lem.102)	believed to be unfeasible (Ex. 107)	believed to be unfeasible (Ex. 107)
SPR	not known to be feasible (Rem.104)	optimal cause known computation unclear (Rem.106)	believed to be unfeasible (Ex. 107)

Table 5.2: Complexity results for co-safety PR with fixed regular effect $rEff$

5.3.1 Checking Co-Safety PR causes

We represent co-safety PR causes $rCause$ as DFAs $\mathcal{A}_{rCause}$ (Def. 11) which accept good prefixes of the represented ω-regular event. Note that, for any ω-regular co-safety property, there is a DFA accepting exactly the minimal good prefixes. So, we will restrict to such DFAs that accept the minimal good prefixes of an ω-regular co-safety property. Such a DFA can never accept a word w as well as a proper prefix v of w.

Let now $\mathcal{M}$ be an MDP, $rEff$ an effect given by the DRA $\mathcal{A}_{rEff}$ and $rCause$ a cause candidate given by a DFA $\mathcal{A}_{rCause}$ as above. In particular, $\mathcal{A}_{rCause}$ accepts exactly the minimal good prefixes for $rCause$. We consider the decision problem, whether $rCause$ is a co-safety SPR cause (resp. co-safety GPR cause). For the special case of Markov chains the check can be done in polynomial time analogously to reachability PR causes by computing the corresponding conditional probabilities [Bai+14]. For co-safety SPR causes this needs a bit more care, as we need to consider conditional probabilities in the product Markov chain $\mathcal{M} \otimes \mathcal{A}_{rCause}$ to specify each minimal good prefix individually. For general MDPs $\mathcal{M}$, we provide a model transformation of using both automata such that the resulting MDP has no end-components and the effect is a reachability property again, similar to Notation 89.

Notation 101 (Canonical MDP for Co-Safety PR Causes). Let $\mathcal{N} = \mathcal{M} \otimes \mathcal{A}_{rEff} \otimes \mathcal{A}_{rCause}$. This product $\mathcal{N}$ is an MDP equipped with two kinds of acceptance conditions. The Rabin acceptance of $\mathcal{A}_{rEff}$ in the second component of each state and the acceptance condition of $\mathcal{A}_{rCause}$ in the third component. Now, let $rCause_{\mathcal{A}_{rEff}}$ be the set of all states of $\mathcal{N}$ whose third component is accepting in $\mathcal{A}_{rCause}$ and which are reachable from the initial state.

As in Notation 89, we construct an MDP $\mathcal{N}'$ by introducing a mode before $rCause_{\mathcal{A}_{rEff}}$ and a mode after $rCause_{\mathcal{A}_{rEff}}$. We then take the MEC-quotient of $\mathcal{N}'$ and add the four fresh terminal states TP, FP, FN, and TN. These states are reachable from states $s_{\mathcal{E}}$ which result from collapsing the MEC $\mathcal{E}$. Depending on whether $\mathcal{E}$ is contained in the before- or after-$rCause_{\mathcal{A}_{rEff}}$ mode and additionally whether the acceptance condition of $\mathcal{A}_{rEff}$ can be realized with prob-

ability 0 or 1, respectively, in $\mathcal{E}$, we define transitions to TP, FP, FN and TN analogous to Notation 89. The resulting MDP is called $\mathcal{M}_{rEff,rCause}$. We emphasize that this MDP contains all states in $rCause_{\mathcal{A}_{rEff}}$ as they are not contained in any end-component. ◁

We start with the observation, that for co-safety GPR causes this reduction fully characterizes the condition **(coG)** .

Lemma 102 (Characterization of Co-Safety GPR Causes). *Let $\mathcal{M}$ be and MDP, $\mathcal{A}_{rEff}$ an DRA, and $\mathcal{A}_{rCause}$ a DFA as above and let $\mathcal{M}_{rEff,rCause}$ be the constructed MDP that contains the set $rCause_{\mathcal{A}_{rEff}}$ of reachable states that have an accepting $\mathcal{A}_{rCause}$-component. Then, $rCause$ is a co-safety GPR cause for $rEff$ in $\mathcal{M}$ if and only if the set of states $rCause_{\mathcal{A}_{rEff}}$ is a GPR cause for $\{TP, FN\}$ in $\mathcal{M}_{rEff,rCause}$.*

Proof. The set $Cause_{\mathcal{K}} = rCause_{\mathcal{A}_{rEff}}$ in $\mathcal{K} = \mathcal{M}_{rEff,rCause}$ satisfies the minimality condition $\Pr^{\max}_{\mathcal{K}}(\neg rCause_{\mathcal{A}_{rEff}} \,\mathrm{U}\, \mathbf{c}) > 0$ for each $\mathbf{c} \in Cause_{\mathcal{K}}$ by construction, since all states in $Cause_{\mathcal{K}}$ are reachable and a run cannot reach two different states in $Cause_{\mathcal{K}}$. Thus, the transformation from $\mathcal{M}$ to $\mathcal{K}$ ensures the minimality condition.

As in the proof of Lemma 91 we now construct a corresponding pair of scheduler between $\mathcal{M}$ and $\mathcal{K}$ such that the probabilities included in (cosafeGPR) are the same under both scheduler in their respective MDP.

For this, let $\mathfrak{T}$ be a scheduler for $\mathcal{K}$. We define a corresponding scheduler $\mathfrak{S}$ for $\mathcal{M}$ by letting $\mathfrak{S}$ mimic the behavior of $\mathfrak{T}$ outside of end-components. As long as $\mathfrak{T}$ moves through the MEC-quotient of $\mathcal{N}'$, the scheduler $\mathfrak{S}$ follows this behavior by leaving MECs through the corresponding actions in $\mathcal{M}$. Whenever $\mathfrak{T}$ moves to one of the states TP and FN through a state $s_{\mathcal{E}}$ obtained from a MEC $\mathcal{E}$ of $\mathcal{M}$, $\mathfrak{S}$ will stay in $\mathcal{E}$ while ensuring with probability 1 that the resulting run is accepted by $\mathcal{A}_{rEff}$. Similarly, if $\mathfrak{T}$ moves to TN or FP, the scheduler $\mathfrak{S}$ for $\mathcal{M}$ stays in the corresponding MEC and realizes the acceptance condition of $\mathcal{A}_{rEff}$ with probability 0.

Vice versa, a scheduler $\mathfrak{S}$ for $\mathcal{M}$ can be mimicked by a scheduler $\mathfrak{T}$ for $\mathcal{K}$ analogously (see also the proof of Lemma 91).

For such a pair of schedulers $\mathfrak{S}$ of $\mathcal{M}$ and $\mathfrak{T}$ of $\mathcal{K}$, we observe that

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(rCause) = \Pr^{\mathfrak{T}}_{\mathcal{K}}(\Diamond Cause_{\mathcal{K}}), \tag{5.4}$$

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(rEff) = \Pr^{\mathfrak{T}}_{\mathcal{K}}(\Diamond\{TP, FN\}), \text{ and} \tag{5.5}$$

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(rCause \wedge rEff) = \Pr^{\mathfrak{T}}_{\mathcal{K}}(\Diamond TP) = \Pr^{\mathfrak{T}}_{\mathcal{K}}(\Diamond Cause_{\mathcal{K}} \wedge \Diamond\{TP, FN\}). \tag{5.6}$$

Hence, $rCause_{\mathcal{A}_{rEff}}$ is a GPR cause for $\{TP, FN\}$ in $\mathcal{M}_{rEff,rCause}$ if and only if $rCause$ is a co-safety GPR cause for $rEff$ in $\mathcal{M}$. □

For co-safety SPR causes this is not a full characterization but only works for one direction.

