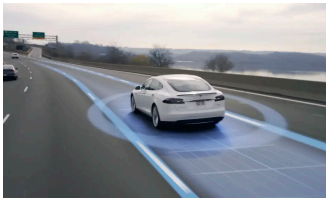
Probabilistic Causality in Markovian Models

Robin Ziemek

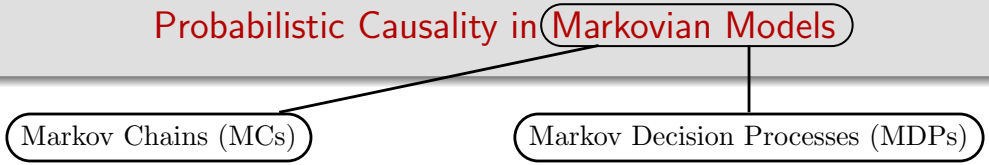
Faculty of Computer Science
Technische Universität Dresden

Defense, 9th August 2024



Probabilistic Causality in Markovian Models

Probabilistic Causality in Markovian Models


Probabilistic Causality in Markovian Models
Markov Chains (MCs)
Markov Decision Processes (MDPs)

- Standard probabilistic models

- Standard probabilistic models
- Represent e.g. probabilistic systems (possibly with non-determinism)

- Standard probabilistic models
- Represent e.g. probabilistic systems (possibly with non-determinism)
- All sorts of properties can be verified by probabilistic model checking

- Standard probabilistic models
- Represent e.g. probabilistic systems (possibly with non-determinism)
- All sorts of properties can be verified by probabilistic model checking

Model Checking

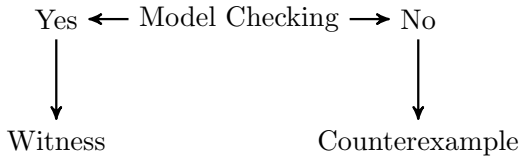
Probabilistic Causality in Markovian Models

- Standard probabilistic models
- Represent e.g. probabilistic systems (possibly with non-determinism)
- All sorts of properties can be verified by probabilistic model checking

Yes ⟵ Model Checking ⟶ No

Probabilistic Causality in Markovian Models

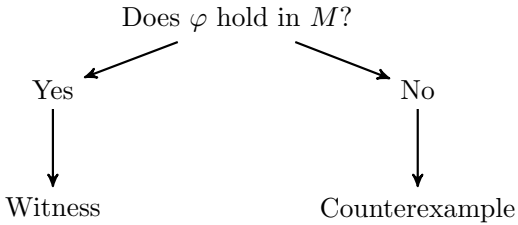
- Standard probabilistic models
- Represent e.g. probabilistic systems (possibly with non-determinism)
- All sorts of properties can be verified by probabilistic model checking



Towards Explanations

Model Checking

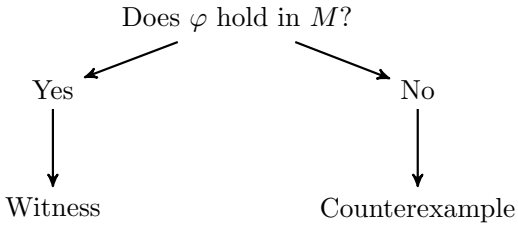
- Given system model M and property specification φ



Towards Explanations

Model Checking

- Given system model M and property specification φ

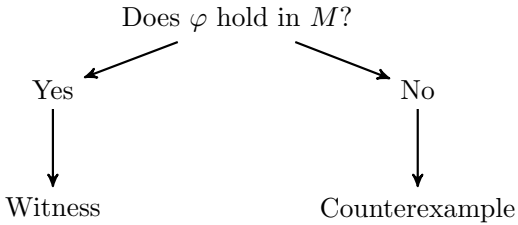


Explanation

Towards Explanations

Model Checking

- Given system model M and property specification φ



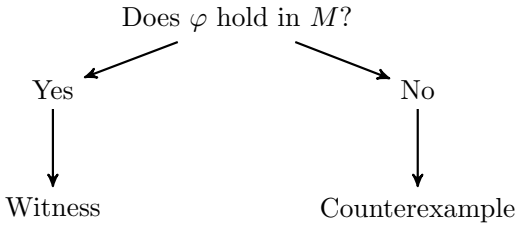
Explanation

- Understanding the system

Towards Explanations

Model Checking

- Given system model M and property specification φ



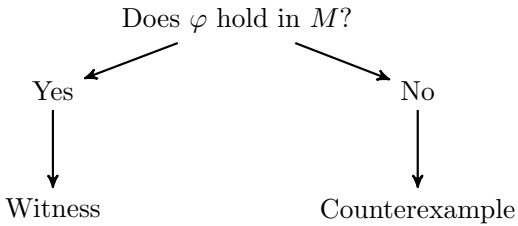
Explanation

- Understanding the system
- Compact and comprehensive

Towards Explanations

Model Checking

- Given system model M and property specification φ



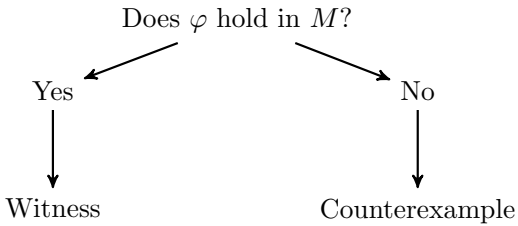
Explanation

- Understanding the system
- Compact and comprehensive
- Inherently difficult to find especially for larger systems

Towards Explanations

Model Checking

- Given system model M and property specification φ



Explanation

- Understanding the system
- Compact and comprehensive
- Inherently difficult to find especially for larger systems

A central concept for understanding is

Causality

Causality

- Causality has been a philosophical topic for centuries

Causality

- Causality has been a philosophical topic for centuries
- Over these years central features of cause-effect relations have been agreed on

Causality

- Causality has been a philosophical topic for centuries
- Over these years central features of cause-effect relations have been agreed on
  - Counterfactual reasoning [1]
    "If the cause would not have happened, the effect would not have happened"
  - Temporal priority

[1] e.g. David Hume. *A Treatise of Human Nature*. John Noon, 1739

Causality

- Causality has been a philosophical topic for centuries
- Over these years central features of cause-effect relations have been agreed on
  - Probability-raising causality [2]
    "*The probability of the effect is higher after the cause happened*"

$$\mathbf{P}(\text{Eff} \mid \text{Cause}) > \mathbf{P}(\text{Eff})$$

  - Temporal priority

[2] e.g. Hans Reichenbach. *The Direction of Time*. Dover Publications, 1956

Causality in Computer Science

- Computer science has taken an interest in Causality

Causality in Computer Science

- Computer science has taken an interest in Causality
- Outstanding work by Halpern & Pearl on counterfactual causality[3,4]

[3] Joseph Y. Halpern and Judea Pearl. "Causes and Explanations: A Structural-Model Approach: Part 1: Causes". In: *Conference in Uncertainty in Artificial Intelligence (UAI)*. 2001, pp. 194–202

[4] Joseph Y. Halpern. "A Modification of the Halpern-Pearl Definition of Causality". In: *Proceedings of IJCAI'15*. AAAI Press, 2015, 3022–3033

Causality in Computer Science

- Computer science has taken an interest in Causality
- Outstanding work by Halpern & Pearl on counterfactual causality[3,4]
- Refinement of probability-raising by Pearl[5] for intelligent systems

[5] Judea Pearl. *Causality*. 2nd. Cambridge University Press, 2009

Causality in Computer Science

- Computer science has taken an interest in Causality
- Outstanding work by Halpern & Pearl on counterfactual causality[3,4]
- Refinement of probability-raising by Pearl[5] for intelligent systems
- We focus on Formal Methods

[5] Judea Pearl. *Causality*. 2nd. Cambridge University Press, 2009

Causality in Formal Methods

- Inspired by Halpern & Pearl and mainly focused on counterfactual

Causality in Formal Methods

- Inspired by Halpern & Pearl and mainly focused on counterfactual
  - Explaining counterexamples by fault localization[6]

[6] Ilan Beer et al. "Explaining counterexamples using causality". In: *Formal Methods in System Design* 40.1 (2012), pp. 20–40

Causality in Formal Methods

- Inspired by Halpern & Pearl and mainly focused on counterfactual
  - Explaining counterexamples by fault localization[6]
  - Relation between counterfactuals and coverage[7]

[7] 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

Causality in Formal Methods

- Inspired by Halpern & Pearl and mainly focused on counterfactual
  - Explaining counterexamples by fault localization[6]
  - Relation between counterfactuals and coverage[7]
  - Causal reasoning in configurable systems[8]

[8] Amjad Ibrahim and Alexander Pretschner. "From Checking to Inference: Actual Causality Computations as Optimization Problems". In: *Proceedings of ATVA'20*. 2020, pp. 343–359

Causality in Formal Methods

- Inspired by Halpern & Pearl and mainly focused on counterfactual
  - Explaining counterexamples by fault localization[6]
  - Relation between counterfactuals and coverage[7]
  - Causal reasoning in configurable systems[8]
- Many more approaches[9]

[9] Christel Baier et al. "From Verification to Causality-Based Explications (Invited Talk)". In: *48th International Colloquium on Automata, Languages, and Programming, (ICALP)*. Vol. 198. LIPIcs. 2021, 1:1–1:20

Causality in Formal Methods

- Inspired by Halpern & Pearl and mainly focused on counterfactual
  - Explaining counterexamples by fault localization[6]
  - Relation between counterfactuals and coverage[7]
  - Causal reasoning in configurable systems[8]
- Many more approaches[9]
- Studies mostly for deterministic models

[9] Christel Baier et al. "From Verification to Causality-Based Explications (Invited Talk)". In: *48th International Colloquium on Automata, Languages, and Programming, (ICALP)*. Vol. 198. LIPIcs. 2021, 1:1–1:20

Causality in Probabilistic Models

- Studies of probabilistic causality in probabilistic models are very sparse

Causality in Probabilistic Models

- Studies of probabilistic causality in probabilistic models are very sparse
- Most approaches concentrate on Markov chains[10,11]

[10] Samantha Kleinberg and Bud Mishra. "The Temporal Logic of Causal Structures". In: *25th Conference on Uncertainty in Artificial Intelligence (UAI)*. 2009, pp. 303–312

[11] 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

Causality in Probabilistic Models

- Studies of probabilistic causality in probabilistic models are very sparse
- Most approaches concentrate on Markov chains[10,11]
  - No thorough studies

[10] Samantha Kleinberg and Bud Mishra. "The Temporal Logic of Causal Structures". In: *25th Conference on Uncertainty in Artificial Intelligence (UAI)*. 2009, pp. 303–312

[11] 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

Causality in Probabilistic Models

- Studies of probabilistic causality in probabilistic models are very sparse
- Most approaches concentrate on Markov chains[10,11]
  - No thorough studies
- No studies for probability-raising causality in Markov decision processes (MDPs)

[10] Samantha Kleinberg and Bud Mishra. "The Temporal Logic of Causal Structures". In: *25th Conference on Uncertainty in Artificial Intelligence (UAI)*. 2009, pp. 303–312

[11] 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

Probabilistic Causality in Markovian Models

- My dissertation is concerned with probabilistic causality in MDPs

Probabilistic Causality in Markovian Models

- My dissertation is concerned with probabilistic causality in MDPs
- Foundational research on formalization of probability-raising causality

Probabilistic Causality in Markovian Models

- My dissertation is concerned with probabilistic causality in MDPs
- Foundational research on formalization of probability-raising causality
- Emphasis on algorithmic studies

Probabilistic Causality in Markovian Models

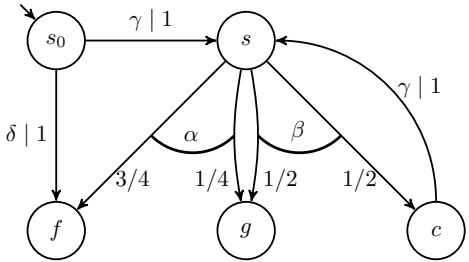
- My dissertation is concerned with probabilistic causality in MDPs
- Foundational research on formalization of probability-raising causality
- Emphasis on algorithmic studies
  - Check cause-effect relations

Probabilistic Causality in Markovian Models

- My dissertation is concerned with probabilistic causality in MDPs
- Foundational research on formalization of probability-raising causality
- Emphasis on algorithmic studies
  - Check cause-effect relations
  - Compute causes for given effect

Markov Decision Process

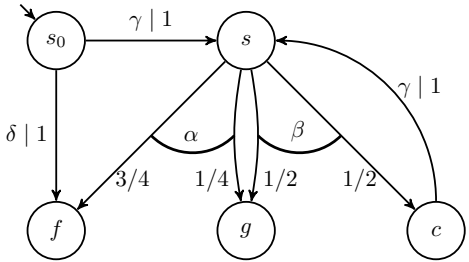
A *Markov decision process* (MDP) $\mathcal{M} = (S, Act, \mathbf{P}, s_0)$ consists of



