Special subclass of Generalized Semi-Markov Decision Processes with discrete time

Alexander Frank

TU Dortmund

September 14, 2020

Table of Contents

Motivation

Decision making, when and why?

- **Event based**

(a) Status Quo

(b) Event

(c) Decision making

Figure: Event based decision making

Decision making, when and why?

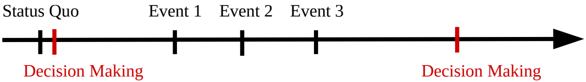
- **Discrete time steps**

(a) Status Quo

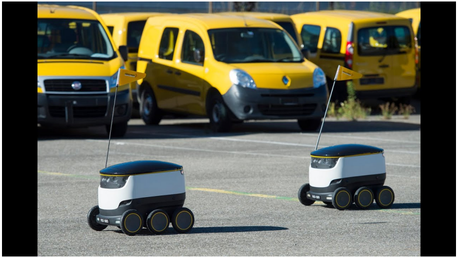

(b) Decision making


(c) Events in period

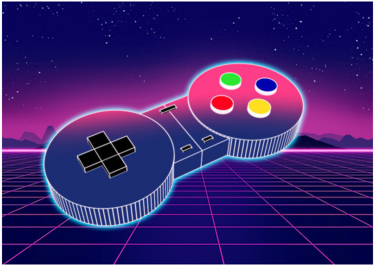
Figure: Decision making for periods



Discrete time steps

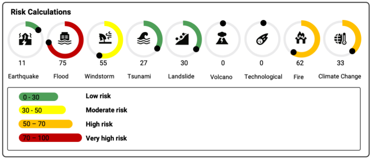


Decentralized problems



Stochastic games

Planning management



Risk calculations

GSMDPs

Example

Actions:	Safe-Modus		Risk-Modus	
Events:	**Prob**	**Effect**	**Prob**	**Effect**
Error 1	10,00 %	-1	15,00 %	-1
Error 2	20,00 %	-1 to -2	25,00 %	-2
Fatal Error	5,00 %	-3	10,00 %	-3 to -4
Self-Repair	35,00 %	2	-	-

Figure: Safe-Modus: Transitions of the example

Example

Actions:	Safe-Modus		Risk-Modus	
Events:	Prob	Effect	Prob	Effect
Error 1	10,00 %	-1	15,00 %	-1
Error 2	20,00 %	-1 to -2	25,00 %	-2
Fatal Error	5,00 %	-3	10,00 %	-3 to -4
Self-Repair	35,00 %	2	-	-

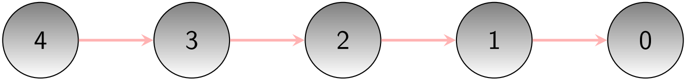


Figure: Safe-Modus: Transitions of the example

Example

Actions:	Safe-Modus		Risk-Modus	
Events:	**Prob**	**Effect**	**Prob**	**Effect**
Error 1	10,00 %	-1	15,00 %	-1
Error 2	20,00 %	-1 to -2	25,00 %	-2
Fatal Error	5,00 %	-3	10,00 %	-3 to -4
Self-Repair	35,00 %	2	-	-

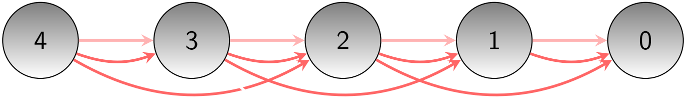


Figure: Safe-Modus: Transitions of the example

Example

Actions:	Safe-Modus		Risk-Modus	
Events:	**Prob**	**Effect**	**Prob**	**Effect**
Error 1	10,00 %	-1	15,00 %	-1
Error 2	20,00 %	-1 to -2	25,00 %	-2
Fatal Error	5,00 %	-3	10,00 %	-3 to -4
Self-Repair	35,00 %	2	-	-

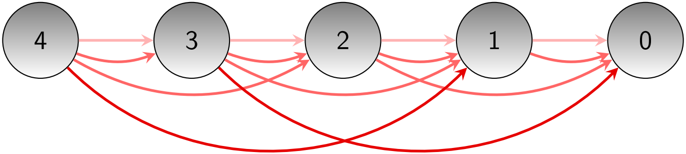


Figure: Safe-Modus: Transitions of the example

Example

Actions:	Safe-Modus		Risk-Modus	
Events:	**Prob**	**Effect**	**Prob**	**Effect**
Error 1	10,00 %	-1	15,00 %	-1
Error 2	20,00 %	-1 to -2	25,00 %	-2
Fatal Error	5,00 %	-3	10,00 %	-3 to -4
Self-Repair	35,00 %	2	-	-