Lemma 103 (Product Construction for Co-Safety SPR Causes). *Let $\mathcal{M}$, $\mathcal{A}_{rEff}$, and $\mathcal{A}_{rCause}$ be as above and let $\mathcal{M}_{rEff,rCause}$ be the constructed MDP that contains the set $rCause_{\mathcal{A}_{rEff}}$ of reachable states that have an accepting $\mathcal{A}_{rCause}$-component. If $rCause$ is a co-safety SPR cause for $rEff$ in $\mathcal{M}$, then the set of states $rCause_{\mathcal{A}_{rEff}}$ is an SPR cause for $\{TP, FN\}$ in $\mathcal{M}_{rEff,rCause}$.*

Proof. Suppose $rCause$ is a regular SPR cause for $rEff$ in $\mathcal{M}$. Let $\mathfrak{c} \in Cause_{\mathcal{K}} = rCause_{\mathcal{A}_{rEff}}$ and let $\mathfrak{T}$ be a scheduler for $\mathcal{K} = \mathcal{M}_{rEff,rCause}$ such that $\Pr^{\mathfrak{T}}_{\mathcal{K}}((\neg Cause)_{\mathcal{K}} \,\mathrm{U}\, \mathfrak{c}) > 0$. Consider the set of paths

$$\Pi = \{\pi \in Paths(\mathcal{K}) \mid \pi \models (\neg Cause)_{\mathcal{K}} \,\mathrm{U}\, \mathfrak{c} \text{ and } \Pr^{\mathfrak{T}}_{\mathcal{K}}(\pi) > 0\}.$$

Let $\mathfrak{S}$ be a scheduler for $\mathcal{M}$ mimicking $\mathfrak{T}$ as described in the proof of Lemma 102. Let Σ_π be the set of $\mathfrak{T}$-paths in $\mathcal{M}$ that correspond to the path π in Π. Thus, a path σ belongs to Σ_π if it moves through the MECs of $\mathcal{M}$ in the same way as the path π moves through the MEC-quotient until σ and π reach the state $\mathfrak{c}$ and if furthermore, σ has positive probability under $\mathfrak{S}$. Now,

$$\Pr^{\mathfrak{T}}_{\mathcal{K}}(\Diamond\{TP, FN\} \mid (\neg Cause)_{\mathcal{K}} \,\mathrm{U}\, \mathfrak{c}) = \frac{\sum_{\pi\in\Pi}\sum_{\sigma\in\Sigma_\pi} \Pr^{\mathfrak{S}}_{\mathcal{M}}(rEff \mid \sigma)\cdot \Pr^{\mathfrak{S}}_{\mathcal{M}}(\sigma)}{\sum_{\pi\in\Pi}\sum_{\sigma\in\Sigma_\pi} \Pr^{\mathfrak{S}}_{\mathcal{M}}(\sigma)}.$$

All the terms $\Pr^{\mathfrak{S}}_{\mathcal{M}}(rEff \mid \sigma)$ are greater than $\Pr^{\mathfrak{S}}_{\mathcal{M}}(rEff) = \Pr^{\mathfrak{T}}_{\mathcal{K}}(\Diamond\{TP, FN\})$. Hence, we conclude that

$$\Pr^{\mathfrak{T}}_{\mathcal{M}_{rEff,rCause}}(\Diamond\{TP, FN\} \mid \neg rCause_{\mathcal{A}_{rEff}} \,\mathrm{U}\, \mathfrak{c}) > \Pr^{\mathfrak{T}}_{\mathcal{M}_{rEff,rCause}}(\Diamond\{TP, FN\}).$$

□

However, for checking co-safety SPR causality this partial characterization is not sufficient. The underlying problem, in which co-safety SPR causes and SPR causes differ, can be seen by revisiting Example 100.

Remark 104 (Inherent Difficulty of Co-Safety SPR Causes). Again, consider the MDP $\mathcal{M}$ from Figure 5.8 with $rEff = \Diamond eff$. In Example 100 we showed that $\mathfrak{c}$ is a reachability SPR cause for $\Diamond eff$. As $\mathfrak{c} \neq eff$ this also implies that $\mathfrak{c}$ is a state-based SPR cause for eff in $\mathcal{M}$. Using the scheduler $\mathfrak{T}$, which chooses α after the path $\pi = \mathfrak{a}\,\mathfrak{b}\,\mathfrak{c}$ and β otherwise the equation

$$\Pr^{\mathfrak{T}}(\Diamond eff \mid \pi) = \frac{1}{2} = \Pr^{\mathfrak{T}}(\Diamond eff).$$

shows, that the minimal good prefix π for $(\neg\mathfrak{c})\,\mathrm{U}\,\mathfrak{c} = \Diamond\mathfrak{c}$ is a witness that **(coS)** does not hold. This also implies that the desired reduction as in Lemma 102 does not work for $\mathcal{M}$. Note that the violation of the condition **(coS)** is only possible in this example if the scheduler behaves differently depending on how state $\mathfrak{c}$ is reached. This different behavior, however, does not have anything to do with the effect, and potentially different residual properties that have to be satisfied to achieve the effect; in the example, the effect is just a reachability property.

Furthermore, we emphasize that the concrete probabilities of the individual paths leading to $\mathbf{c}$ are important for the violation. In general, this imposes a major challenge for checking the condition **(coS)** , which is not yet known to have a solution. Similar problems arise when trying to check the existence of a co-safety SPR cause. A witness for such existence might be one individual path, potentially only together with a scheduler that realizes this path with very low probability. ◁

5.3.2 Quality and Optimality of Co-Safety PR Causes

We define the quality measures $recall$, $covrat$, and $fscore$ for co-safety PR causes, analogous to Section 5.2.2 by defining the values of the confusion matrix (cf. Figure 3.14) and then taking the infimum over values $quality^{\mathfrak{S}}$ for all schedulers for which $quality^{\mathfrak{S}}$ is defined ($quality \in \{recall, covrat, fscore\}$). With the construction of $\mathcal{M}_{rEff,rCause}$ and the correspondence between schedulers $\mathfrak{S}$ for $\mathcal{M}$ and $\mathfrak{T}$ for $\mathcal{M}_{rEff,rCause}$ satisfying Equations (5.4)-(5.6) established in the proof of Lemma 102, we obtain analogous results to the setting of reachability PR causes:

Corollary 105 (Computing Quality Measures of Co-Safety PR Causes). *Let $\mathcal{M}$ be an MDP and $rEff \subseteq S^\omega$ a non-empty ω-regular language given as DRA $\mathcal{A}_{rEff}$. Given a co-safety SPR/GPR cause $rCause \subseteq S^\omega$ by a DFA $\mathcal{A}_{rCause}$, we can compute $recall(rCause)$, $covrat(rCause)$ and $fscore(rCause)$ in polynomial time.*

Proof. The proof uses the same argumentation as in Lemma 94 to establish that the pairs of schedulers $\mathfrak{S}$ for $\mathcal{M}$ and $\mathfrak{T}$ for $\mathcal{M}_{rEff,rCause}$ retain the same values for the probabilities of the confusion matrix. By reduction we can then use Lemma 60 and Theorem 67. □

Quality-Optimal Co-Safety PR Causes

Already for reachability PR causes, we have seen that without further restrictions on the allowed causes, they might be trivial and intuitively violate the idea of temporal priority (cf. Section 5.2.2). Hence, also here, we impose an additional condition, a variation of the condition (TempPrio) used above. In line with the difference between the definitions of reachability PR causes and co-safety PR causes, we require that after any minimal good prefix σ of a co-safety cause $rCause$, the probability that effect $rEff$ will occur is not guaranteed to be 1. In particular, we require that

$$\Pr^{\min}(rEff \mid \sigma) < 1 \text{ for all minimal good prefixes } \sigma \text{ of } rCause. \qquad \text{(TempPrio2)}$$

Unfortunately, we will observe that there are more obstacles for finding quality-optimal co-safety PR causes.