Markov Decision Process

A *Markov decision process* (MDP) $\mathcal{M} = (S, Act, \mathbf{P}, s_0)$ consists of

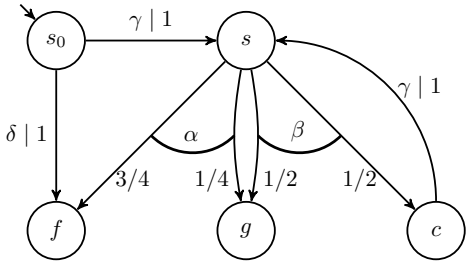
- a finite set of states S,



Markov Decision Process

A *Markov decision process* (MDP) $\mathcal{M} = (S, Act, \mathbf{P}, s_0)$ consists of

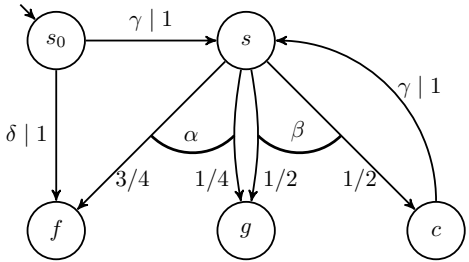
- a finite set of states S,
- a finite set of actions Act,



Markov Decision Process

A *Markov decision process* (MDP) $\mathcal{M} = (S, Act, \mathbf{P}, s_0)$ consists of

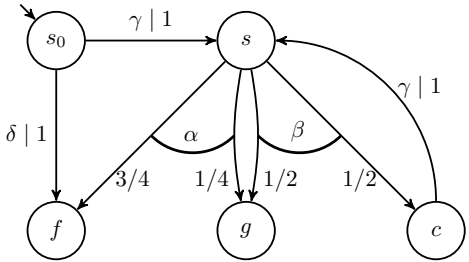
- a finite set of states S,
- a finite set of actions Act,
- an initial state s_0,



Markov Decision Process

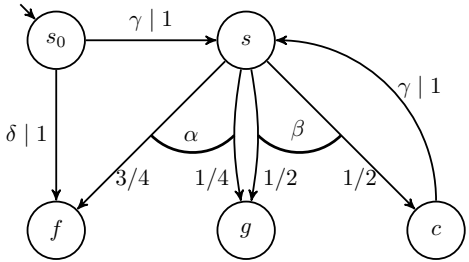
A *Markov decision process* (MDP) $\mathcal{M} = (S, Act, \mathbf{P}, s_0)$ consists of

- a finite set of states S,
- a finite set of actions Act,
- an initial state s_0,
- and a probabilistic transition function $\mathbf{P}\colon S \times Act \times S \to [0,1]$.



Markov Decision Process – Scheduler

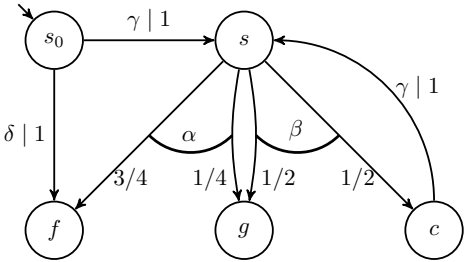
A *scheduler* of an MDP $\mathcal{M}$ is a function $\mathfrak{S}$ with



Markov Decision Process – Scheduler

A *scheduler* of an MDP $\mathcal{M}$ is a function $\mathfrak{S}$ with

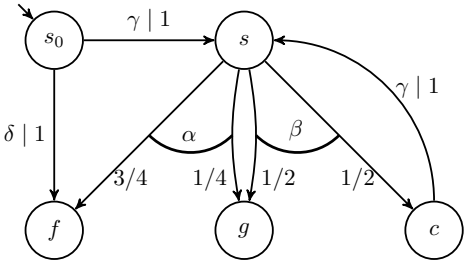
- Input: finite path in $\mathcal{M}$



Markov Decision Process – Scheduler

A *scheduler* of an MDP $\mathcal{M}$ is a function $\mathfrak{S}$ with

- Input: finite path in $\mathcal{M}$
- Output: distribution over enabled actions

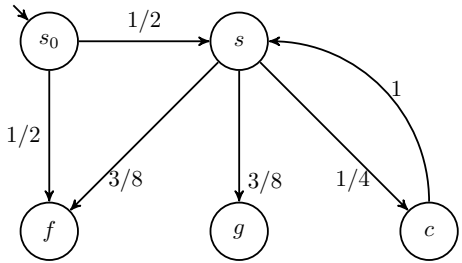


Markov Decision Process – Scheduler

A *scheduler* of an MDP $\mathcal{M}$ is a function $\mathfrak{S}$ with

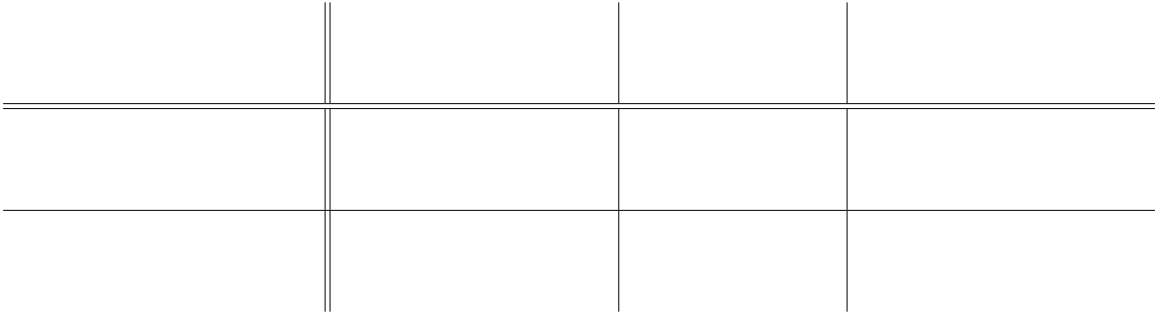
- Input: finite path in $\mathcal{M}$
- Output: distribution over enabled actions

The purely probabilistic model is a *discrete time Markov chain* (DTMC)



Probability-Raising Causality in MDPs

Setting for probability-raising conditions depends on definition of cause C and effect E



Probability-Raising Causality in MDPs

Setting for probability-raising conditions depends on definition of cause C and effect E

	state-based C and E sets of states		

Probability-Raising Causality in MDPs

Setting for probability-raising conditions depends on definition of cause C and effect E

	state-based C and E sets of states		*path-based* C and E path properties

Probability-Raising Causality in MDPs

Setting for probability-raising conditions depends on definition of cause C and effect E

	state-based C and E sets of states	*mixed* C set of states E path property	*path-based* C and E path properties

Probability-Raising Causality in MDPs

Setting for probability-raising conditions depends on definition of cause C and effect E

	state-based C and E sets of states	*mixed* C set of states E path property	*path-based* C and E path properties
global PR C viewed as unit			

Probability-Raising Causality in MDPs

Setting for probability-raising conditions depends on definition of cause C and effect E

	state-based C and E sets of states	*mixed* C set of states E path property	*path-based* C and E path properties
global PR C viewed as unit			
strict PR elements in C individually			

Probability-Raising Causality in MDPs

Setting for probability-raising conditions depends on definition of cause C and effect E

	state-based C and E sets of states	*mixed* C set of states E path property	*path-based* C and E path properties
global PR C viewed as unit	GPR cause		
strict PR elements in C individually			

Probability-Raising Causality in MDPs

Setting for probability-raising conditions depends on definition of cause C and effect E

	state-based C and E sets of states	*mixed* C set of states E path property	*path-based* C and E path properties
global PR C viewed as unit	GPR cause		
strict PR elements in C individually	SPR cause		

Probability-Raising Causality in MDPs

Setting for probability-raising conditions depends on definition of cause C and effect E

	state-based C and E sets of states	*mixed* C set of states E path property	*path-based* C and E path properties
global PR C viewed as unit	GPR cause	reachability GPR cause	
strict PR elements in C individually	SPR cause	reachability SPR cause	

Probability-Raising Causality in MDPs

Setting for probability-raising conditions depends on definition of cause C and effect E

	state-based C and E sets of states	*mixed* C set of states E path property	*path-based* C and E path properties
global PR C viewed as unit	GPR cause	reachability GPR cause	co-safety GPR cause
strict PR elements in C individually	SPR cause	reachability SPR cause	co-safety SPR cause

Probability-Raising Causality in MDPs

Setting for probability-raising conditions depends on definition of cause C and effect E

	state-based C and E sets of states	*mixed* C set of states E path property	*path-based* C and E path properties
global PR C viewed as unit	GPR cause	reachability GPR cause	co-safety GPR cause
strict PR elements in C individually	SPR cause	reachability SPR cause	co-safety SPR cause

Probability-Raising Causality in MDPs

Setting for probability-raising conditions depends on definition of cause C and effect E

	state-based C and E sets of states	*mixed* C set of states E path property	*path-based* C and E path properties
global PR C viewed as unit	GPR cause	reachability GPR cause	co-safety GPR cause
strict PR elements in C individually	SPR cause	reachability SPR cause	co-safety SPR cause

Global Probability-Raising Causes

Given $\mathcal{M}$ and set of terminal effect states E,

Global Probability-Raising Causes

Given $\mathcal{M}$ and set of terminal effect states E, a set of states C is a *GPR cause* iff

Global Probability-Raising Causes

Given $\mathcal{M}$ and set of terminal effect states E, a set of states C is a *GPR cause* iff

(GPR) For each scheduler $\mathfrak{S}$ reaching C:

$$\Pr^{\mathfrak{S}}(\Diamond E \mid \Diamond C) > \Pr^{\mathfrak{S}}(\Diamond E).$$

Global Probability-Raising Causes

Given $\mathcal{M}$ and set of terminal effect states E, a set of states C is a *GPR cause* iff

(GPR) For each scheduler $\mathfrak{S}$ reaching C:

$$\Pr^{\mathfrak{S}}(\ \Diamond E \mid \Diamond C\ ) \ > \ \Pr^{\mathfrak{S}}(\Diamond E).$$

- E being terminal ensures temporal priority of C

Global Probability-Raising Causes

Given $\mathcal{M}$ and set of terminal effect states E, a set of states C is a *GPR cause* iff

(GPR) For each scheduler $\mathfrak{S}$ reaching C:

$$\Pr^{\mathfrak{S}}(\ \Diamond E \mid \Diamond C\ ) \ > \ \Pr^{\mathfrak{S}}(\Diamond E).$$

- E being terminal ensures temporal priority of C
- C is treated as a unit by (GPR)

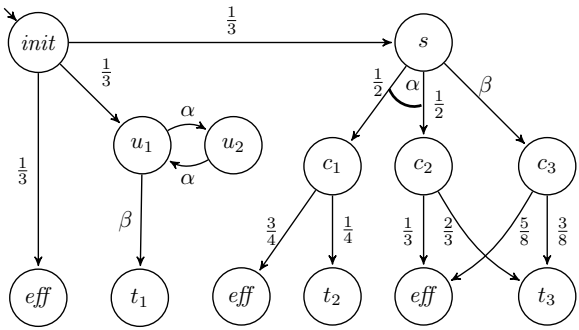
Global Probability-Raising Causes

Given $\mathcal{M}$ and set of terminal effect states E, a set of states C is a *GPR cause* iff

(GPR) For each scheduler $\mathfrak{S}$ reaching C:

$$\Pr^{\mathfrak{S}}(\ \Diamond E \mid \Diamond C\ ) \ > \ \Pr^{\mathfrak{S}}(\Diamond E).$$

- E being terminal ensures temporal priority of C
- C is treated as a unit by (GPR)



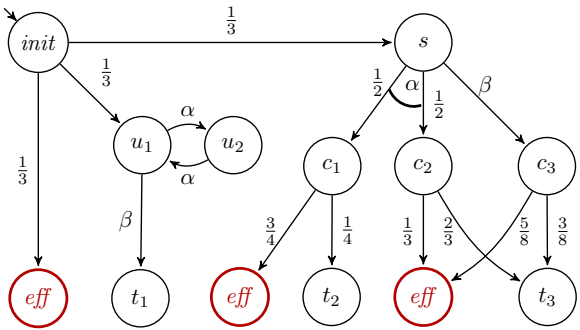
Global Probability-Raising Causes

Given $\mathcal{M}$ and set of terminal effect states E, a set of states C is a *GPR cause* iff

(GPR) For each scheduler $\mathfrak{S}$ reaching C:

$$\Pr^{\mathfrak{S}}(\ \Diamond E \mid \Diamond C\ ) \ > \ \Pr^{\mathfrak{S}}(\Diamond E).$$

- E being terminal ensures temporal priority of C
- C is treated as a unit by (GPR)