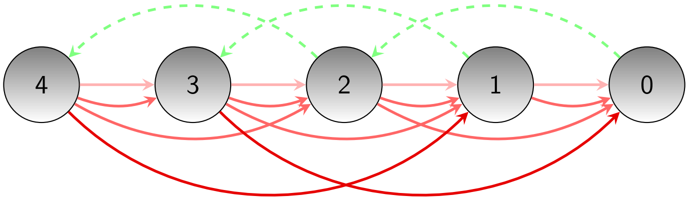


Figure: Safe-Modus: Transitions of the example

GSMDP Formulation

A GSMDP is given by a 7-tuple $(\mathcal{S}, \mathcal{A}, \mathcal{E}, \mathbf{C}, \mathbf{P}, \mathbf{R}, F)$ with

- $\mathcal{S}$ a set of states
- $\mathcal{A}$ a set of actions
- $\mathcal{E}$ a set of events (decreasing priority)
- $\mathbf{C}$: $\mathcal{S} \times \mathcal{A} \times \mathcal{E} \to \{0, 1\}$ mapping for (in)active events
- $\mathbf{P}$: $\mathcal{S} \times \mathcal{E} \to \mathcal{S}$ transition function from state to state by an event
- $\mathbf{R}$: $\mathcal{S} \times \mathcal{E}_0 \times \mathcal{S} \to \mathbb{R}$ rewards for state transitions by an event
- F : $\mathcal{E} \times \mathbb{N} \to [0, 1]$ probabilities for events depending on time T

Problem:

π^* : $\mathcal{S} \times T_1 \times ... \times T_{\mathcal{E}} \to \mathcal{A}$ with $T_i = \mathbb{N}$ optimal policy with best rewards!

GSMDP Formulation

A GSMDP is given by a 7-tuple $(\mathcal{S}, \mathcal{A}, \mathcal{E}, \mathbf{C}, \mathbf{P}, \mathbf{R}, F)$ with

- $\mathcal{S}$ a set of states
- $\mathcal{A}$ a set of actions
- $\mathcal{E}$ a set of events (decreasing priority)
- $\mathbf{C}$: $\mathcal{S} \times \mathcal{A} \times \mathcal{E} \rightarrow \{0, 1\}$ mapping for (in)active events
- $\mathbf{P}$: $\mathcal{S} \times \mathcal{E} \rightarrow \mathcal{S}$ transition function from state to state by an event
- $\mathbf{R}$: $\mathcal{S} \times \mathcal{E}_0 \times \mathcal{S} \rightarrow \mathbb{R}$ rewards for state transitions by an event
- F : $\mathcal{E} \times \mathbb{N} \rightarrow [0, 1]$ probabilities for events depending on time T

Problem:

π^* : $\mathcal{S} \times T_1 \times ... \times T_{\mathcal{E}} \rightarrow \mathcal{A}$ with $T_i = \mathbb{N}$ optimal policy with best rewards!

GSMDP Formulation

A GSMDP is given by a 7-tuple $(\mathcal{S}, \mathcal{A}, \mathcal{E}, \mathbf{C}, \mathbf{P}, \mathbf{R}, F)$ with

- $\mathcal{S}$ a set of states
- $\mathcal{A}$ a set of actions
- $\mathcal{E}$ a set of events (decreasing priority)
- $\mathbf{C}$: $\mathcal{S} \times \mathcal{A} \times \mathcal{E} \rightarrow \{0, 1\}$ mapping for (in)active events
- $\mathbf{P}$: $\mathcal{S} \times \mathcal{E} \rightarrow \mathcal{S}$ transition function from state to state by an event
- $\mathbf{R}$: $\mathcal{S} \times \mathcal{E}_0 \times \mathcal{S} \rightarrow \mathbb{R}$ rewards for state transitions by an event
- F : $\mathcal{E} \times \mathbb{N} \rightarrow [0, 1]$ probabilities for events depending on time T

Problem:

π^* : $\mathcal{S} \times T_1 \times ... \times T_{\mathcal{E}} \rightarrow \mathcal{A}$ with $T_i = \mathbb{N}$ optimal policy with best rewards!