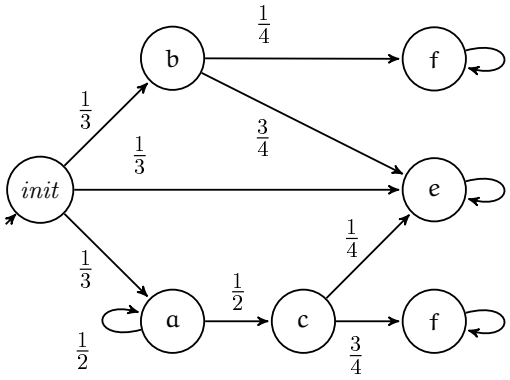

Figure 5.9: Markov chain $\mathfrak{M}$ where $\Diamond \mathbf{e}$ has no recall-optimal co-safety PR-cause.

Remark 106 (Canonical Co-Safety SPR Cause). Following the observations from Theorem 70 we can define a canonical co-safety PR cause which is an optimal co-safety SPR cause for both recall and coverage ratio. In this fully path-based setting this canonical cause consists of all minimal paths which are singleton co-safety SPR causes. This set obviously yields a path-wise coverage over all possible probability-raising paths. However, since we are not aware of a feasible way to check **(coS)** , the computation of this cause is unclear. ◁

For co-safety GPR causes, the following example illustrates that there may be no recall-optimal causes which respect (TempPrio2). Intuitively, the problem is that causes can be pushed arbitrarily close towards a violation of the PR condition while increasing the recall.

Example 107 (No Quality-Optimal Co-Safety PR Cause). In this example we consider a Markov chain $\mathfrak{M}$ with a state $\mathbf{e}$ such that the effect $\mathit{rEff} = \Diamond \mathbf{e}$ has co-safety GPR causes which respect the condition (TempPrio2), but there are no recall-optimal co-safety GPR causes that respect (TempPrio2).

Consider the Markov chain $\mathfrak{M}$ depicted in Figure 5.9 with state space $\mathbf{S}$ and the ω-regular effect $\mathit{rEff} = \Diamond \mathbf{e}$. In $\mathfrak{M}$, we have

$$\Pr(\mathit{rEff}) = \frac{1}{3} \cdot \frac{3}{4} + \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{4} = \frac{2}{3}.$$

Furthermore, clearly the co-safety language $\mathit{init}\, \mathfrak{b}\, \mathbf{S}^{\omega}$ with the unique minimal good prefix $\mathit{init}\, \mathfrak{b}$ is a co-safety GPR cause for rEff in $\mathfrak{M}$ as $\Pr(\mathit{rEff} \mid \mathit{init}\, \mathfrak{b}) = 3/4$.

Note that there cannot be a co-safety GPR cause rCause' that does not have $\mathit{init}\, \mathfrak{b}$ as a minimal good prefix. By (TempPrio2) the minimal good prefixes can not end in $\mathbf{e}$. Also, init is clearly also no candidate for a minimal good prefix of rCause' since this would imply that rCause consists of all paths of $\mathfrak{M}$ and cannot satisfy the condition **(coG)** . If $\mathit{init}\, \mathfrak{b}$ is not a minimal good prefix, all minimal good prefixes have to end in $\mathfrak{a}$, $\mathfrak{c}$, or $\mathfrak{f}$. Afterwards, the probability to reach $\mathbf{e}$ is at most $1/4$. So, $\Pr_{\mathfrak{M}}(\mathit{rEff} \mid \mathit{rCause}') \leqslant 1/4$ as this is a weighted average of values $\Pr_{\mathfrak{M}}(\mathit{rEff} \mid \sigma) \leqslant 1/4$ for the minimal good prefixes σ of rCause'.

Thus, all regular GPR causes have the minimal good prefix $\mathit{init}\,\mathfrak{b}$ together with potentially further minimal good prefixes. Let now

$$rCause_{\mathfrak{p}} = \mathit{init}\,\mathfrak{b}\,\mathrm{S}^{\omega} \cup \mathrm{A}_{\mathfrak{p}}$$

where $\mathrm{A}_{\mathfrak{p}}$ is a regular subset of the paths $\{\mathit{init}\,\mathfrak{a}^{\mathfrak{k}}\,\mathfrak{c} \mid \mathfrak{k} \geqslant 1\}$ such that all the paths in $\mathrm{A}_{\mathfrak{p}}$ together have probability mass $\frac{1}{3}\mathfrak{p}$. Note that we can find such a set $\mathrm{A}_{\mathfrak{p}}$ for a dense set of values $\mathfrak{p} \in [0,1]$. We compute

$$\Pr(rCause_{\mathfrak{p}}) = \frac{1}{3}(1+\mathfrak{p}) \quad \text{and} \quad \Pr(rCause_{\mathfrak{p}} \wedge rEff) = \frac{1}{4} + \frac{1}{12}\cdot\mathfrak{p}.$$

Therefore, we obtain that $rCause_{\mathfrak{p}}$ is a co-safety GPR cause if and only if

$$\begin{aligned} \Pr(rEff \mid rCause_{\mathfrak{p}}) \; &= \frac{\frac{1}{4} + \frac{1}{12}\cdot\mathfrak{p}}{\frac{1}{3}(1+\mathfrak{p})} > \frac{2}{3} = \Pr(rEff) \\ \iff 9 + 3\cdot\mathfrak{p} > 8 + \mathfrak{p} \iff \mathfrak{p} &< \frac{1}{5}, \end{aligned}$$

where the first equivalence is obtained by multiplying the inequality with $12(1+\mathfrak{p})$. The recall of $rCause_{\mathfrak{p}}$ is

$$recall(rCause_{\mathfrak{p}}) = \Pr(rCause_{\mathfrak{p}} \mid rEff) = \frac{\frac{1}{4} + \frac{1}{12}\cdot\mathfrak{p}}{2/3} = \frac{3}{8} + \frac{1}{8}\cdot\mathfrak{p}.$$

So, among the co-safety GPR causes of the form $rCause_{\mathfrak{p}}$, there is no recall-optimal one. For $\mathfrak{p}$ tending to $1/5$ from below, the recall always increases.

Even for approximating a recall-optimal co-safety GPR cause this shows that the computation becomes very complicated. Given an approximation constant $\varepsilon > 0$ an ε-recall-optimal co-safety GPR cause must take a very complicated form. It has to select paths of the form $\mathit{init}\,\mathfrak{a}^{\mathfrak{k}}\,\mathfrak{c}$ that have probability $\frac{1}{3\cdot 2^{\mathfrak{k}}}$ such that their probability adds up to a value less than, but close to $\frac{1}{15}$. ◁

We have seen that for recall-optimal (and hence coverage ratio-optimal) co-safety SPR causes for an effect given by $\mathcal{A}_{rEff}$, we can provide a characterization by defining a path-based canonical cause. However, it is unclear how this canonical cause can be computed, since we do not have a method to check the co-safety SPR condition **(coS)** . Moreover, there might be some paths ending in a given state in the product of $\mathfrak{M}$ and $\mathcal{A}_{rEff}$ that belong to the canonical cause while other paths may end in that state do not. For co-safety GPR causes, we have even seen that there might be no (non-trivial) optimal causes. Furthermore, Example 107 shows that and that causes close to the optimum can take a very complicated shape even for simple reachability properties as the effect. Since the f-score is a more involved quality measure than the recall, it can be expected that finding f-score optimal co-safety PR causes involves even more complications.