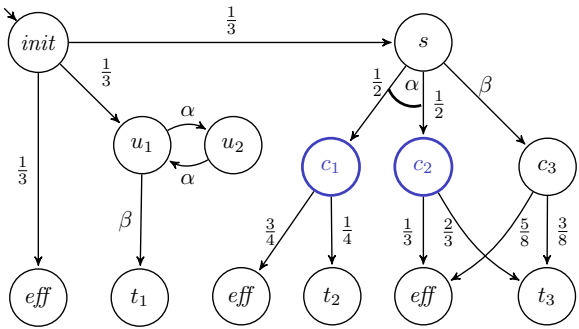
Global Probability-Raising Causes

Given $\mathcal{M}$ and set of terminal effect states E, a set of states C is a *GPR cause* iff

(GPR) For each scheduler $\mathfrak{S}$ reaching C:

$$\Pr^{\mathfrak{S}}(\ \Diamond E \mid \Diamond C\ ) \ > \ \Pr^{\mathfrak{S}}(\Diamond E).$$

- E being terminal ensures temporal priority of C
- C is treated as a unit by (GPR)
- In the example $C = \{c_1, c_2\}$ is GPR cause



Probability-Raising Causality in MDPs

Setting for probability-raising conditions depends on definition of cause C and effect E

	state-based C and E sets of states	*mixed* C set of states E path property	*path-based* C and E path properties
global PR C viewed as unit	GPR cause	reachability GPR cause	co-safety GPR cause
strict PR elements in C individually	SPR cause	reachability SPR cause	co-safety SPR cause

Strict Probability-Raising Causes

Given $\mathcal{M}$ and set of terminal effect states E, a set of states C is a *SPR cause* iff

Strict Probability-Raising Causes

Given $\mathcal{M}$ and set of terminal effect states E, a set of states C is a *SPR cause* iff

(SPR) For each $c \in C$
and scheduler $\mathfrak{S}$ reaching c:

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

Strict Probability-Raising Causes

Given $\mathcal{M}$ and set of terminal effect states E, a set of states C is a *SPR cause* iff

(SPR) For each $c \in C$
and scheduler $\mathfrak{S}$ reaching c:

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

- PR for each individual element

Strict Probability-Raising Causes

Given $\mathcal{M}$ and set of terminal effect states E, a set of states C is a *SPR cause* iff

(SPR) For each $c \in C$
and scheduler $\mathfrak{S}$ reaching c:

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

- PR for each individual element
- Each $c \in C$ is treated locally by (SPR)

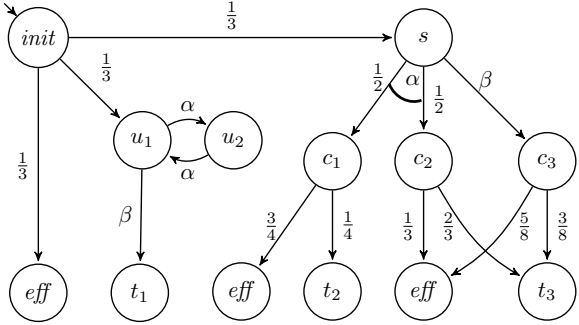
Strict Probability-Raising Causes

Given $\mathcal{M}$ and set of terminal effect states E, a set of states C is a *SPR cause* iff

(SPR) For each $c \in C$
and scheduler $\mathfrak{S}$ reaching c:

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

- PR for each individual element
- Each $c \in C$ is treated locally by (SPR)



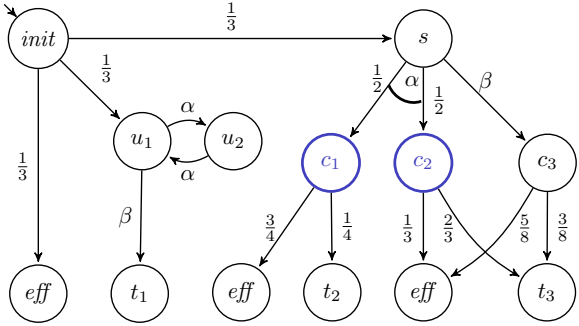
Strict Probability-Raising Causes

Given $\mathcal{M}$ and set of terminal effect states E, a set of states C is a *SPR cause* iff

(SPR) For each $c \in C$
and scheduler $\mathfrak{S}$ reaching c:

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

- PR for each individual element
- Each $c \in C$ is treated locally by (SPR)
- In example $C = \{c_1, c_2\}$ is no SPR cause



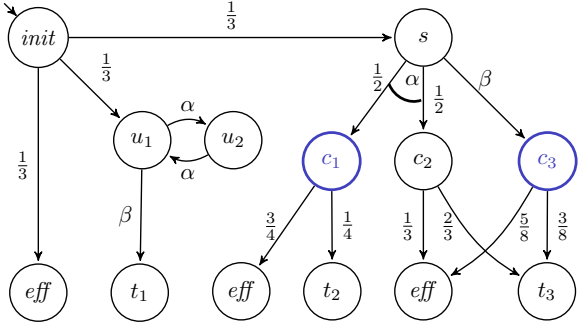
Strict Probability-Raising Causes

Given $\mathcal{M}$ and set of terminal effect states E, a set of states C is a *SPR cause* iff

(SPR) For each $c \in C$
and scheduler $\mathfrak{S}$ reaching c:

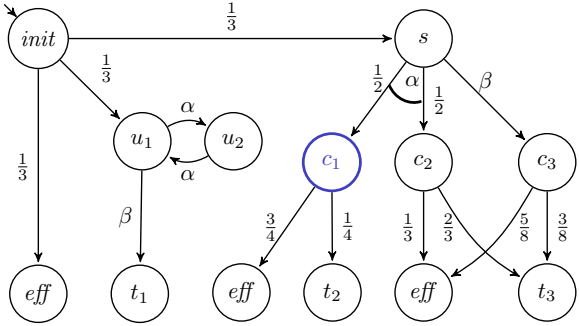
$$\Pr^{\mathfrak{S}}(\ \Diamond E \mid (\neg C)\,\mathrm{U}\,c\ ) \ > \ \Pr^{\mathfrak{S}}(\Diamond E).$$

- PR for each individual element
- Each $c \in C$ is treated locally by (SPR)
- In example $C = \{c_1, c_2\}$ is no SPR cause
- $C = \{c_1, c_3\}$ is an SPR cause



Singleton PR Causes

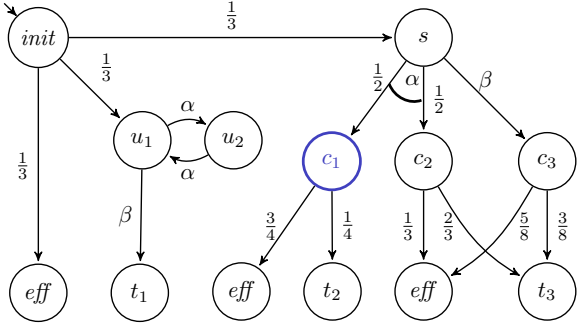
For the example MDP consider the effect set as labeled and $C = \{c_1\}$.



Singleton PR Causes

For the example MDP consider the effect set as labeled and $C = \{c_1\}$.

(GPR) For each sensible $\mathfrak{S}$: $\Pr^{\mathfrak{S}}(\Diamond E | \Diamond C) > \Pr^{\mathfrak{S}}(\Diamond E)$.
(SPR) For each $c \in C$ and sensible $\mathfrak{S}$:
$\Pr^{\mathfrak{S}}(\Diamond E | (\neg C)\,\mathrm{U}\,c) > \Pr^{\mathfrak{S}}(\Diamond E)$.

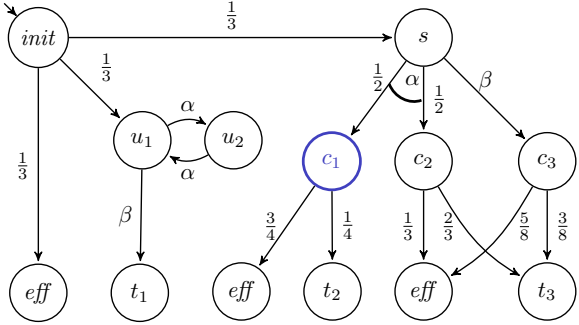


Singleton PR Causes

For the example MDP consider the effect set as labeled and $C = \{c_1\}$.

(GPR) For each sensible $\mathfrak{S}$: $\Pr^{\mathfrak{S}}(\Diamond E | \Diamond C) > \Pr^{\mathfrak{S}}(\Diamond E)$.
(SPR) For each $c \in C$ and sensible $\mathfrak{S}$:
$\Pr^{\mathfrak{S}}(\Diamond E | (\neg C)\,\mathrm{U}\,c) > \Pr^{\mathfrak{S}}(\Diamond E)$.

- The scheduler taking β in s is not considered in (GPR) or (SPR)

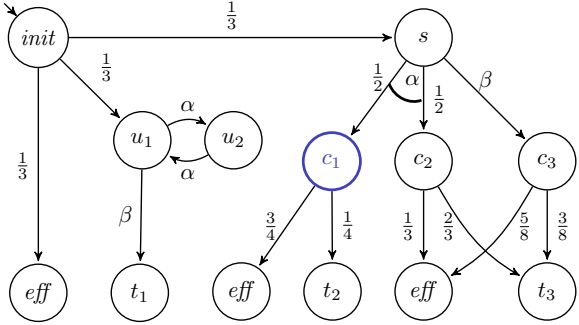


Singleton PR Causes

For the example MDP consider the effect set as labeled and $C = \{c_1\}$.

(GPR) For each sensible $\mathfrak{S}$: $\Pr^{\mathfrak{S}}(\Diamond E | \Diamond C) > \Pr^{\mathfrak{S}}(\Diamond E)$.
(SPR) For each $c \in C$ and sensible $\mathfrak{S}$:
$\Pr^{\mathfrak{S}}(\Diamond E | (\neg C)\,\mathrm{U}\,c) > \Pr^{\mathfrak{S}}(\Diamond E)$.

- The scheduler taking β in s is not considered in (GPR) or (SPR)
- $\Pr^{\min}(\Diamond E \mid \Diamond C) = \frac{3}{4} > \frac{13}{24} = \Pr^{\max}(\Diamond E)$



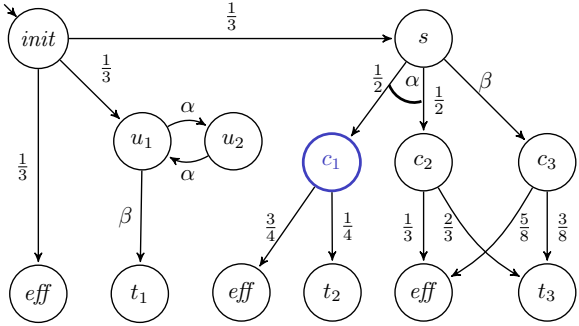
Singleton PR Causes

For the example MDP consider the effect set as labeled and $C = \{c_1\}$.

(GPR) For each sensible $\mathfrak{S}$: $\Pr^{\mathfrak{S}}(\Diamond E | \Diamond C) > \Pr^{\mathfrak{S}}(\Diamond E)$.
(SPR) For each $c \in C$ and sensible $\mathfrak{S}$:
$\Pr^{\mathfrak{S}}(\Diamond E | (\neg C)\,\mathrm{U}\,c) > \Pr^{\mathfrak{S}}(\Diamond E)$.

- The scheduler taking β in s is not considered in (GPR) or (SPR)
- $\Pr^{\min}(\Diamond E \mid \Diamond C) = \frac{3}{4} > \frac{13}{24} = \Pr^{\max}(\Diamond E)$

$\Longrightarrow$ $C = \{c_1\}$ is both SPR and GPR cause



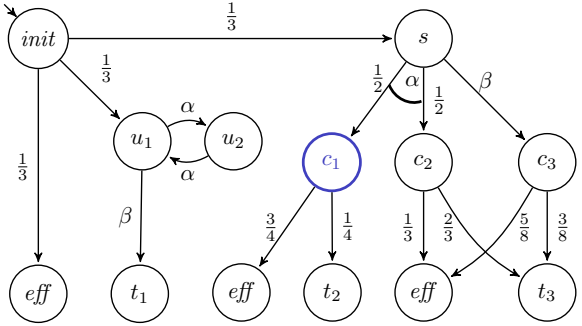
Singleton PR Causes

For the example MDP consider the effect set as labeled and $C = \{c_1\}$.

(GPR) For each sensible $\mathfrak{S}$: $\Pr^{\mathfrak{S}}(\Diamond E | \Diamond C) > \Pr^{\mathfrak{S}}(\Diamond E)$.
(SPR) For each $c \in C$ and sensible $\mathfrak{S}$:
$\Pr^{\mathfrak{S}}(\Diamond E | (\neg C)\,\mathrm{U}\,c) > \Pr^{\mathfrak{S}}(\Diamond E)$.