GSMDP Formulation

A GSMDP is given by a 7-tuple $(\mathcal{S}, \mathcal{A}, \mathcal{E}, \mathbf{C}, \mathbf{P}, \mathbf{R}, F)$ with

- $\mathcal{S}$ a set of states
- $\mathcal{A}$ a set of actions
- $\mathcal{E}$ a set of events (decreasing priority)
- $\mathbf{C}$: $\mathcal{S} \times \mathcal{A} \times \mathcal{E} \to \{0, 1\}$ mapping for (in)active events
- $\mathbf{P}$: $\mathcal{S} \times \mathcal{E} \to \mathcal{S}$ transition function from state to state by an event
- $\mathbf{R}$: $\mathcal{S} \times \mathcal{E}_0 \times \mathcal{S} \to \mathbb{R}$ rewards for state transitions by an event
- F : $\mathcal{E} \times \mathbb{N} \to [0, 1]$ probabilities for events depending on time T

Problem:

π^* : $\mathcal{S} \times T_1 \times ... \times T_{\mathcal{E}} \to \mathcal{A}$ with $T_i = \mathbb{N}$ optimal policy with best rewards!

GSMDP Formulation

A GSMDP is given by a 7-tuple $(\mathcal{S}, \mathcal{A}, \mathcal{E}, \mathbf{C}, \mathbf{P}, \mathbf{R}, F)$ with

- $\mathcal{S}$ a set of states
- $\mathcal{A}$ a set of actions
- $\mathcal{E}$ a set of events (decreasing priority)
- $\mathbf{C}$: $\mathcal{S} \times \mathcal{A} \times \mathcal{E} \rightarrow \{0, 1\}$ mapping for (in)active events
- $\mathbf{P}$: $\mathcal{S} \times \mathcal{E} \rightarrow \mathcal{S}$ transition function from state to state by an event
- $\mathbf{R}$: $\mathcal{S} \times \mathcal{E}_0 \times \mathcal{S} \rightarrow \mathbb{R}$ rewards for state transitions by an event
- F : $\mathcal{E} \times \mathbb{N} \rightarrow [0, 1]$ probabilities for events depending on time T

Problem:

π^* : $\mathcal{S} \times T_1 \times ... \times T_{\mathcal{E}} \rightarrow \mathcal{A}$ with $T_i = \mathbb{N}$ optimal policy with best rewards!

GSMDP Formulation

A GSMDP is given by a 7-tuple $(\mathcal{S}, \mathcal{A}, \mathcal{E}, \mathbf{C}, \mathbf{P}, \mathbf{R}, F)$ with

- $\mathcal{S}$ a set of states
- $\mathcal{A}$ a set of actions
- $\mathcal{E}$ a set of events (decreasing priority)
- $\mathbf{C}$: $\mathcal{S} \times \mathcal{A} \times \mathcal{E} \rightarrow \{0, 1\}$ mapping for (in)active events
- $\mathbf{P}$: $\mathcal{S} \times \mathcal{E} \rightarrow \mathcal{S}$ transition function from state to state by an event
- $\mathbf{R}$: $\mathcal{S} \times \mathcal{E}_0 \times \mathcal{S} \rightarrow \mathbb{R}$ rewards for state transitions by an event
- F : $\mathcal{E} \times \mathbb{N} \rightarrow [0, 1]$ probabilities for events depending on time T

Problem:

π^* : $\mathcal{S} \times T_1 \times ... \times T_{\mathcal{E}} \rightarrow \mathcal{A}$ with $T_i = \mathbb{N}$ optimal policy with best rewards!

GSMDP Formulation

A GSMDP is given by a 7-tuple $(\mathcal{S}, \mathcal{A}, \mathcal{E}, \mathbf{C}, \mathbf{P}, \mathbf{R}, F)$ with

- $\mathcal{S}$ a set of states
- $\mathcal{A}$ a set of actions
- $\mathcal{E}$ a set of events (decreasing priority)
- $\mathbf{C}$: $\mathcal{S} \times \mathcal{A} \times \mathcal{E} \to \{0, 1\}$ mapping for (in)active events
- $\mathbf{P}$: $\mathcal{S} \times \mathcal{E} \to \mathcal{S}$ transition function from state to state by an event
- $\mathbf{R}$: $\mathcal{S} \times \mathcal{E}_0 \times \mathcal{S} \to \mathbb{R}$ rewards for state transitions by an event
- F : $\mathcal{E} \times \mathbb{N} \to [0, 1]$ probabilities for events depending on time T

Problem:

π^* : $\mathcal{S} \times T_1 \times ... \times T_{\mathcal{E}} \to \mathcal{A}$ with $T_i = \mathbb{N}$ optimal policy with best rewards!