Concluding this chapter, we note that extending the state-based PR conditions is not straight forward and imposes additional problems, with the most prevalent one being temporal priority. Whenever both the cause and effect are ω-regular properties, it is not clear how to formulate the temporal priority of the cause. Even if in theory these properties are concerning infinite runs, intuitively they may hold already at some finite time point, for example when the probability of satisfaction is 1. This makes it difficult to impose a temporal aspect to the property. However, we can require temporal priority to each individual word of the property as in (TempPrio2) which then has the consequence that quality-optimal causes might not exist as in the above example.

6 Conclusion

In this thesis we introduced and discussed two notions of *probability-raising cause* in MDPs. For this we formalized the probability-raising condition and pair it with temporal priority. This resulted in a strict and a global condition for probabilistic causality. We considered several quality notions for causes in order to distinguish between the predictive power of causes. Moreover, we covered fundamental algorithmic problems like checking the conditions or computing optimal causes. Furthermore, we extended the state-based setting for causality to ω-regular events as cause and effect.

In this final section we conclude the thesis by returning to our notion of probability-raising cause with a critical look. For this, we will discuss the conditions imposed by GPR and SPR causes on the non-determinism of the MDP. This discussion will include further directions of research for probabilistic causality in probabilistic models as well as possible modifications of our notions. Moreover, we will discuss the interplay between the strict and global probability-raising condition in both meaning and algorithmic properties. As a final consideration we take a look at our discussed extensions to path-based events as cause and effect in the context of Markov chains.

Further Criteria for Probabilistic Causes

In this thesis we work with a notion of probabilistic causality in MDPs which is independent from the non-determinism by letting the (PR) condition hold *for all* scheduler of the MDP. This implies that the considered event Cause is *always* a probabilistic cause for the event Eff,independent from the choices of the user and the environment. We have discussed the soundness and meaning of this definition in length in Sections 3.1 and 3.2. However, from a formal standpoint this is a strong condition on the MDP since it is sufficient if only one scheduler exists which refutes the (PR) condition in order for Cause to not have the claimed causal relation to Eff. We have already stated a weaker probability-raising condition by defining potential PR causes in Section 5.1, where Cause has a causal relation with Eff

whenever there is at least one scheduler for which *PR* holds. This is the other side of the spectrum where the causality condition may be to weak as it can be the case that for "most" scheduler the *PR* condition does not hold (also see Example 80).

Thus, it makes sense to consider the strength of a cause as the "amount" of situations in which it is exemplifying probability-raising of the effect. Consistently with above argumentation general and potential causes are then defining opposite sites of this spectrum. For general PR causes the probability-raising happens in *all* situations (for all schedulers of the MDP) whereas for potential PR causes this only need to hold for *at least one* situation. Formally, we can define the causal strength of a potential PR cause is a fraction of the amount of scheduler which are probability-raising scheduler by Definition 78 to the amount of all possible scheduler of the MDP. This idea motivates notions of probabilistic causes in MDPs which are stronger than potential PR causes but weaker than general PR causes. For example, a potential PR cause can be considered *genuine* if its causal strength exceeds a certain threshold $\mathfrak{f} \in [0, 1]$. This threshold then governs the minimal "strength" of a genuine PR cause. For $\mathfrak{f} = 0.5$ a set $Cause$ would then be considered probabilistically causal for Eff if there are more scheduler which are above a set threshold than there are below. Extending this idea leads to the weighting of scheduler by a suitable quality-measure like the f-score or Mathews correlation coefficient (MCC) for a notion of "weighted strength". For the MCC a natural threshold would for example be 0 as an MCC above 0 for a set $Cause$ implies that $Cause$ is a GPR cause (as can be seen in Lemma 50). These considerations are part of ongoing research.

Non-Strict Inequality in the PR Condition. The approach to probability-raising causality in this thesis is in line with the classical notion from philosophical literature which uses a strict inequality in the (PR) condition. As a consequence causes might not exist as argued in Remark 17. However, we can also consider a relaxed PR condition by only requiring a non-strict inequality. This is apparently a minor change that broadens the choice of causes. Indeed, the proposed algorithms for checking the GPR and SPR condition for state-based effects (Chapter 4.1) can easily be modified for the relaxed definition. As the algorithms of both extended settings discussed in Sections 5.2 and 5.3 stem mainly from a reduction to state-based effects this also holds for reachability and co-safety causes of regular effects. However, a non-strict inequality in the PR condition also leads to a questionable notion of causality, as e.g. $\{init\}$ would always be a recall- and ratio-optimal cause. Therefore, the relaxed PR condition needs to be complemented by other side constraints in order to be used properly. An example would be to require the non-strict inequality for all schedulers that reach a cause with positive probability and also requiring the existence of a PR scheduler with the strict inequality. This might be a useful alternative definition which agrees with Definition 12 for Markov chains.

Relaxing the Minimality Condition. As many causality notions in the literature use some kind of minimality constraint, we also include the condition $\Pr^{\max}_{\mathcal{M}}((\neg \mathit{Cause})\,\mathrm{U}\,\mathbf{c}) > 0$ for all states of Cause in the state-based setting and for reachability PR causes of regular effects. This condition can easily be checked, and obviously depends on the structure of the considered MDP. However, our minimality condition could also be dropped without affecting the algorithmic results presented here. This can be useful when the task is to identify components or agents which are responsible for the occurrences of undesired effects. In these cases the cause candidates are fixed (e.g., for each agent $\mathfrak{i}$, the set of states controlled by agent $\mathfrak{i}$), but some of them might violate the minimality condition.

Strict vs. Global Probability-Raising Causes

In this thesis we distinguish between two types of causes, namely global and strict PR causes. The global PR condition **(G)** takes the set Cause (either a set of states or paths) as a unit and then imposes the (PR) condition. The strict PR condition **(S)** considers the individual elements of Cause and imposes the PR condition individually. Thus for SPR causes each individual element is a cause by definition and the notion is stable under union. This also establishes a clear implication from SPR causality to GPR causality (Proposition 15). However, GPR causes also take the distribution under which Cause is reached under consideration for the PR condition, which is not the case for SPR causes. This is one of the reasons why GPR causes are algorithmically more involved as can be seen by the coNP complexity for the GPR check in comparison to the P complexity of the SPR check. For finding quality-optimal causes the more general GPR condition is computationally hard (Theorem 77) while for SPR causes there can be simpler optimal causes (Theorem 70) computable in polynomial time. On the other side, this also means that the GPR condition is more flexible to achieve better quality-measures.

When we extend PR causality to ω-regular effects given by deterministic Rabin automata (DRA) to reachability PR causes (Def. 88), the SPR condition loses its algorithmic advantage over the GPR condition. The caveat here is that the strict PR condition translates to a global PR condition after a product of the MDP and the DRA. This increases the algorithmic complexity of reachability SPR causes to the level of reachability GPR causes. In the further extension to ω-regular co-safety PR causes in Definition 99, we also note that the global PR condition comes more natural than the splitting of elements in the strict PR condition. Here, the co-safety SPR condition considers each individual path specified by a finite automaton, which is somewhat unnatural. For example, this also leads to phenomena as the non-existence of quality-optimal PR causes (discussed in Example 107). Rather, it is more natural to consider such a co-safety path property as a whole. The co-safety GPR condition then basically states that the co-safety property causes the effect property, instead of saying all possible realizations of the co-safety path property cause the effect.