- The scheduler taking β in s is not considered in (GPR) or (SPR)
- $\Pr^{\min}(\Diamond E \mid \Diamond C) = \frac{3}{4} > \frac{13}{24} = \Pr^{\max}(\Diamond E)$

$\Longrightarrow$ $C = \{c_1\}$ is both SPR and GPR cause

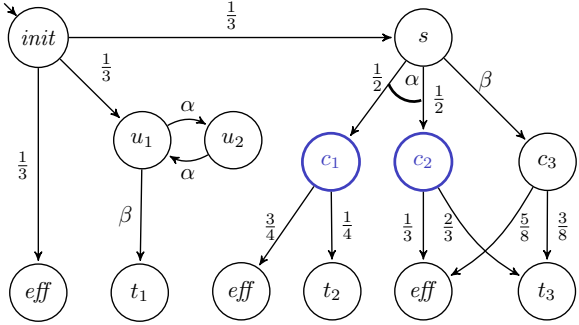


[Prop] For singletons (GPR) and (SPR) are equivalent.

SPR vs. GPR Causality

For the example MDP consider the effect set as labeled and $C = \{c_1, c_2\}$.

(GPR) For each sensible $\mathfrak{S}$: $\Pr^{\mathfrak{S}}(\Diamond E | \Diamond C) > \Pr^{\mathfrak{S}}(\Diamond E)$.
(SPR) For each $c \in C$ and sensible $\mathfrak{S}$:
$\Pr^{\mathfrak{S}}(\Diamond E | (\neg C)\,\mathrm{U}\,c) > \Pr^{\mathfrak{S}}(\Diamond E)$.

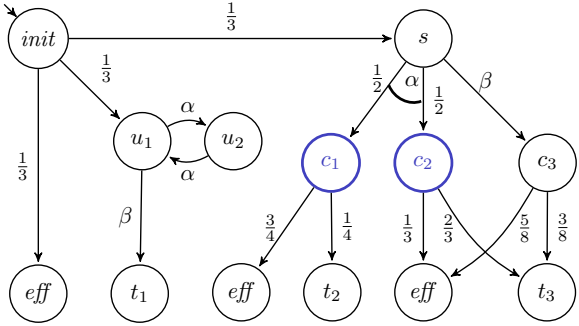


SPR vs. GPR Causality

For the example MDP consider the effect set as labeled and $C = \{c_1, c_2\}$.

(GPR) For each sensible $\mathfrak{S}$: $\Pr^{\mathfrak{S}}(\Diamond E | \Diamond C) > \Pr^{\mathfrak{S}}(\Diamond E)$.
(SPR) For each $c \in C$ and sensible $\mathfrak{S}$:
$\Pr^{\mathfrak{S}}(\Diamond E | (\neg C)\,\mathrm{U}\,c) > \Pr^{\mathfrak{S}}(\Diamond E)$.

- $C = \{c_1, c_2\}$ is a GPR cause

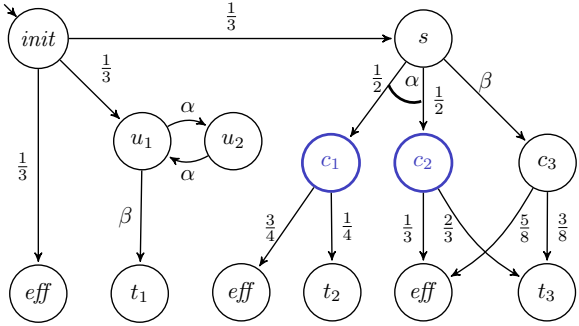


SPR vs. GPR Causality

For the example MDP consider the effect set as labeled and $C = \{c_1, c_2\}$.

(GPR) For each sensible $\mathfrak{S}$: $\Pr^{\mathfrak{S}}(\Diamond E | \Diamond C) > \Pr^{\mathfrak{S}}(\Diamond E)$.
(SPR) For each $c \in C$ and sensible $\mathfrak{S}$:
$\Pr^{\mathfrak{S}}(\Diamond E | (\neg C)\,\mathrm{U}\,c) > \Pr^{\mathfrak{S}}(\Diamond E)$.

- $C = \{c_1, c_2\}$ is a GPR cause
  but c_2 does not satisfy (SPR)

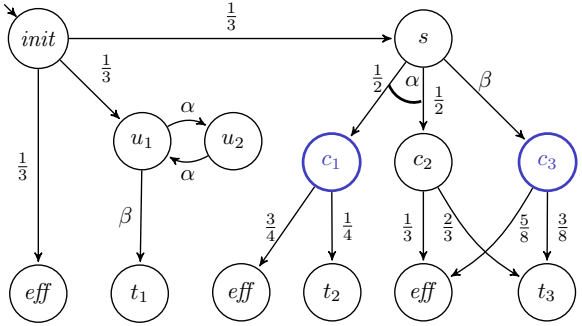


SPR vs. GPR Causality

For the example MDP consider the effect set as labeled and $C = \{c_1, c_3\}$.

(GPR) For each sensible $\mathfrak{S}$: $\Pr^{\mathfrak{S}}(\Diamond E | \Diamond C) > \Pr^{\mathfrak{S}}(\Diamond E)$.
(SPR) For each $c \in C$ and sensible $\mathfrak{S}$:
$\Pr^{\mathfrak{S}}(\Diamond E | (\neg C) \,\mathrm{U}\, c) > \Pr^{\mathfrak{S}}(\Diamond E)$.

- $C = \{c_1, c_2\}$ is a GPR cause
  but c_2 does not satisfy (SPR)
- $C = \{c_1, c_3\}$ is SPR and GPR cause

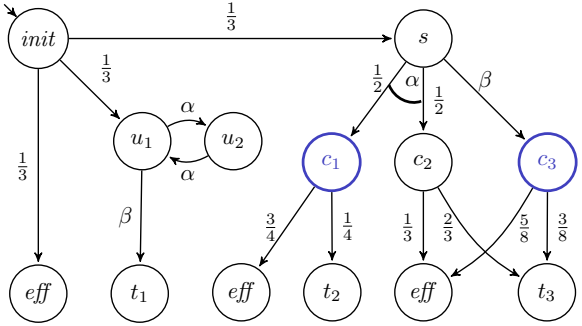


SPR vs. GPR Causality

For the example MDP consider the effect set as labeled and $C = \{c_1, c_3\}$.

(GPR) For each sensible $\mathfrak{S}$: $\Pr^{\mathfrak{S}}(\Diamond E | \Diamond C) > \Pr^{\mathfrak{S}}(\Diamond E)$.
(SPR) For each $c \in C$ and sensible $\mathfrak{S}$:
$\Pr^{\mathfrak{S}}(\Diamond E | (\neg C) \,\mathrm{U}\, c) > \Pr^{\mathfrak{S}}(\Diamond E)$.

- $C = \{c_1, c_2\}$ is a GPR cause
  but c_2 does not satisfy (SPR)
- $C = \{c_1, c_3\}$ is SPR and GPR cause



[Prop] Every SPR cause is also a GPR cause but not vice versa.

Algorithmic Studies

Problems for state-based probability-raising causes

PR causality in MDPs			
GPR cause			
SPR cause			

Algorithmic Studies

Problems for state-based probability-raising causes

PR causality in MDPs	Check condition C and E are given		
GPR cause			
SPR cause			

Algorithmic Studies

Problems for state-based probability-raising causes

PR causality in MDPs	Check condition C and E are given	Compute cause C for given effect E	
GPR cause			
SPR cause			

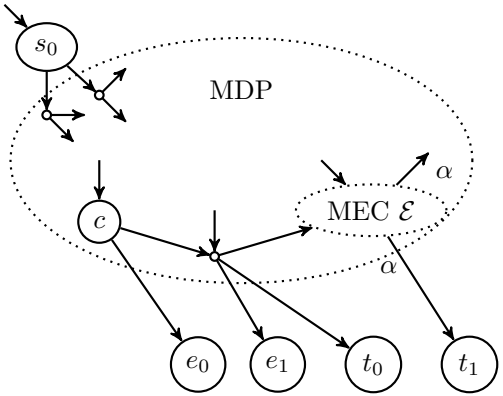
Algorithmic Studies

Problems for state-based probability-raising causes

PR causality in MDPs	Check condition C and E are given	Compute cause C for given effect E	
GPR cause	?		
SPR cause	?		

The Canonical MDP – Cutting at the Cause

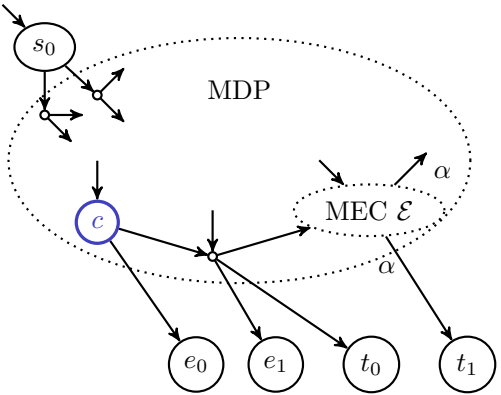
For a given cause candidate C in $\mathcal{M}$ we build the *canonical MDP* $\mathcal{M}_C$



The Canonical MDP – Cutting at the Cause

For a given cause candidate C in $\mathcal{M}$ we build the *canonical MDP* $\mathcal{M}_C$

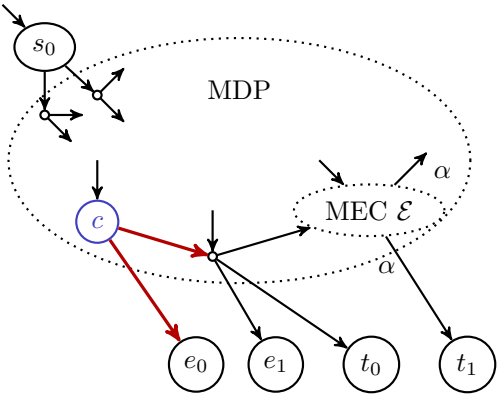
1. $\mathcal{M} \to \mathcal{M}_{<C>}$:



The Canonical MDP – Cutting at the Cause

For a given cause candidate C in $\mathcal{M}$ we build the *canonical MDP* $\mathcal{M}_C$

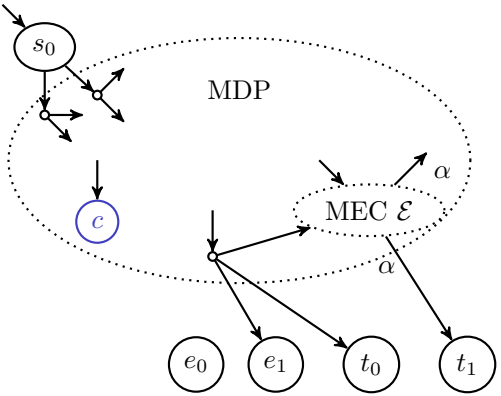
1. $\mathcal{M} \to \mathcal{M}_{<C>}$:
   - Delete enabled actions from $c \in C$



The Canonical MDP – Cutting at the Cause

For a given cause candidate C in $\mathcal{M}$ we build the *canonical MDP* $\mathcal{M}_C$

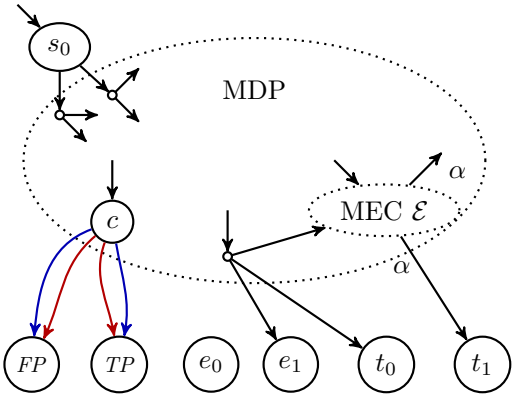
1. $\mathcal{M} \to \mathcal{M}_{<C>}$:
   - Delete enabled actions from $c \in C$
   - Actions α_{min} and α_{max} and states TP and FP are added



The Canonical MDP – Cutting at the Cause

For a given cause candidate C in $\mathcal{M}$ we build the *canonical MDP* $\mathcal{M}_C$

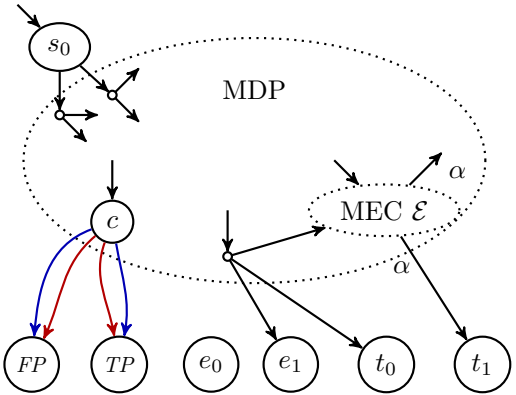
1. $\mathcal{M} \to \mathcal{M}_{<C>}$:
   - Delete enabled actions from $c \in C$
   - Actions α_{min} and α_{max} and states TP and FP are added