Special subclass of a GSMDP

A GSMDP is given by a 7-tuple $(\mathcal{S}, \mathcal{A}, \mathcal{E}, \mathbf{C}, \mathbf{P}, \mathbf{R}, F)$ with

- $\mathcal{S}$ a set of states
- $\mathcal{A}$ a set of actions
- $\mathcal{E}$ a set of events (decreasing priority)
- $\mathbf{C}$: $\mathcal{S} \times \mathcal{A} \times \mathcal{E} \to \{0, 1\}$ mapping for (in)active events
- $\mathbf{P}$: $\mathcal{S} \times \mathcal{E} \to \mathcal{S}$ transition function from state to state by an event
- $\mathbf{R}$: $\mathcal{S} \times \mathcal{E}_0 \times \mathcal{S} \to \mathbb{R}$ rewards for state transitions by an event
- F : $\mathcal{E} \times \mathbb{N} \to [0, 1]$ probabilities for events depending on time T

$\to$ If one or more events are triggered, all are reseted next period!

Problem:

π^* : $\mathcal{S} \to \mathcal{A}$ optimal policy with best rewards!

Solution approaches

Exact and approximative proceedings

Transform the GSMDP to a MDP with similar behaviour

What is the problem?

- Infinite state space for $\mathcal{S}_{MDP}$!
- An exponential number of path or event combinations $\Gamma \in \mathbf{O}(2^{|\mathcal{E}|})$, for the computation of $\mathbf{P}_{MDP}, \mathbf{R}_{MDP}$!

$\Rightarrow$ Building a MDP has an exponential running time

Exact and approximative proceedings

Transform the GSMDP to a MDP with similar behaviour

What is the problem?

- Infinite state space for $\mathcal{S}_{MDP}$!
- An exponential number of path or event combinations $\Gamma \in \mathbf{O}(2^{|\mathcal{E}|})$, for the computation of $\mathbf{P}_{MDP}, \mathbf{R}_{MDP}$!

$\Rightarrow$ Building a MDP has an exponential running time

Exact and approximative proceedings

Transform the GSMDP to a MDP with similar behaviour

What is the problem?

- Infinite state space for $\mathcal{S}_{MDP}$!
- An exponential number of path or event combinations $\Gamma \in \mathbf{O}(2^{|\mathcal{E}|})$, for the computation of $\mathbf{P}_{MDP}, \mathbf{R}_{MDP}$!

$\Rightarrow$ Building a MDP has an exponential running time

Exact and approximative proceedings

Transform the GSMDP to a MDP with similar behaviour

What is the problem?

- Infinite state space for $\mathcal{S}_{MDP}$!
- An exponential number of path or event combinations $\Gamma \in \mathbf{O}(2^{|\mathcal{E}|})$, for the computation of $\mathbf{P}_{MDP}, \mathbf{R}_{MDP}$!

$\Rightarrow$ Building a MDP has an exponential running time

Exact proceeding - Limit for no event

t is the time step

$e \in \mathcal{E}$ an event

$e_0 \in \mathcal{E}_0 \backslash \mathcal{E}$ means no event is triggered

$\mathbb{P}(e|t)$ the distribution function

- Steps with no triggering events: $\mathbb{P}(\mathbf{e_0}|\mathbf{t}) \searrow \mathbf{0}$
  - $\mathbb{P}(e|t-1) \leq \mathbb{P}(e|t)$
  - $\lim_{t\to\infty} \mathbb{P}(e|t) = 1$
  - $\mathbb{P}(e_0|t) := \prod_{e=1}^{|\mathcal{E}|}(1 - \mathbb{P}(e|t))$
  - $\lim_{t\to\infty} \mathbb{P}(e_0|t) = 0$
  - $\forall \varepsilon > 0 \; \exists t_0 \in \mathbb{N} \; \forall t \geq t_0 : \mathbb{P}(e_0|t) < \varepsilon$
- So $t \in (t_0, \infty)$ is negligible