Optimal Causes. Since there can be many probability-raising causes for a given effect we have introduced quality-measures to distinguish them by their predictive power measured by statistical means. For this we interpret PR causes as binary classifiers which predict the effect. As we have seen by our algorithmic considerations that quality-measures also are a tool for computing causes as they give us a criteria for which we can find optimal causes. Without such criteria the computed PR causes from the existence check may be very trivial (Lemma 48 and Remark 69). However, we show in Theorem 77 computational hardness for finding optimal causes for the global PR condition. Also, for more involved measures like the f-score it is also algorithmically difficult to compute optimal SPR causes (Thm. 75). On one hand, this opens up further research directions on the practical side to find heuristics and methods to efficiently compute quality-optimal causes. On the other hand, this also motivates to look for further criteria for "good" causes.

Extension to Path-Based Settings in Markov Chains

We discussed the algorithmic problems arising from possible extensions to path-based events as cause and effect in Sections 5.2 and 5.3. It is clear that any method to check or derive probabilistic causes that are a proper extension of our state-based PR causes must be at least as complex. However, this lower bound is not a problem when considering Markov chains as the base model in most cases. As noted throughout the thesis, the complexity of state-based probability-raising causality in DTMCs seems to be reasonably low (also confer Table 4.2). Most notably, this does not change when considering path-based notions. A product construction of the original DTMC with an automata which represent causes and effects only introduces a polynomial overhead in complexity but reduces causes and effects to state-based reachability properties, as shown in 102. Thus, this allows us to consider more involved queries on the causes.

One approach is to introduce a state-based cost function on the Markov chain, which could, for example, model the costs for monitoring the system or the cost of failure at the current point. In our publications [Zie+22] and [Bai+21c] we consider a notion of causality which combines the PR condition together with a complete covering of the effect. That is, almost all paths which lead to the effect have a prefix which reaches the cause. In the setting of this thesis, this condition corresponds to a recall of 1. Requiring these two conditions makes the so-called $\mathbf{p}$-causes of [Bai+21c] suitable when causes are used for monitoring purposes. Assuming the effect describes a faulty behavior or another critical event the system should avoid. Then, a $\mathbf{p}$-cause can be used to trigger an alarm as soon as the path is part of the cause, which is before the critical event occurs.

Returning to the state-based cost function, tailoring causes to be cost-optimal with respect to different ways of incurring cost is thoroughly explored in our publications [Zie+22] and [Bai+21c]. It turns out that most algorithmic problems that arise when computing cost-

optimal causes can be reduced to a stochastic shortest path problem in a simple MDP constructed from the original DTMC. The idea is, that the introduced non-deterministic actions govern which states (or paths) are part of the cost-optimal cause, similar to the game construction used for the computation of f-score optimal SPR causes in DTMCs (Theorem 73). When considering partial expected cost, the computation of cost-optimal causes is PP-hard even for non-negative weight functions using a reduction to the cost problem in acyclic MCs introduced in [HK15a].

We can extend the consideration of cost-functions of [Bai+21c] to proper MDPs as in the setting of this thesis. The resulting computational problems then search for cost-optimal PR causes with a set recall of 1. Since for general expected cost the DTMC setting is using a reduction to a stochastic shortest path problem in an MDP, for proper MDPs this will result in stochastic shortest path games as in the decision problem for f-score optimal PR causes (Theorem 75). For partial expected cost in proper MDPs the computation is obviously also PP-hard, but expectedly even harder. A complete formal consideration of cost-optimal probability-raising causes in MDPs in the spirit of above descriptions remains as future work.

Bibliography

[ÁB18] Erika Ábrahám and Borzoo Bonakdarpour. “HyperPCTL: A Temporal Logic for Probabilistic Hyperproperties”. In: *15th International Conference on Quantitative Evaluation of Systems (QEST)*. 2018, pp. 20–35.

[Alf97] Luca de Alfaro. “Formal Verification of Probabilistic Systems”. PhD Thesis. Stanford, USA: Stanford University, 1997.

[Alf99] Luca de Alfaro. “Computing Minimum and Maximum Reachability Times in Probabilistic Systems”. In: *10th International Conference on Concurrency Theory (CONCUR)*. Ed. by Jos C. M. Baeten and Sjouke Mauw. Vol. 1664. Lecture Notes in Computer Science. Springer, 1999, pp. 66–81. URL: https://doi.org/10.1007/3-540-48320-9_7.

[Bai+14] Christel Baier, Joachim Klein, Sascha Klüppelholz and Steffen Märcker. “Computing Conditional Probabilities in Markovian Models Efficiently”. In: *20th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS)*. Ed. by Erika Ábrahám and Klaus Havelund. Vol. 8413. Lecture Notes in Computer Science. Springer, 2014, pp. 515–530. URL: https://doi.org/10.1007/978-3-642-54862-8_43.

[Bai+17] Christel Baier, Joachim Klein, Sascha Klüppelholz and Sascha Wunderlich. “Maximizing the Conditional Expected Reward for Reaching the Goal”. In: *Tools and Algorithms for the Construction and Analysis of Systems*. Ed. by Axel Legay and Tiziana Margaria. Berlin, Heidelberg: Springer Berlin Heidelberg, 2017, pp. 269–285.

[Bai+18] Christel Baier, Nathalie Bertrand, Clemens Dubslaff, Daniel Gburek and Ocan Sankur. “Stochastic Shortest Paths and Weight-Bounded Properties in Markov Decision Processes”. In: *33rd Annual ACM/IEEE Symposium on Logic in Computer*

Science, LICS 2018, Oxford, UK, July 09-12, 2018. Ed. by Anuj Dawar and Erich Grädel. ACM, 2018, pp. 86–94. URL: https://doi.org/10.1145/3209108.3209184.

[Bai+21a] Christel Baier, Norine Coenen, Bernd Finkbeiner, Florian Funke, Simon Jantsch and Julian Siber. "Causality-Based Game Solving". In: *Computer Aided Verification*. Ed. by Alexandra Silva and K. Rustan M. Leino. Cham: Springer International Publishing, 2021, pp. 894–917.

[Bai+21b] Christel Baier, Clemens Dubslaff, Florian Funke, Simon Jantsch, Rupak Majumdar, Jakob Piribauer and Robin Ziemek. "From Verification to Causality-Based Explications (Invited Talk)". In: *48th International Colloquium on Automata, Languages, and Programming, (ICALP)*. Ed. by Nikhil Bansal, Emanuela Merelli and James Worrell. Vol. 198. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021, 1:1–1:20.

[Bai+21c] Christel Baier, Florian Funke, Simon Jantsch, Jakob Piribauer and Robin Ziemek. "Probabilistic Causes in Markov Chains". In: *Automated Technology for Verification and Analysis*. Ed. by Zhe Hou and Vijay Ganesh. Springer International Publishing, 2021, pp. 205–221. ISBN: 978-3-030-88885-5.