The Canonical MDP – Cutting at the Cause

For a given cause candidate C in $\mathcal{M}$ we build the *canonical MDP* $\mathcal{M}_C$

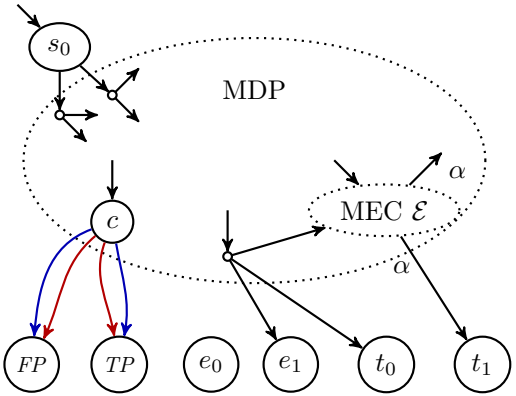
1. $\mathcal{M} \to \mathcal{M}_{<C>}$:
   - Delete enabled actions from $c \in C$
   - Actions $\alpha_{\min}$ and $\alpha_{\max}$ and states TP and FP are added
   - Correspond to minimum/maximum probability to reach E from $c \in C$



The Canonical MDP – MEC-Quotient

For a given cause candidate C in $\mathcal{M}$ we build the *canonical MDP* $\mathcal{M}_C$

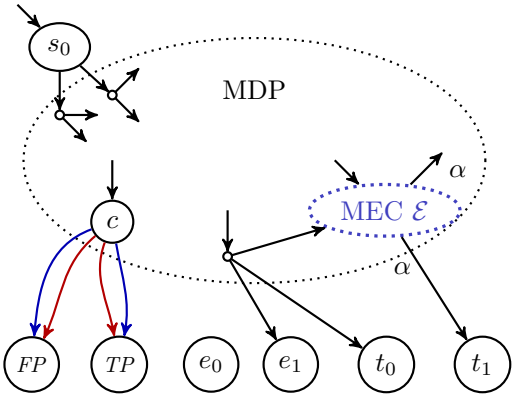
2. $\mathcal{M}_{<C>} \to \mathcal{M}_C$:



The Canonical MDP – MEC-Quotient

For a given cause candidate C in $\mathcal{M}$ we build the *canonical MDP* $\mathcal{M}_C$

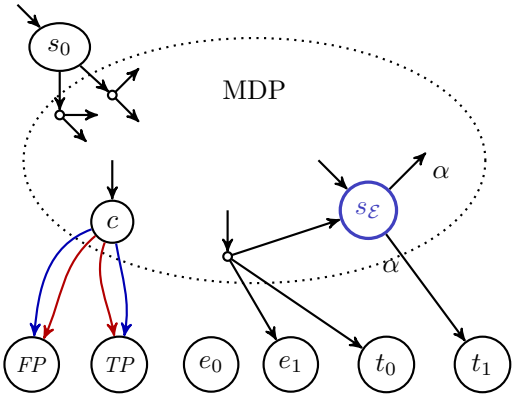
2. $\mathcal{M}_{<C>} \to \mathcal{M}_C$:
   - Collapse MECs $\mathcal{E}$ to single states $s_{\mathcal{E}}$ while keeping outgoing actions of $\mathcal{E}$



The Canonical MDP – MEC-Quotient

For a given cause candidate C in $\mathcal{M}$ we build the *canonical MDP* $\mathcal{M}_C$

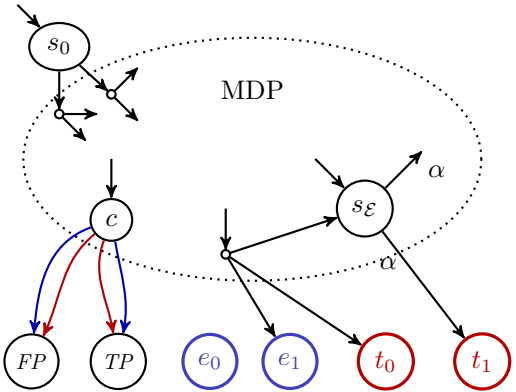
2. $\mathcal{M}_{<C>} \to \mathcal{M}_C$:
   - Collapse MECs $\mathcal{E}$ to single states $s_{\mathcal{E}}$ while keeping outgoing actions of $\mathcal{E}$



The Canonical MDP – MEC-Quotient

For a given cause candidate C in $\mathcal{M}$ we build the *canonical MDP* $\mathcal{M}_C$

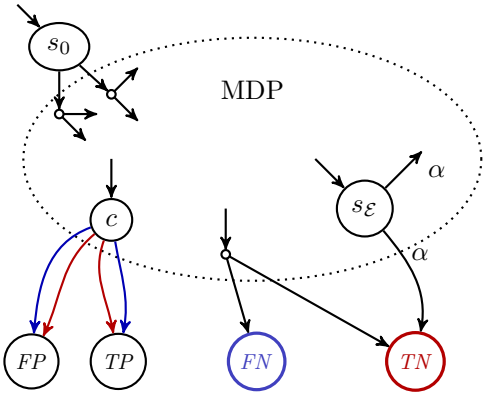
2. $\mathcal{M}_{<C>} \to \mathcal{M}_C$:
   - Collapse MECs $\mathcal{E}$ to single states $s_{\mathcal{E}}$ while keeping outgoing actions of $\mathcal{E}$
   - Collapse E to *FN* and remaining terminal states to *TN*



The Canonical MDP – MEC-Quotient

For a given cause candidate C in $\mathcal{M}$ we build the *canonical MDP* $\mathcal{M}_C$

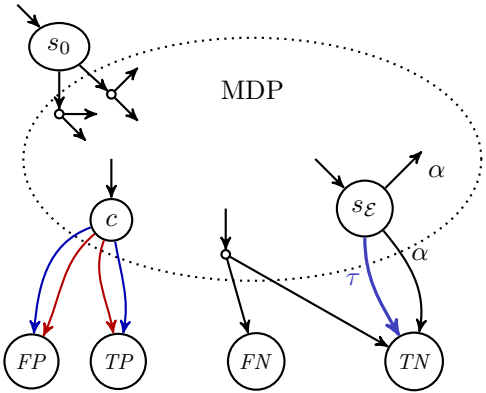
2. $\mathcal{M}_{<C>} \to \mathcal{M}_C$:
   - Collapse MECs $\mathcal{E}$ to single states $s_{\mathcal{E}}$ while keeping outgoing actions of $\mathcal{E}$
   - Collapse E to *FN* and remaining terminal states to *TN*



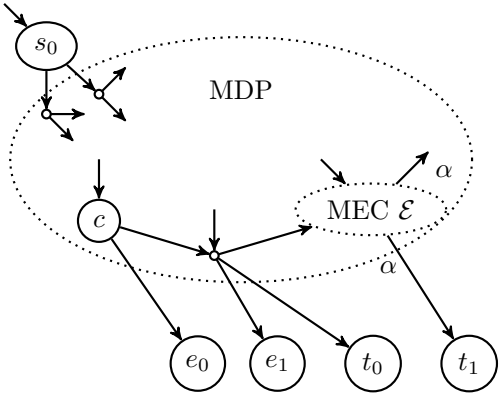
The Canonical MDP – MEC-Quotient

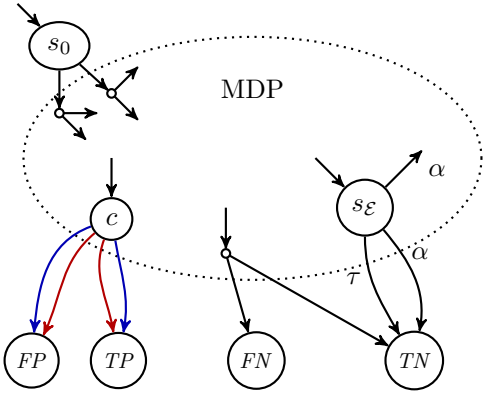
For a given cause candidate C in $\mathcal{M}$ we build the *canonical MDP* $\mathcal{M}_C$

2. $\mathcal{M}_{<C>} \to \mathcal{M}_C$:
   - Collapse MECs $\mathcal{E}$ to single states $s_{\mathcal{E}}$ while keeping outgoing actions of $\mathcal{E}$
   - Collapse E to FN and remaining terminal states to TN
   - Add action τ leading from $s_{\mathcal{E}}$ to TN



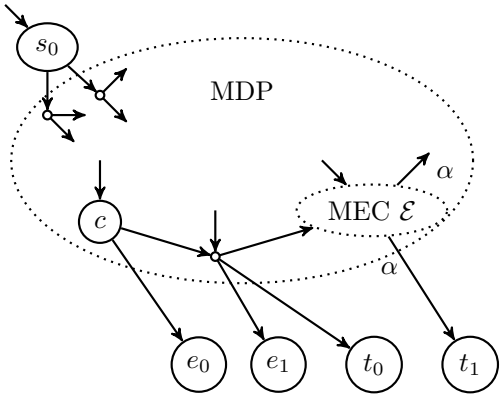
The Canonical MDP – Theorem

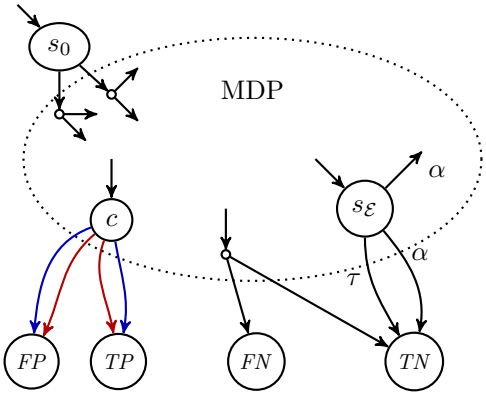




The Canonical MDP – Theorem

C is a SPR/GPR cause for E in $\mathcal{M}$ $\iff$ C is a SPR/GPR cause for $\{TP, FN\}$ in $\mathcal{M}_C$.





Algorithmic Studies

Problems for state-based probability-raising causes

PR causality in MDPs	Check condition C and E are given	Compute cause C for given effect E	
GPR cause	?		
SPR cause	?		

Algorithmic Studies

Problems for state-based probability-raising causes

PR causality in MDPs	Check condition C and E are given	Compute cause C for given effect E	
GPR cause			
SPR cause	?		

Algorithmic Studies – Checking SPR

(SPR) For each $c \in C$ and scheduler $\mathfrak{S}$ reaching c:

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

Algorithmic Studies – Checking SPR

(SPR) For each $c \in C$ and scheduler $\mathfrak{S}$ reaching c:

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

- Check for each $c \in C$ individually

Algorithmic Studies – Checking SPR

(SPR) For each $c \in C$ and scheduler $\mathfrak{S}$ reaching c:

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

- Check for each $c \in C$ individually
- Only consider scheduler which minimize $\Pr_c^{\mathfrak{S}}(\Diamond E)$ for every $c \in C$

Algorithmic Studies – Checking SPR

(SPR) For each $c \in C$ and scheduler $\mathfrak{S}$ reaching c:

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

- Check for each $c \in C$ individually
- Only consider scheduler which minimize $\Pr_c^{\mathfrak{S}}(\Diamond E)$ for every $c \in C$

  Delete action $\alpha_{\max}$ to get the MDP $\mathcal{N}$

Algorithmic Studies – Checking SPR

(SPR) For each $c \in C$ and scheduler $\mathfrak{S}$ reaching c:

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

- Check for each $c \in C$ individually
- Only consider scheduler which minimize $\Pr_c^{\mathfrak{S}}(\Diamond E)$ for every $c \in C$

  Delete action $\alpha_{\max}$ to get the MDP $\mathcal{N}$
- Compare reachability probabilities in $\mathcal{N}$

Algorithmic Studies – Checking SPR

(SPR) For each $c \in C$ and scheduler $\mathfrak{S}$ reaching c:

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

- Check for each $c \in C$ individually
- Only consider scheduler which minimize $\Pr_c^{\mathfrak{S}}(\Diamond E)$ for every $c \in C$

  Delete action α_{max} to get the MDP $\mathcal{N}$
- Compare reachability probabilities in $\mathcal{N}$

$$\underbrace{\Pr^{\mathsf{min}}_{\mathcal{M},c}(\Diamond E)}_{\text{minimal probability to reach } E \text{ from } c} \quad \bowtie \quad \underbrace{\Pr^{\mathsf{max}}_{\mathcal{N},s_0}(\Diamond E)}_{\text{maximal probability to reach } E \text{ in } \mathcal{N}}$$

Algorithmic Studies – Checking SPR

(SPR) For each $c \in C$ and scheduler $\mathfrak{S}$ reaching c:

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

- Check for each $c \in C$ individually
- Only consider scheduler which minimize $\Pr_c^{\mathfrak{S}}(\Diamond E)$ for every $c \in C$

  Delete action α_{max} to get the MDP $\mathcal{N}$
- Compare reachability probabilities in $\mathcal{N}$

$$\underbrace{\Pr^{\mathsf{min}}_{\mathcal{M},c}(\Diamond E)}_{\text{minimal probability to reach } E \text{ from } c} \quad \bowtie \quad \underbrace{\Pr^{\mathsf{max}}_{\mathcal{N},s_0}(\Diamond E)}_{\text{maximal probability to reach } E \text{ in } \mathcal{N}}$$

- Results in polynomial time

Algorithmic Studies

Problems for state-based probability-raising causes

PR causality in MDPs	Check condition C and E are given	Compute cause C for given effect E	
GPR cause			
SPR cause	$\in \mathrm{P}$		

Algorithmic Studies

Problems for state-based probability-raising causes

PR causality in MDPs	Check condition C and E are given	Compute cause C for given effect E	
GPR cause	?		
SPR cause	$\in \mathbf{P}$		

Algorithmic Studies – Checking GPR

(GPR) For each scheduler $\mathfrak{S}$ reaching C:

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond E \mid \Diamond C) > \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond E).$$

In $\mathcal{M}_C$ a terminal state is reached almost surely

Algorithmic Studies – Checking GPR

(GPR) For each scheduler $\mathfrak{S}$ reaching C:

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond E \mid \Diamond C) > \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond E).$$

In $\mathcal{M}_C$ a terminal state is reached almost surely

- Use state-action frequencies $x_{s,\alpha}$ to build a constraint system for scheduler $\mathfrak{S}$ of $\mathcal{M}_C$ [12]

[12] Lodewijk Kallenberg. *Lecture Notes Markov Decision Problems – version 2020.* Feb. 2020

Algorithmic Studies – Checking GPR

(GPR) For each scheduler $\mathfrak{S}$ reaching C:

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond E \mid \Diamond C) > \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond E).$$

In $\mathcal{M}_C$ a terminal state is reached almost surely

- Use state-action frequencies $x_{s,\alpha}$ to build a constraint system for scheduler $\mathfrak{S}$ of $\mathcal{M}_C$ [12]

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

$$x_s = \sum_{(t,\alpha)} x_{t,\alpha} \cdot P(t,\alpha,s) \qquad \text{for } s \in S \tag{S2}$$

[12] Lodewijk Kallenberg. *Lecture Notes Markov Decision Problems – version 2020.* Feb. 2020

Algorithmic Studies – Checking GPR

(GPR) For each scheduler $\mathfrak{S}$ reaching C:

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond E \mid \Diamond C) > \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond E).$$

In $\mathcal{M}_C$ the *refusal* of (GPR) is equivalent to

Algorithmic Studies – Checking GPR

(GPR) For each scheduler $\mathfrak{S}$ reaching C:

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond E \mid \Diamond C) > \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond E).$$

In $\mathcal{M}_C$ the *refusal* of (GPR) is equivalent to

$$x_{FN} \cdot x_{FP} - x_{TN} \cdot x_{TP} \geqslant 0 \tag{S3}$$

$$x_{\text{Cause}} > 0 \tag{S4}$$

Algorithmic Studies – Checking GPR

(GPR) For each scheduler $\mathfrak{S}$ reaching C:

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond E \mid \Diamond C) > \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond E).$$

Equations (S1)–(S4) express the existence of a scheduler which refuses (GPR)

Algorithmic Studies – Checking GPR

(GPR) For each scheduler $\mathfrak{S}$ reaching C:

$$\Pr_{\mathcal{M}}^{\mathfrak{S}}(\Diamond E \mid \Diamond C) > \Pr_{\mathcal{M}}^{\mathfrak{S}}(\Diamond E).$$

Equations (S1)–(S4) express the existence of a scheduler which refuses (GPR)

- quadratic system of inequalities by

$$x_{FN} \cdot x_{FP} - x_{TN} \cdot x_{TP} \geqslant 0 \qquad \text{(S3)}$$

Algorithmic Studies – Checking GPR

(GPR) For each scheduler $\mathfrak{S}$ reaching C:

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond E \mid \Diamond C) > \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond E).$$

Equations (S1)–(S4) express the existence of a scheduler which refuses (GPR)

- quadratic system of inequalities by

$$x_{FN} \cdot x_{FP} - x_{TN} \cdot x_{TP} \geqslant 0 \qquad \text{(S3)}$$

- Similar to quadratic programming[13], this can be decided in NP

[13] Stephen A. Vavasis. "Quadratic programming is in NP". In: *Information Processing Letters* 36.2 (1990), pp. 73–77

Algorithmic Studies – Checking GPR

(GPR) For each scheduler $\mathfrak{S}$ reaching C:

$$\Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond E \mid \Diamond C) > \Pr^{\mathfrak{S}}_{\mathcal{M}}(\Diamond E).$$

Equations (S1)–(S4) express the existence of a scheduler which refuses (GPR)

- quadratic system of inequalities by

$$x_{FN} \cdot x_{FP} - x_{TN} \cdot x_{TP} \geqslant 0 \qquad \text{(S3)}$$

- Similar to quadratic programming[13], this can be decided in NP
- Results in coNP for deciding (GPR)

[13] Stephen A. Vavasis. "Quadratic programming is in NP". In: *Information Processing Letters* 36.2 (1990), pp. 73–77

Algorithmic Studies

Problems for state-based probability-raising causes

PR causality in MDPs	Check condition C and E are given	Compute cause C for given effect E	
GPR cause	$\in$ coNP		
SPR cause	$\in$ P		

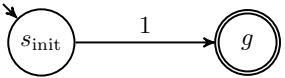
Algorithmic Studies

Problems for state-based probability-raising causes

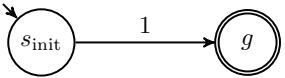
PR causality in MDPs	Check condition C and E are given	Compute cause C for given effect E	
GPR cause	$\in$ coNP		
SPR cause	$\in$ P		

Existence of PR causes

Existence of PR causes

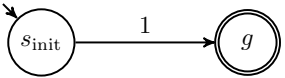


Existence of PR causes



Given an MDP $\mathcal{M}$ and an effect E, the following are equivalent

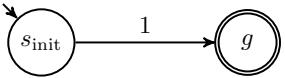
Existence of PR causes



Given an MDP $\mathcal{M}$ and an effect E, the following are equivalent

(a) There is an SPR cause C for E

Existence of PR causes

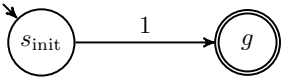


Given an MDP $\mathcal{M}$ and an effect E, the following are equivalent

(a) There is an SPR cause C for E

(b) There is a GPR cause C for E

Existence of PR causes



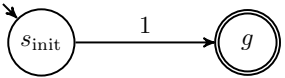
Given an MDP $\mathcal{M}$ and an effect E, the following are equivalent

(a) There is an SPR cause C for E

(b) There is a GPR cause C for E

(c) There is a singleton cause C for E

Existence of PR causes



Given an MDP $\mathcal{M}$ and an effect E, the following are equivalent

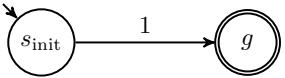
(a) There is an SPR cause C for E

(b) There is a GPR cause C for E

(c) There is a singleton cause C for E

- Existence of PR causes can be decided in poly-time

Existence of PR causes



Given an MDP $\mathcal{M}$ and an effect E, the following are equivalent

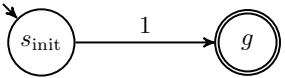
(a) There is an SPR cause C for E

(b) There is a GPR cause C for E

(c) There is a singleton cause C for E

- Existence of PR causes can be decided in poly-time
- *But* singletons as causes might not be desirable

Existence of PR causes



Given an MDP $\mathcal{M}$ and an effect E, the following are equivalent

(a) There is an SPR cause C for E

(b) There is a GPR cause C for E

(c) There is a singleton cause C for E

- Existence of PR causes can be decided in poly-time
- *But* singletons as causes might not be desirable
- Further criteria for computation are needed

Quality Measures – Confusion Matrix

Consider C as a classifier telling when a path becomes causal

Quality Measures – Confusion Matrix

Consider C as a classifier telling when a path becomes causal

- Leads to confusion matrix

Path hits	E	$\neg E$
C	True positive (tp) C correctly predicted E	False positive (fp) C falsely predicted E
$\neg C$	False negative (fn) C falsely not predicted E	True negative (tn) C correctly not predicted E

Quality Measures – Confusion Matrix

Consider C as a classifier telling when a path becomes causal

- Leads to confusion matrix

Path hits	E	$\neg E$
C	True positive (tp) C correctly predicted E	False positive (fp) C falsely predicted E
$\neg C$	False negative (fn) C falsely not predicted E	True negative (tn) C correctly not predicted E

- Measures from statistical analysis can rate the quality of a cause

Quality Measures – Definitions

Given MDP $\mathcal{M}$, effect E, PR cause C and scheduler $\mathfrak{S}$,

Quality Measures – Definitions

Given MDP $\mathcal{M}$, effect E, PR cause C and scheduler $\mathfrak{S}$, define e.g.

Quality Measures – Definitions

Given MDP $\mathcal{M}$, effect E, PR cause C and scheduler $\mathfrak{S}$, define e.g.

$$\mathit{recall}^{\mathfrak{S}}(\mathsf{Cause}) \;=\; \frac{\mathsf{tp}^{\mathfrak{S}}}{\mathsf{tp}^{\mathfrak{S}}+\mathsf{fn}^{\mathfrak{S}}} \;=\; \Pr^{\mathfrak{S}}_{\mathcal{M}}(\ \Diamond\mathsf{Cause} \mid \Diamond\mathsf{Eff}\ )$$

Quality Measures – Definitions

Given MDP $\mathcal{M}$, effect E, PR cause C and scheduler $\mathfrak{S}$, define e.g.

$$
\begin{aligned}
\mathit{recall}^{\mathfrak{S}}(\mathsf{Cause}) &= \frac{\mathsf{tp}^{\mathfrak{S}}}{\mathsf{tp}^{\mathfrak{S}}+\mathsf{fn}^{\mathfrak{S}}} &&= \Pr^{\mathfrak{S}}_{\mathcal{M}}(\ \Diamond\mathsf{Cause} \mid \Diamond\mathsf{Eff}\ ) \\
\mathit{fscore}^{\mathfrak{S}}(\mathsf{Cause}) &= \frac{2\cdot\mathsf{tp}^{\mathfrak{S}}}{2\cdot\mathsf{tp}^{\mathfrak{S}}+\mathsf{fp}^{\mathfrak{S}}+\mathsf{fn}^{\mathfrak{S}}}
\end{aligned}
$$

Quality Measures – Definitions

Given MDP $\mathcal{M}$, effect E, PR cause C and scheduler $\mathfrak{S}$, define e.g.

$$\begin{aligned} \mathit{recall}^{\mathfrak{S}}(\mathsf{Cause}) &= \frac{\mathsf{tp}^{\mathfrak{S}}}{\mathsf{tp}^{\mathfrak{S}}+\mathsf{fn}^{\mathfrak{S}}} &= \Pr^{\mathfrak{S}}_{\mathcal{M}}(\ \Diamond\mathsf{Cause} \mid \Diamond\mathsf{Eff}\ ) \\ \mathit{fscore}^{\mathfrak{S}}(\mathsf{Cause}) &= \frac{2\cdot\mathsf{tp}^{\mathfrak{S}}}{2\cdot\mathsf{tp}^{\mathfrak{S}}+\mathsf{fp}^{\mathfrak{S}}+\mathsf{fn}^{\mathfrak{S}}} \end{aligned}$$

- Take worst case under sensible schedulers for a general definition e.g.