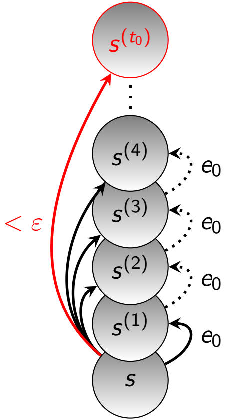


Figure: Cut of pseudo-states with an irrelevant probability

Exact proceeding - Limit for no event

t is the time step

$e \in \mathcal{E}$ an event

$e_0 \in \mathcal{E}_0 \backslash \mathcal{E}$ means no event is triggered

$\mathbb{P}(e|t)$ the distribution function

- Steps with no triggering events: $\mathbb{P}(\mathbf{e_0}|\mathbf{t}) \searrow \mathbf{0}$
  - $\mathbb{P}(e|t-1) \leq \mathbb{P}(e|t)$
  - $\lim_{t\to\infty} \mathbb{P}(e|t) = 1$
  - $\mathbb{P}(e_0|t) := \prod_{e=1}^{|\mathcal{E}|}(1-\mathbb{P}(e|t))$
  - $\lim_{t\to\infty} \mathbb{P}(e_0|t) = 0$
  - $\forall \varepsilon > 0\ \exists t_0 \in \mathbb{N}\ \forall t \geq t_0 : \mathbb{P}(e_0|t) < \varepsilon$
- So $t \in (t_0, \infty)$ is negligible

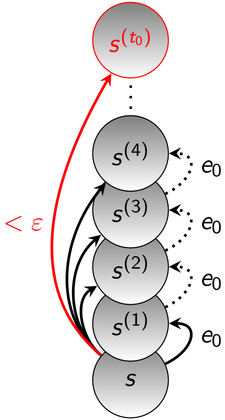


Figure: Cut of pseudo-states with an irrelevant probability

Exact proceeding - Limit for no event

t is the time step

$e \in \mathcal{E}$ an event

$e_0 \in \mathcal{E}_0 \backslash \mathcal{E}$ means no event is triggered

$\mathbb{P}(e|t)$ the distribution function

- Steps with no triggering events: $\mathbb{P}(\mathbf{e_0}|\mathbf{t}) \searrow \mathbf{0}$
  - $\mathbb{P}(e|t-1) \leq \mathbb{P}(e|t)$
  - $\lim_{t\to\infty} \mathbb{P}(e|t) = 1$
  - $\mathbb{P}(e_0|t) := \prod_{e=1}^{|\mathcal{E}|}(1-\mathbb{P}(e|t))$
  - $\lim_{t\to\infty} \mathbb{P}(e_0|t) = 0$
  - $\forall \varepsilon > 0\ \exists t_0 \in \mathbb{N}\ \forall t \geq t_0 : \mathbb{P}(e_0|t) < \varepsilon$
- So $t \in (t_0, \infty)$ is negligible

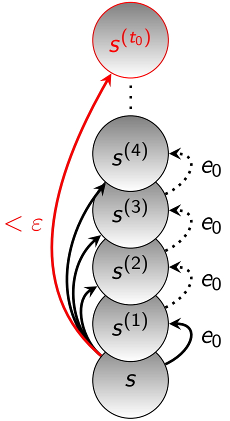


Figure: Cut of pseudo-states with an irrelevant probability

Exact proceeding - Limit for no event

t is the time step

$e \in \mathcal{E}$ an event

$e_0 \in \mathcal{E}_0 \backslash \mathcal{E}$ means no event is triggered

$\mathbb{P}(e|t)$ the distribution function

- Steps with no triggering events: $\mathbb{P}(\mathbf{e_0}|\mathbf{t}) \searrow \mathbf{0}$
  - $\mathbb{P}(e|t-1) \leq \mathbb{P}(e|t)$
  - $\lim_{t\to\infty} \mathbb{P}(e|t) = 1$
  - $\mathbb{P}(e_0|t) := \prod_{e=1}^{|\mathcal{E}|}(1 - \mathbb{P}(e|t))$
  - $\lim_{t\to\infty} \mathbb{P}(e_0|t) = 0$
  - $\forall \varepsilon > 0 \; \exists t_0 \in \mathbb{N} \; \forall t \geq t_0 : \mathbb{P}(e_0|t) < \varepsilon$
- So $t \in (t_0, \infty)$ is negligible