[Bai+22a] Christel Baier, Clemens Dubslaff, Florian Funke, Simon Jantsch, Jakob Piribauer and Robin Ziemek. "Operational Causality – Necessarily Sufficient and Sufficiently Necessary". In: *A Journey from Process Algebra via Timed Automata to Model Learning : Essays Dedicated to Frits Vaandrager on the Occasion of His 60th Birthday*. Ed. by Nils Jansen, Mariëlle Stoelinga and Petra van den Bos. Springer Nature Switzerland, 2022, pp. 27–45. ISBN: 978-3-031-15629-8.

[Bai+22b] Christel Baier, Florian Funke, Jakob Piribauer and Robin Ziemek. "On probability-raising causality in Markov decision processes". In: *Foundations of Software Science and Computation Structures*. 2022, pp. 40–60. ISBN: 978-3-030-99253-8. DOI: 10.1007/978-3-030-99253-8_3.

[BBS95] Andrew G. Barto, Steven J. Bradtke and Satinder P. Singh. "Learning to act using real-time dynamic programming". In: *Artificial Intelligence* 72.1 (1995), pp. 81–138. ISSN: 0004-3702. DOI: https://doi.org/10.1016/0004-3702(94)00011-O. URL: https://www.sciencedirect.com/science/article/pii/000437029400011O.

[BCH23] Sander Beckers, Hana Chockler and Joseph Y. Halpern. *A Causal Analysis of Harm*. 2023. arXiv: 2210.05327 [cs.AI].

[Bee+12] Ilan Beer, Shoham Ben-David, Hana Chockler, Avigail Orni and Richard J. Trefler. "Explaining counterexamples using causality". In: *Formal Methods in System Design* 40.1 (2012), pp. 20–40.

[Bee+15] Adrian Beer, Stephan Heidinger, Uwe Kühne, Florian Leitner-Fischer and Stefan Leue. "Symbolic Causality Checking Using Bounded Model Checking". In: *Model Checking Software*. Ed. by Bernd Fischer and Jaco Geldenhuys. Cham: Springer International Publishing, 2015, pp. 203–221.

[BFM21] Christel Baier, Florian Funke and Rupak Majumdar. "A Game-Theoretic Account of Responsibility Allocation". In: *30th International Joint Conference on Artificial Intelligence (IJCAI)*. Ed. by Zhi-Hua Zhou. ijcai.org, 2021, pp. 1773–1779. URL: https://doi.org/10.24963/ijcai.2021/244.

[BGC09] Christel Baier, Marcus Größer and Frank Ciesinski. "Quantitative analysis under fairness constraints". In: *International Symposium on Automated Technology for Verification and Analysis*. Springer. 2009, pp. 135–150.

[BH12] M. Braham and M. van Hees. "An Anatomy of Moral Responsibility". In: *Mind* 121 (483) (2012), pp. 601–634.

[BK08] Christel Baier and Joost-Pieter Katoen. *Principles of Model Checking (Representation and Mind Series)*. The MIT Press, Cambridge, MA, 2008. ISBN: 026202649X, 9780262026499.

[BPZ24] Christel Baier, Jakob Piribauer and Robin Ziemek. "Foundations of probability-raising causality in Markov decision processes". In: *Logical Methods in Computer Science* Volume 20, Issue 1 (Jan. 2024). DOI: 10.46298/lmcs-20(1:4)2024.

[BT91] Dimitri P. Bertsekas and John N. Tsitsiklis. "An Analysis of Stochastic Shortest Path Problems". In: *Mathemathics of Operations Research* 16.3 (1991), 580–595.

[Can88] John F. Canny. "Some Algebraic and Geometric Computations in PSPACE". In: *20th Annual ACM Symposium on Theory of Computing (STOC)*. ACM, 1988, pp. 460–467.

[Car79] Nancy Cartwright. "Causal Laws and Effective Strategies". In: *Noûs* 13.4 (1979), pp. 419–437.

[CGL18] Georgiana Caltais, Sophie Linnea Guetlein and Stefan Leue. "Causality for General LTL-definable Properties". In: *3rd Workshop on formal reasoning about Causation, Responsibility, and Explanations in Science and Technology (CREST)*. 2018, pp. 1–15.

[CGP99] E. M. Clarke, O. Grumberg and D. Peled. *Model Checking*. MIT Press, 1999.

[CH04] Hana Chockler and Joseph Y. Halpern. "Responsibility and Blame: A Structural-Model Approach". In: *J. Artif. Int. Res.* 22.1 (Oct. 2004), 93–115. ISSN: 1076-9757.

[CHK08] Hana Chockler, Joseph Y. Halpern and Orna Kupferman. "What causes a system to satisfy a specification?" In: *ACM Transactions on Computational Logic* 9.3 (2008), 20:1–20:26.

[CJ20] Davide Chicco and Giuseppe Jurman. "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation". In: *BMC genomics* 21.1 (2020), pp. 1–13.

[Coe+22a] Norine Coenen, Raimund Dachselt, Bernd Finkbeiner, Hadar Frenkel, Christopher Hahn, Tom Horak, Niklas Metzger and Julian Siber. "Explaining Hyperproperty Violations". In: *Computer Aided Verification*. Ed. by Sharon Shoham and Yakir Vizel. Cham: Springer International Publishing, 2022, pp. 407–429.

[Coe+22b] Norine Coenen, Bernd Finkbeiner, Hadar Frenkel, Christopher Hahn, Niklas Metzger and Julian Siber. "Temporal Causality in Reactive Systems". In: *Automated Technology for Verification and Analysis*. Ed. by Ahmed Bouajjani, Lukáš Holík and Zhilin Wu. Cham: Springer International Publishing, 2022, pp. 208–224. ISBN: 978-3-031-19992-9.

[Cpe] *Center for Perspicuous Computing*. accessed on 06.02.2024. URL: https://www.perspicuous-computing.science/research/.

[CR21] S. Marc Cohen and C. D. C. Reeve. "Aristotle's Metaphysics". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2021. Metaphysics Research Lab, Stanford University, 2021.

[DFT20] Rayna Dimitrova, Bernd Finkbeiner and Hazem Torfah. "Probabilistic Hyperproperties of Markov Decision Processes". In: *18th International Symposium on Automated Technology for Verification and Analysis (ATVA)*. Ed. by Dang Van Hung and Oleg Sokolsky. Vol. 12302. Lecture Notes in Computer Science. Springer, 2020, pp. 484–500. URL: https://doi.org/10.1007/978-3-030-59152-6_27.

[Dub21] Clemens Dubslaff. "Quantitative Analysis of Configurable and Reconfigurable Systems". PhD thesis. TU Dresden, Institute for Theoretical Computer Science, Germany, 2021. URL: https://nbn-resolving.org/urn:nbn:de:bsz:14-qucosa2-785434.

[Dub+22] Clemens Dubslaff, Kallistos Weis, Christel Baier and Sven Apel. "Causality in Configurable Software Systems". In: *Proceedings of the 44th International Conference on Software Engineering*. ICSE '22. Pittsburgh, Pennsylvania: Association for Computing Machinery, 2022, 325–337. ISBN: 9781450392211. DOI: 10.1145/3510003.3510200. URL: https://doi.org/10.1145/3510003.3510200.

[Eel91] Ellery Eells. *Probabilistic Causality*. Cambridge Studies in Probability, Induction and Decision Theory. Cambridge University Press, 1991.

[ES83] Ellery Eells and Elliott Sober. "Probabilistic Causality and the Question of Transitivity". In: *Philosophy of Science* 50.1 (1983), pp. 35–57. ISSN: 00318248, 1539767X. URL: http://www.jstor.org/stable/187403 (visited on 06/02/2024).