Quality Measures – Definitions

Given MDP $\mathcal{M}$, effect E, PR cause C and scheduler $\mathfrak{S}$, define e.g.

$$\begin{aligned} \mathit{recall}^{\mathfrak{S}}(\mathsf{Cause}) &= \frac{\mathsf{tp}^{\mathfrak{S}}}{\mathsf{tp}^{\mathfrak{S}}+\mathsf{fn}^{\mathfrak{S}}} = \Pr^{\mathfrak{S}}_{\mathcal{M}}(\ \Diamond\mathsf{Cause} \mid \Diamond\mathsf{Eff}\ ) \\ \mathit{fscore}^{\mathfrak{S}}(\mathsf{Cause}) &= \frac{2\cdot\mathsf{tp}^{\mathfrak{S}}}{2\cdot\mathsf{tp}^{\mathfrak{S}}+\mathsf{fp}^{\mathfrak{S}}+\mathsf{fn}^{\mathfrak{S}}} \end{aligned}$$

- Take worst case under sensible schedulers for a general definition e.g.

$$\mathit{fscore}(\mathsf{Cause}) = \inf_{\mathfrak{S}} \mathit{fscore}^{\mathfrak{S}}(\mathsf{Cause})$$

Algorithmic Studies

Problems for state-based probability-raising causes

PR causality in MDPs	Check condition C and E are given	Compute cause C for given effect E	
GPR cause	$\in$ coNP		
SPR cause	$\in$ P		

Algorithmic Studies

Problems for state-based probability-raising causes

PR causality in MDPs	Check condition C and E are given	Compute cause C for given effect E	
		recall-optimal	*fscore*-optimal
GPR cause	$\in$ coNP		
SPR cause	$\in$ P		

Algorithmic Studies

Problems for state-based probability-raising causes

PR causality in MDPs	Check condition C and E are given	Compute cause C for given effect E	
		recall-optimal	*fscore*-optimal
GPR cause	$\in$ coNP	poly-space NP-hard	poly-space NP-hard
SPR cause	$\in$ P	poly-time	poly-space threshold $\in$ NP $\cap$ coNP

Algorithmic Studies

Problems for state-based probability-raising causes

PR causality in MDPs	Check condition C and E are given	Compute cause C for given effect E	
		recall-optimal	*fscore*-optimal
GPR cause	$\in$ coNP	poly-space NP-hard	poly-space NP-hard
SPR cause	$\in$ P	poly-time	poly-space threshold $\in$ NP $\cap$ coNP

Algorithmic Studies

Problems for state-based probability-raising causes

PR causality in MDPs	Check condition C and E are given	Compute cause C for given effect E	
		recall-optimal	*fscore*-optimal
GPR cause	$\in$ coNP	poly-space NP-hard	poly-space NP-hard
SPR cause	$\in$ P	poly-time	poly-space threshold $\in$ NP $\cap$ coNP

Algorithmic Studies

Problems for state-based probability-raising causes

PR causality in MDPs	Check condition C and E are given	Compute cause C for given effect E	
		recall-optimal	*fscore*-optimal
GPR cause	$\in$ coNP	poly-space NP-hard	poly-space NP-hard
SPR cause	$\in$ P	poly-time	poly-space threshold $\in$ NP $\cap$ coNP

Algorithmic Studies

Problems for state-based probability-raising causes

PR causality in DTMCs	Check condition C and E are given	Compute cause C for given effect E	
		recall-optimal	*fscore*-optimal
GPR cause	$\in$ P	NP-complete	NP-complete
SPR cause	$\in$ P	poly-time	poly-time

Algorithmic Studies

Problems for state-based probability-raising causes

PR causality in DTMCs	Check condition C and E are given	Compute cause C for given effect E	
		recall-optimal	*fscore*-optimal
GPR cause	$\in$ P	NP-complete	NP-complete
SPR cause	$\in$ P	poly-time to exp-time	poly-time

For $recall = 1$ with additional cost constraints in Robin Ziemek et al. "Probabilistic causes in Markov chains". In: *Innovations in Systems and Software Engineering* (2022)

Potential PR Causes – Existential Quantification

(GPR) For all scheduler $\mathfrak{S}$ reaching C: $\Pr^{\mathfrak{S}}(\ \Diamond E \mid \Diamond C\ ) > \Pr^{\mathfrak{S}}(\Diamond E)$.

(SPR) For each $c \in C$ and all scheduler $\mathfrak{S}$ reaching c:

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

Potential PR Causes – Existential Quantification

(GPR) There is a scheduler $\mathfrak{S}$ reaching C: $\Pr^{\mathfrak{S}}(\ \Diamond E \mid \Diamond C\ ) > \Pr^{\mathfrak{S}}(\Diamond E)$.

(SPR) For each $c \in C$ and all scheduler $\mathfrak{S}$ reaching c:

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

Potential PR Causes – Existential Quantification

(GPR) There is a scheduler $\mathfrak{S}$ reaching C: $\Pr^{\mathfrak{S}}(\ \Diamond E \mid \Diamond C\ ) > \Pr^{\mathfrak{S}}(\Diamond E)$.
(SPR) There is a scheduler $\mathfrak{S}$ reaching C such that for c with $\Pr((\neg C)\,\mathrm{U}\,c) > 0$:

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

Potential PR Causes – Existential Quantification

(GPR) There is a scheduler $\mathfrak{S}$ reaching C: $\Pr^{\mathfrak{S}}(\ \Diamond E \mid \Diamond C\ ) > \Pr^{\mathfrak{S}}(\Diamond E)$.
(SPR) There is a scheduler $\mathfrak{S}$ reaching C such that for c with $\Pr((\neg C)\,\mathrm{U}\,c) > 0$:

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

- Suitable if the non-determinism has strong influence on the system

Potential PR Causes – Existential Quantification

(GPR) There is a scheduler $\mathfrak{S}$ reaching C: $\Pr^{\mathfrak{S}}(\ \Diamond E \mid \Diamond C\ ) > \Pr^{\mathfrak{S}}(\Diamond E)$.
(SPR) There is a scheduler $\mathfrak{S}$ reaching C such that for c with $\Pr((\neg C)\,\mathrm{U}\,c) > 0$:

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

- Suitable if the non-determinism has strong influence on the system
- Conditions can be checked with similar complexity

Potential PR Causes – Existential Quantification

(GPR) There is a scheduler $\mathfrak{S}$ reaching C: $\Pr^{\mathfrak{S}}(\ \Diamond E \mid \Diamond C\ ) > \Pr^{\mathfrak{S}}(\Diamond E)$.
(SPR) There is a scheduler $\mathfrak{S}$ reaching C such that for c with $\Pr((\neg C)\,\mathrm{U}\,c) > 0$:

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

- Suitable if the non-determinism has strong influence on the system
- Conditions can be checked with similar complexity
  - Potential SPR in P
  - Potential GPR in NP

Potential PR Causes – Existential Quantification

(GPR) There is a scheduler $\mathfrak{S}$ reaching C: $\Pr^{\mathfrak{S}}(\ \Diamond E \mid \Diamond C\ ) > \Pr^{\mathfrak{S}}(\Diamond E)$.
(SPR) There is a scheduler $\mathfrak{S}$ reaching C such that for c with $\Pr((\neg C)\,\mathrm{U}\,c) > 0$:

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

- Suitable if the non-determinism has strong influence on the system
- Conditions can be checked with similar complexity
  - Potential SPR in P
  - Potential GPR in NP
- Computing PR causes optimal for quality-measures can be misguided
  - There may be only a few probability-raising scheduler

Potential PR Causes – Existential Quantification

(GPR) There is a scheduler $\mathfrak{S}$ reaching C: $\Pr^{\mathfrak{S}}(\ \Diamond E \mid \Diamond C\ ) \ > \ \Pr^{\mathfrak{S}}(\Diamond E)$.
(SPR) There is a scheduler $\mathfrak{S}$ reaching C such that for c with $\Pr((\neg C)\,\mathrm{U}\,c) > 0$:

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

- Suitable if the non-determinism has strong influence on the system
- Conditions can be checked with similar complexity
  - Potential SPR in P
  - Potential GPR in NP
- Computing PR causes optimal for quality-measures can be misguided
  - There may be only a few probability-raising scheduler
- Other criteria for computation are needed

Probability-Raising Causality in MDPs

Setting for probability-raising conditions depends on definition of cause C and effect E

	state-based C and E sets of states	*mixed* C set of states E path property	*path-based* C and E path properties
global PR C viewed as unit	GPR cause	reachability GPR cause	co-safety GPR cause
strict PR elements in C individually	SPR cause	reachability SPR cause	co-safety SPR cause

Probability-Raising Causality in MDPs

Setting for probability-raising conditions depends on definition of cause C and effect E

	state-based C and E sets of states	*mixed* C set of states E path property	*path-based* C and E path properties
global PR C viewed as unit	GPR cause	reachability GPR cause	co-safety GPR cause
strict PR elements in C individually	SPR cause	reachability SPR cause	co-safety SPR cause

Extension to Mixed Setting

- Effect is ω-regular language rEff $\subseteq S^\omega$ represented by DRA $\mathcal{A}_{\mathsf{rEff}}$

Extension to Mixed Setting

- Effect is ω-regular language rEff $\subseteq S^\omega$ represented by DRA $\mathcal{A}_{\mathsf{rEff}}$
- Cause is set of states C

Extension to Mixed Setting

- Effect is ω-regular language rEff $\subseteq S^\omega$ represented by DRA $\mathcal{A}_{\text{rEff}}$
- Cause is set of states C
- We get reachability PR conditions of the form:

$$\Pr^{\mathfrak{S}}(\text{rEff} \mid \Diamond C) \;>\; \Pr^{\mathfrak{S}}(\text{rEff}) \qquad \text{(rGPR)}$$

$$Pr^{\mathfrak{S}}(\text{rEff} \mid (\neg C)\,\mathrm{U}\,c) \;>\; \Pr^{\mathfrak{S}}(\text{rEff}) \qquad \text{(rSPR)}$$

Extension to Mixed Setting

- Effect is ω-regular language rEff $\subseteq S^{\omega}$ represented by DRA $\mathcal{A}_{\text{rEff}}$
- Cause is set of states C
- We get reachability PR conditions of the form:

$$\Pr^{\mathfrak{S}}(\text{rEff} \mid \Diamond C) \;>\; \Pr^{\mathfrak{S}}(\text{rEff}) \qquad \text{(rGPR)}$$

$$Pr^{\mathfrak{S}}(\text{rEff} \mid (\neg C)\,\mathrm{U}\,c) \;>\; \Pr^{\mathfrak{S}}(\text{rEff}) \qquad \text{(rSPR)}$$

- Reduction to state-based setting through product MDP $\mathcal{M} \otimes \mathcal{A}_{\text{rEff}}$,

Extension to Mixed Setting

- Effect is ω-regular language rEff $\subseteq S^\omega$ represented by DRA $\mathcal{A}_{\text{rEff}}$
- Cause is set of states C
- We get reachability PR conditions of the form:

$$\Pr^{\mathfrak{S}}(\text{rEff} \mid \Diamond C) \;>\; \Pr^{\mathfrak{S}}(\text{rEff}) \qquad \text{(rGPR)}$$

$$Pr^{\mathfrak{S}}(\text{rEff} \mid (\neg C)\,\mathrm{U}\,c) \;>\; \Pr^{\mathfrak{S}}(\text{rEff}) \qquad \text{(rSPR)}$$

- Reduction to state-based setting through product MDP $\mathcal{M} \otimes \mathcal{A}_{\text{rEff}}$, where reachability SPR reduces to state-based GPR

Probability-Raising Causality in MDPs

Setting for probability-raising conditions depends on definition of cause C and effect E

	state-based C and E sets of states	*mixed* C set of states E path property	*path-based* C and E path properties
global PR C viewed as unit	GPR cause	reachability GPR cause	co-safety GPR cause
strict PR elements in C individually	SPR cause	reachability SPR cause	co-safety SPR cause

Extension to Path-Based Setting

- Effect is ω-regular language rEff $\subseteq S^\omega$ represented by DRA $\mathcal{A}_{\mathsf{rEff}}$

Extension to Path-Based Setting

- Effect is ω-regular language rEff $\subseteq S^\omega$ represented by DRA $\mathcal{A}_{\mathsf{rEff}}$
- Cause is ω-regular co-safety language rCause $\subseteq S^\omega$ represented by DFA $\mathcal{A}_{\mathsf{rCause}}$

Extension to Path-Based Setting

- Effect is ω-regular language rEff $\subseteq S^\omega$ represented by DRA $\mathcal{A}_{\mathsf{rEff}}$
- Cause is ω-regular co-safety language rCause $\subseteq S^\omega$ represented by DFA $\mathcal{A}_{\mathsf{rCause}}$
- We get co-safety PR conditions of the form

$$\Pr^{\mathfrak{S}}(\mathsf{rEff} \mid \mathsf{rCause}) > \Pr^{\mathfrak{S}}(\mathsf{rEff}) \qquad \text{(cosafeGPR)}$$

$$\Pr^{\mathfrak{S}}(\mathsf{rEff} \mid \sigma) > \Pr^{\mathfrak{S}}(\mathsf{rEff}) \qquad \text{(cosafeSPR)}$$

where σ is minimal good prefix for rCause

Extension to Path-Based Setting

- Effect is ω-regular language rEff $\subseteq S^\omega$ represented by DRA $\mathcal{A}_{\mathsf{rEff}}$
- Cause is ω-regular co-safety language rCause $\subseteq S^\omega$ represented by DFA $\mathcal{A}_{\mathsf{rCause}}$
- We get co-safety PR conditions of the form

$$\Pr^{\mathfrak{S}}(\mathsf{rEff} \mid \mathsf{rCause}) \;>\; \Pr^{\mathfrak{S}}(\mathsf{rEff}) \qquad \text{(cosafeGPR)}$$

$$\Pr^{\mathfrak{S}}(\mathsf{rEff} \mid \sigma) \;>\; \Pr^{\mathfrak{S}}(\mathsf{rEff}) \qquad \text{(cosafeSPR)}$$

  where σ is minimal good prefix for rCause
- SPR condition becomes difficult when using reduction via product MDP