[Fei+11] Joan Feigenbaum, James A. Hendler, Aaron D. Jaggard, Daniel J. Weitzner and Rebecca N. Wright. "Accountability and Deterrence in Online Life". In: *Proceedings of WebSci '11*. New York, NY, USA: ACM, 2011.

[FG11] Luke Fenton-Glynn. "A Probabilistic Analysis of Causation". In: *The British Journal for the Philosophy of Science* 62 (2011).

[FJB20] Florian Funke, Simon Jantsch and Christel Baier. "Farkas Certificates and Minimal Witnesses for Probabilistic Reachability Constraints". In: *Tools and Algorithms for the Construction and Analysis of Systems*. Ed. by Armin Biere and David Parker. Cham: Springer International Publishing, 2020, pp. 324–345.

[FS23] Bernd Finkbeiner and Julian Siber. "Counterfactuals Modulo Temporal Logics". In: (Nov. 2023). DOI: 10.60882/cispa.24614664.v1.

[GJ79] M. R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979. ISBN: 0-7167-1044-7.

[GJ90] Michael R. Garey and David S. Johnson. *Computers and Intractability; A Guide to the Theory of NP-Completeness*. USA: W. H. Freeman & Co., 1990. ISBN: 0716710455.

[GM13] Gregor Gößler and Daniel Le Métayer. "A General Trace-Based Framework of Logical Causality". In: *Formal Aspects of Component Software - 10th International Symposium (FACS)*. 2013, pp. 157–173.

[Gro+06] Alex Groce, Sager Chaki, Daniel Kroening and Ofer Strichman. "Error explanation with distance metrics". In: *International Journal on Software Tools for Technology Transfer*. 2006, pp. 229–247. DOI: 10.1007/s10009-005-0202-0.

[Gro+20] Timo P. Gros, Holger Hermanns, Jörg Hoffmann, Michaela Klauck and Marcel Steinmetz. "Deep Statistical Model Checking". In: *Formal Techniques for Distributed Objects, Components, and Systems*. Ed. by Alexey Gotsman and Ana Sokolova. Cham: Springer International Publishing, 2020, pp. 96–114. ISBN: 978-3-030-50086-3.

[Hal04] Ned Hall. "Two Concepts of Causation". In: *Causation and Counterfactuals*. Ed. by John Collins, Ned Hall and Laurie Paul. MIT Press, 2004, pp. 225–276.

[Hal15] Joseph Y. Halpern. "A Modification of the Halpern-Pearl Definition of Causality". In: *Proceedings of IJCAI'15*. Buenos Aires, Argentina: AAAI Press, 2015, 3022–3033. ISBN: 9781577357384.

[Hal16] Joseph Y. Halpern. *Actual Causality*. MIT Press, 2016.

[HIH23] Sarah Hiller, Jonas Israel and Jobst Heitzig. "An Axiomatic Approach to Formalized Responsibility Ascription". In: *PRIMA 2022: Principles and Practice of Multi-Agent Systems*. Ed. by Reyhan Aydoğan, Natalia Criado, Jérôme Lang, Victor Sanchez-Anguix and Marc Serramia. Cham: Springer International Publishing, 2023, pp. 435–457.

[Hit21] Christopher Hitchcock. "Probabilistic Causation". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Spring 2021. Metaphysics Research Lab, Stanford University, 2021.

[HK15a] Christoph Haase and Stefan Kiefer. "The Odds of Staying on Budget". In: *Automata, Languages, and Programming*. Springer Berlin Heidelberg, 2015, pp. 234–246. ISBN: 978-3-662-47666-6.

[HK15b] Yuxiao Huang and Samantha Kleinberg. "Fast and Accurate Causal Inference from Time Series Data". In: *Proceedings of FLAIRS 2015*. AAAI Press, 2015, pp. 49–54.

[HP01] Joseph Y. Halpern and Judea Pearl. "Causes and Explanations: A Structural-Model Approach: Part 1: Causes". In: *17th Conference in Uncertainty in Artificial Intelligence (UAI)*. 2001, pp. 194–202.

[HP05a] Joseph Y. Halpern and Judea Pearl. "Causes and Explanations: A Structural-Model Approach. Part I: Causes". In: *The British Journal for the Philosophy of Science* 56.4 (2005), pp. 843–887.

[HP05b] Joseph Y. Halpern and Judea Pearl. "Causes and Explanations: A Structural-Model Approach. Part II: Explanations". In: *The British Journal for the Philosophy of Science* 56.4 (2005), pp. 889–911.

[Hum39] David Hume. *A Treatise of Human Nature*. John Noon, 1739.

[Hum48] David Hume. *An Enquiry Concerning Human Understanding*. London, 1748.

[Ibr+20] Amjad Ibrahim, Alexander Pretschner, Tobias Klesel, Ehsan Zibaei, Severin Kacianka and Alexander Pretschner. "Actual Causality Canvas: A General Framework for Explanation-Based Socio-Technical Constructs". In: *Proceedings of ECAI'20*. IOS Press Ebooks, 2020, pp. 2978 –2985.

[IP20] Amjad Ibrahim and Alexander Pretschner. "From Checking to Inference: Actual Causality Computations as Optimization Problems". In: *Proceedings of ATVA'20*. Cham: Springer Internat. Publishing, 2020, pp. 343–359. ISBN: 978-3-030-59152-6.

[Kal20] Lodewijk Kallenberg. *Lecture Notes Markov Decision Problems - version 2020*. Feb. 2020.

[KH11] Samantha Kleinberg and George Hripcsak. "A review of causal inference for biomedical informatics". In: *J Biomed Inform.* 44.6 (2011), pp. 1102–12.

[Kle10] Samantha Kleinberg. "An algorithmic enquiry concerning causality". PhD thesis. New York University, 2010.

[Kle11] Samantha Kleinberg. "A Logic for Causal Inference in Time Series with Discrete and Continuous Variables". In: *Proceedings of IJCAI'11*. 2011, pp. 943–950.

[Kle12] Semantha Kleinberg. *Causality, Probability and Time*. Cambridge University Press, 2012.

[KLL11] Matthias Kuntz, Florian Leitner-Fischer and Stefan Leue. "From Probabilistic Counterexamples via Causality to Fault Trees". In: *30th International Conference on Computer Safety, Reliability, and Security (SAFECOMP)*. Ed. by Francesco Flammini, Sandro Bologna and Valeria Vittorini. Vol. 6894. Lecture Notes in Computer Science. Springer, 2011, pp. 71–84. URL: https://doi.org/10.1007/978-3-642-24270-0_6.

[KM09] Samantha Kleinberg and Bud Mishra. "The Temporal Logic of Causal Structures". In: *25th Conference on Uncertainty in Artificial Intelligence (UAI)*. 2009, pp. 303–312.

[KM10] Samantha Kleinberg and Bud Mishra. "The Temporal Logic of Token Causes". In: *Proceedings of KR'10*. Toronto, Ontario, Canada: AAAI Press, 2010, 575–577. ISBN: 1577354516.

[KNP11] M. Kwiatkowska, G. Norman and D. Parker. "PRISM 4.0: Verification of Probabilistic Real-time Systems". In: *Proc. 23rd International Conference on Computer Aided Verification (CAV'11)*. Ed. by G. Gopalakrishnan and S. Qadeer. Vol. 6806. LNCS. Springer, 2011, pp. 585–591.

[KP23] Milad Kazemi and Nicola Paoletti. *Causal Temporal Reasoning for Markov Decision Processes*. 2023. arXiv: 2212.08712 [cs.AI].

[KTK80] Mikhail K Kozlov, Sergei P Tarasov and Leonid G Khachiyan. "The polynomial solvability of convex quadratic programming". In: *USSR Computational Mathematics and Mathematical Physics* 20.5 (1980), pp. 223–228.

[Lew73] David Lewis. "Counterfactuals and comparative possibility". In: *Journal of Philosophical Logic* 2.4 (1973), pp. 418–446.

[LL13] Florian Leitner-Fischer and Stefan Leue. "Causality Checking for Complex System Models". In: *14th International Conference on Verification, Model Checking, and Abstract Interpretation, 14th International Conference (VMCAI)*. 2013, pp. 248–267.

[Mär20] Steffen Märcker. "Model checking techniques for design and analysis of future hardware and software systems". PhD thesis. TU Dresden, Germany, 2020. URL: https://d-nb.info/1232958204.

[Mil17] Tim Miller. "Explanation in Artificial Intelligence: Insights from the Social Sciences". In: *Artificial Intelligence* 267 (June 2017).

[MK16] Christopher A Merck and Samantha Kleinberg. "Causal Explanation under Indeterminism: A Sampling Approach". In: *30th Conference on Artificial Intelligence (AAAI)*. AAAI'16. Phoenix, Arizona: AAAI Press, 2016, 1037–1043.

[MNH21] Marco Marabelli, Sue Newell and Val Handunge. "The lifecycle of algorithmic decision-making systems: Organizational choices and ethical challenges". In: *The Journal of Strategic Information Systems* 30 (July 2021). DOI: 10.1016/j.jsis.2021.101683.

[MP95] Zohar Manna and Amir Pnueli. *The Temporal Logic of Reactive and Concurrent Systems: Safety*. Springer-Verlag, 1995. DOI: 10.1007/978-1-4612-0931-7.

[Nam01] Kedar S. Namjoshi. "Certifying Model Checkers". In: *13th International Conference on Computer Aided Verification (CAV)*. Vol. 2102. Lecture Notes in Computer Science. Springer, 2001, pp. 2–13. URL: https://doi.org/10.1007/3-540-44585-4_2.

[New22] BBC News. *'Self-driving' cars to be allowed on UK roads this year*. accessed on 06.02.2024. 2022. URL: https://www.bbc.com/news/technology-56906145.

[PB99] Stephen D Patek and Dimitri P Bertsekas. "Stochastic shortest path games". In: *SIAM Journal on Control and Optimization* 37.3 (1999), pp. 804–824.

[Pea09] Judea Pearl. *Causality*. 2nd. Cambridge University Press, 2009.

[PH13] L. A. Paul and Ned Hall. *Causation: A User?s Guide*. Oxford: Oxford University Press UK, 2013.

[Pow11] David Powers. "Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation". In: *Journal of Machine Learning Technologies* 2.1 (2011), pp. 37–63. ISSN: 2229-3981.

[PPB23] Julie Parreaux, Jakob Piribauer and Christel Baier. "Counterfactual Causality for Reachability and Safety based on Distance Functions". In: *Proceedings of the Fourteenth International Symposium on Games, Automata, Logics, and Formal Verification, Udine, Italy, 18-20th September 2023*. Ed. by Antonis Achilleos and Dario Della Monica. Vol. 390. Electronic Proceedings in Theoretical Computer Science. Open Publishing Association, 2023, pp. 132–149. DOI: 10.4204/EPTCS.390.9.

[Put94] Martin Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York, NY: John Wiley & Sons, Inc., 1994.

[PV91] Panos M Pardalos and Stephen A Vavasis. "Quadratic programming with one negative eigenvalue is NP-hard". In: *Journal of Global optimization* 1.1 (1991), pp. 15–22.

[Rei56] Hans Reichenbach. *The Direction of Time*. Dover Publications, 1956.

[Sco23] Richard Mackenzie-Gray Scott. *Automated Decision-Making and the Challenge of Implementing Existing Laws*. en. 2023. DOI: 10.17176/20231005-233624-0.

[Sel94] Alan L. Selman. "A taxonomy of complexity classes of functions". In: *Journal of Computer and System Sciences* 48.2 (1994), pp. 357–381. ISSN: 0022-0000. DOI: https://doi.org/10.1016/S0022-0000(05)80009-1.

[Sim51] E. H. Simpson. "The Interpretation of Interaction in Contingency Tables". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 13.2 (1951), pp. 238–241. ISSN: 00359246. URL: http://www.jstor.org/stable/2984065 (visited on 06/02/2024).

[Sip13] Michael Sipser. *Introduction to the Theory of Computation*. Third. Boston, MA: Course Technology, 2013. ISBN: 113318779X.

[Sky81] Brian Skyrms. "Causal Necessity". In: *Philosophy of Science* 48.2 (1981), pp. 329–335.

[Sup70] Patrick Suppes. *A Probabilistic Theory of Causality*. Amsterdam: North-Holland Pub. Co., 1970.

[Tho90] Wolfgang Thomas. "Automata on Infinite Objects". In: *Formal Models and Semantics*. Ed. by JAN VAN LEEUWEN. Handbook of Theoretical Computer Science. Amsterdam: Elsevier, 1990, pp. 133–191. ISBN: 978-0-444-88074-1. DOI: https://doi.org/10.1016/B978-0-444-88074-1.50009-3. URL: https://www.sciencedirect.com/science/article/pii/B9780444880741500093.

[Vav90] Stephen A. Vavasis. "Quadratic programming is in NP". In: *Information Processing Letters* 36.2 (1990), pp. 73–77. ISSN: 0020-0190.

[Yoo+21] Woongjae Yoo, Jacob K. Zieba, Nora J. Foegeding, Teresa P. Torres, Catherine D. Shelton, Nicolas G. Shealy, Austin J. Byndloss, Stephanie A. Cevallos, Erik Gertz, Connor R. Tiffany, Julia D. Thomas, Yael Litvak, Henry Nguyen, Erin E. Olsan, Brian J. Bennett, Jeffrey C. Rathmell, Amy S. Major, Andreas J. Bäumler and Mariana X. Byndloss. "High-fat diet induced colonocyte dysfunction escalates microbiota-derived trimethylamine N-oxide". In: *Science* 373.6556 (2021), pp. 813–818. DOI: 10.1126/science.aba3683.

[ZCK18] Min Zheng, Jan Claassen and Samantha Kleinberg. "Automated Identification of Causal Moderators in Time-Series Data". In: *Proceedings of 2018 ACM SIGKDD Workshop on Causal Disocvery*. Ed. by Thuc Duy Le, Kun Zhang, Emre Kıcıman, Aapo Hyvärinen and Lin Liu. Vol. 92. Proceedings of Machine Learning Research. PMLR, 2018, pp. 4–22. URL: https://proceedings.mlr.press/v92/zheng18a.html.

[Zhe+20] Min Zheng, Jessecae Marsh, Jeffrey Nickerson and Samantha Kleinberg. “How causal information affects decisions”. In: *Cognitive Research: Principles and Implications* 5 (Dec. 2020). DOI: 10.1186/s41235-020-0206-z.

[Zie+22] Robin Ziemek, Jakob Piribauer, Florian Funke, Simon Jantsch and Christel Baier. “Probabilistic causes in Markov chains”. In: *Innovations in Systems and Software Engineering* (2022). ISSN: 1614-5054.

[ZK17] Min Zheng and Samantha Kleinberg. “A Method for Automating Token Causal Explanation and Discovery”. In: *Proceedings of FLAIRS’17*. 2017.