Extension to Path-Based Setting

- Effect is ω-regular language rEff $\subseteq S^\omega$ represented by DRA $\mathcal{A}_{\mathsf{rEff}}$
- Cause is ω-regular co-safety language rCause $\subseteq S^\omega$ represented by DFA $\mathcal{A}_{\mathsf{rCause}}$
- We get co-safety PR conditions of the form

$$\Pr^{\mathfrak{S}}(\mathsf{rEff} \mid \mathsf{rCause}) \;>\; \Pr^{\mathfrak{S}}(\mathsf{rEff}) \qquad \text{(cosafeGPR)}$$

$$\Pr^{\mathfrak{S}}(\mathsf{rEff} \mid \sigma) \;>\; \Pr^{\mathfrak{S}}(\mathsf{rEff}) \qquad \text{(cosafeSPR)}$$

  where σ is minimal good prefix for rCause
- SPR condition becomes difficult when using reduction via product MDP
- Quality-optimal causes may not exist

Summary

- Formal notions of state-based and path-based probability-raising causality in MDPs

Summary

- Formal notions of state-based and path-based probability-raising causality in MDPs
- Algorithmic foundations for state-based setting

Summary

- Formal notions of state-based and path-based probability-raising causality in MDPs
- Algorithmic foundations for state-based setting
  - Check cause-effect relation
  - Compute quality-optimal causes

Summary

- Formal notions of state-based and path-based probability-raising causality in MDPs
- Algorithmic foundations for state-based setting
  - Check cause-effect relation
  - Compute quality-optimal causes
- Existential quantification for PR condition

Summary

- Formal notions of state-based and path-based probability-raising causality in MDPs
- Algorithmic foundations for state-based setting
  - Check cause-effect relation
  - Compute quality-optimal causes
- Existential quantification for PR condition
- Extension of the algorithmic framework to path-based setting

Summary

- Formal notions of state-based and path-based probability-raising causality in MDPs
- Algorithmic foundations for state-based setting
  - Check cause-effect relation
  - Compute quality-optimal causes
- Existential quantification for PR condition
- Extension of the algorithmic framework to path-based setting
  - With some difficult open problems

Further Research

Further Research

- Go from theoretical foundations to application

Further Research

- Go from theoretical foundations to application
  - Implementation is in work

Further Research

- Go from theoretical foundations to application
  - Implementation is in work
  - Collaboration with domain experts on explanations

Further Research

- Go from theoretical foundations to application
  - Implementation is in work
  - Collaboration with domain experts on explanations
- Other criteria for computation

  e.g. relative portion of PR scheduler

Further Research

- Go from theoretical foundations to application
  - Implementation is in work
  - Collaboration with domain experts on explanations
- Other criteria for computation

  e.g. relative portion of PR scheduler
- Extend the framework to stochastic games

Further Research

- Go from theoretical foundations to application
  - Implementation is in work
  - Collaboration with domain experts on explanations
- Other criteria for computation

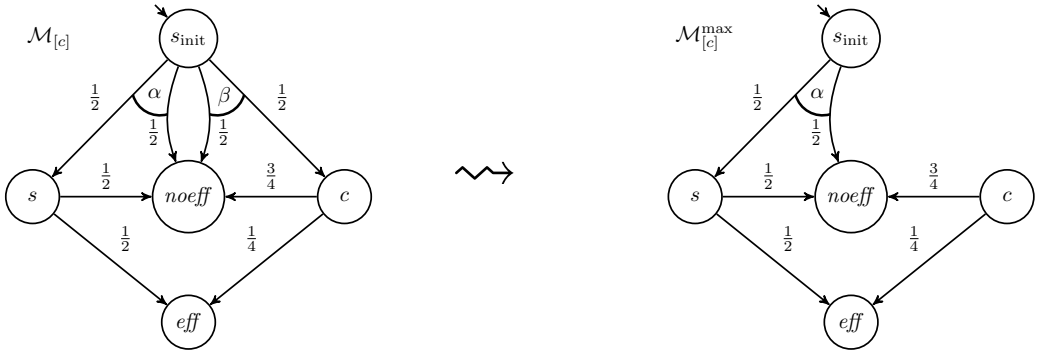
  e.g. relative portion of PR scheduler
- Extend the framework to stochastic games

1. Introduction
   - Motivation
   - Markov Decision Process
2. Probability-Raising Causality in MDPs
   - Notions of Cause-Effect Relations
   - Simple Properties
   - Algorithmic Studies
3. Variants of PR Causality
   - Potential PR Causes
   - Reachability PR Causes
   - Co-Safety PR Causes
4. Conclusion

Backup Table of Contents

5. SPR Check for Equality Case
6. Halpern & Pearl Definition
7. Kleinbergs Probabilistic Causes in DTMCs
8. p-Causes in DTMCs
   - Cost Functions
9. Algorithmic Results for Path-Based PR Causes

Example for Equality Case of SPR Check



- $w_c = \Pr_c^{\min}(\Diamond \text{eff}) = \frac{1}{4} = \Pr^{\max}(\Diamond \text{eff})$ and c not reachable in $\mathcal{M}^{\max}_{[c]}$ implies case 3.2
- Consider $\mathfrak{S}$ taking β with probability $\varepsilon > 0$

$\Rightarrow \Pr^{\mathfrak{S}}(\Diamond \text{eff} \mid \Diamond c) = \frac{1}{4} > (1-\varepsilon)\frac{1}{4} + \varepsilon \cdot \frac{1}{4} = \Pr^{\mathfrak{S}}(\Diamond \text{eff})$

Halpern & Pearl: Counterfactual Causality [14]

For a causal model $W = (U, V\{f_x\}_{x \in V})$ a formula φ and a context $\overrightarrow{u}$ an assignment $\overrightarrow{X} = \overrightarrow{x}$ is an *actual cause* of φ in $(W, \overrightarrow{u})$ if

AC1 $(W, \overrightarrow{u})$ satisfies $\overrightarrow{X} = \overrightarrow{x}$ and φ

AC2 $[\overrightarrow{X} \leftarrow \overrightarrow{x}']\neg\varphi$ holds in $(W, \overrightarrow{u})$

AC3 $\overrightarrow{X} = \overrightarrow{x}$ is minimal

[14] Joseph Y. Halpern and Judea Pearl. "Causes and Explanations: A Structural-Model Approach: Part 1: Causes". In: *Conference in Uncertainty in Artificial Intelligence (UAI)*. 2001, pp. 194–202

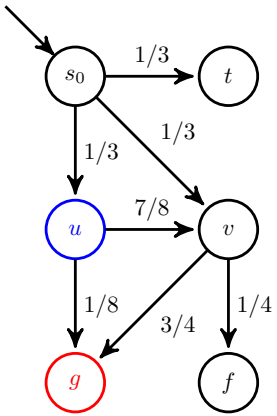
Kleinberg et. al.[15]: State-based PR in DTMCs

Probability-raising in Markov chains

- Effect and cause given as PCTL formulas
  e.g. $\varphi_{\text{eff}} = \forall \Diamond \Pr_{\geqslant 0.9}(\Box g)$
- Equivalent: sets of states Eff and Cause
  e.g. $\text{Eff} = \{s \in S \mid s \models \varphi_{\text{eff}}\} = \{g\}$

Cause $\subseteq S \setminus \text{Eff}$ is *prima facie cause* of $\Diamond$Eff if:

- $\forall c \in \text{Cause}: \Pr_c(\Diamond \text{Eff}) \geqslant p$
- $\Pr_{s_0}(\Diamond \text{Eff}) < p$



[15] Samantha Kleinberg and Bud Mishra. "The Temporal Logic of Causal Structures". In: *25th Conference on Uncertainty in Artificial Intelligence (UAI)*. 2009, pp. 303–312

p-Causes: Probability-Raising in DTMCs

Building on Kleinberg we define p-causes in a DTMC M

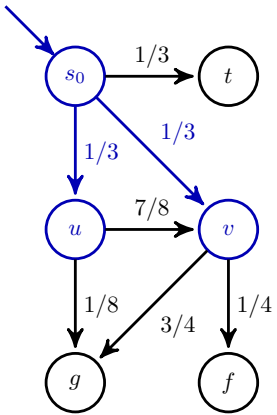
- Effect is ω-regular language $\mathcal{L} \subseteq S^\omega$
- Cause is a set of finite paths in M

Cause is a p-cause for $\mathcal{L}$ if it is prefix-free and

- almost every $\pi \in \mathcal{L}$ has a prefix in Cause
- for every $\hat{\pi} \in$ Cause we have $\Pr(\mathcal{L} \mid \hat{\pi}) \geqslant p$.

Example

- Consider $\mathcal{L} = \Diamond g$ and $p = 0.75$
- Then Cause $= \{s_0u,\ s_0v\}$ is a p-cause



p-Causes: Probability-Raising in DTMCs

Can be reduced to *state-based effects*

- Given deterministic Rabin automaton $\mathcal{A}$ for $\mathcal{L}$
- Consider product $M \otimes \mathcal{A}$ and union of all accepting BSCCs U
- For all finite paths $\hat{\pi}$ of M:

$$\Pr_M(\mathcal{L} \mid \hat{\pi}) = \Pr_{M \otimes \mathcal{A}}\big(\Diamond U \mid a(\hat{\pi})\big)$$

- It suffices to consider sets of terminal effect states

Cost Functions

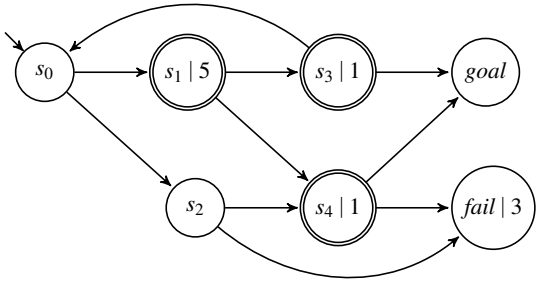
- Systems consume resources with budget restrictions
- Cost can be modeled as state weights $c : S \to \mathbb{Q}$
- Induce a cost function on finite executions $c : S^* \to \mathbb{Q}$ by

$$c(s_0 \cdots s_n) := \sum_{i=0}^{n} c(s_i) \quad \text{(accumulative)} \qquad \text{or} \qquad c(s_0 \cdots s_n) = c(s_n) \quad (\textit{instantaneous})$$

Cost Variants

Given a weight function $c : S \to \mathbb{Q}$ modeling costs

- Optimize expected (accumulated) cost in all scenarios
- Optimize partial expected (accumulated) cost (only if cause is reached)
- Optimize for lowest maximal cost of any path in the cause



Complexity of Computing Optimal p-Causes

Summary of complexity results for computing cost-optimal p-causes

	expcost	pexpcost	maxcost
non-negative weights accumulated	poly-time	exponential time PP-hard	poly-time
arbitrary weights accumulated	poly-time	PP-hard	pseudo-polyn. in $\mathtt{NP} \cap \mathtt{coNP}$
arbitrary weights instantaneous	poly-time	poly-time	poly-time

Algorithmic Studies

Problems for mixed probability-raising causes

PR causality in MDPs	Check condition C and E are given	Compute cause C for given effect E	
		recall-optimal	*fscore*-optimal
rGPR cause	$\in$ coNP	poly-space NP-hard	
rSPR cause	$\in$ coNP	$\in \mathrm{PF}^{\mathrm{NP}}$ threshold $\in$ coNP	poly-space threshold $\in \Sigma_2^{\mathrm{P}}$

Algorithmic Studies

Problems for path-based probability-raising causes

PR causality in MDPs	Check condition C and E are given	Compute cause C for given effect E	
		recall-optimal	*fscore*-optimal
cosafeGPR cause	$\in$ coNP	in general no optimal causes	open
cosafeSPR cause	difficult by checking for each prefix	optimal cause known computation unclear	open

Reduction to State-Based Effects

Effect is represented by a deterministic Rabin automaton (DRA) $\mathcal{A}_{\mathsf{rEff}}$

- Take the product $\mathcal{M} \otimes \mathcal{A}_{\mathsf{rEff}}$ and classify MECs as in the confusion matrix
- PR states in $\mathcal{M}$ correspond to PR sets of states in $\mathcal{M} \otimes \mathcal{A}_{\mathsf{rEff}}$
- rSPR causes in $\mathcal{M}$ correspond to GPR causes in $\mathcal{M} \otimes \mathcal{A}_{\mathsf{rEff}}$