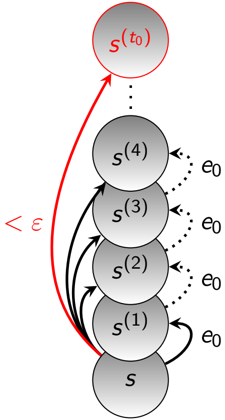


Figure: Cut of pseudo-states with an irrelevant probability

Exact proceeding - Limit for no event

t is the time step

$e \in \mathcal{E}$ an event

$e_0 \in \mathcal{E}_0 \backslash \mathcal{E}$ means no event is triggered

$\mathbb{P}(e|t)$ the distribution function

- Steps with no triggering events: $\mathbb{P}(\mathbf{e_0}|\mathbf{t}) \searrow \mathbf{0}$
  - $\mathbb{P}(e|t-1) \leq \mathbb{P}(e|t)$
  - $\lim_{t\to\infty} \mathbb{P}(e|t) = 1$
  - $\mathbb{P}(e_0|t) := \prod_{e=1}^{|\mathcal{E}|}(1 - \mathbb{P}(e|t))$
  - $\lim_{t\to\infty} \mathbb{P}(e_0|t) = 0$
  - $\forall \varepsilon > 0 \; \exists t_0 \in \mathbb{N} \; \forall t \geq t_0 : \mathbb{P}(e_0|t) < \varepsilon$
- So $t \in (t_0, \infty)$ is negligible

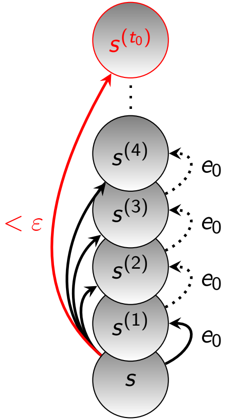


Figure: Cut of pseudo-states with an irrelevant probability

Exact proceeding - Limit for no event

t is the time step

$e \in \mathcal{E}$ an event

$e_0 \in \mathcal{E}_0 \backslash \mathcal{E}$ means no event is triggered

$\mathbb{P}(e|t)$ the distribution function

- Steps with no triggering events: $\mathbb{P}(\mathbf{e_0}|\mathbf{t}) \searrow \mathbf{0}$
  - $\mathbb{P}(e|t-1) \leq \mathbb{P}(e|t)$
  - $\lim_{t\to\infty} \mathbb{P}(e|t) = 1$
  - $\mathbb{P}(e_0|t) := \prod_{e=1}^{|\mathcal{E}|}(1 - \mathbb{P}(e|t))$
  - $\lim_{t\to\infty} \mathbb{P}(e_0|t) = 0$
  - $\forall \varepsilon > 0 \; \exists t_0 \in \mathbb{N} \; \forall t \geq t_0 : \mathbb{P}(e_0|t) < \varepsilon$
- So $t \in (t_0, \infty)$ is negligible

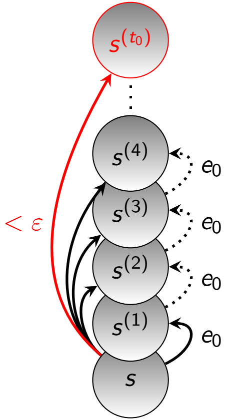


Figure: Cut of pseudo-states with an irrelevant probability

Exact proceeding - Influence of events

- **Zero-steps**
  - Events $e \in \mathcal{E}$ have a chance to be triggered
  - Build $\Gamma' \subseteq \Gamma$ with all paths for evaluation
  - $\gamma =< s_t, x_1, ..., x_{|\mathcal{E}|} >\in \Gamma'$ with $x_i \in \{e_i, \overline{e_i}\}$ is well defined
  - $\mathbf{C}(s, a, e)$ deactivates some events after passing some states!

γ_1 : e_1, e_2, e_3, e_4

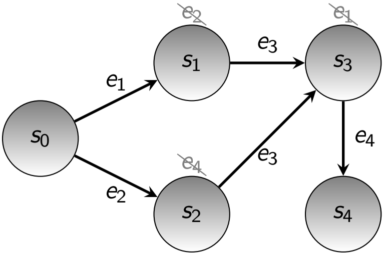


Figure: Different events lead to seperate paths

Exact proceeding - Influence of events

- **Zero-steps**
  - Events $e \in \mathcal{E}$ have a chance to be triggered
  - Build $\Gamma' \subseteq \Gamma$ with all paths for evaluation
  - $\gamma =< s_t, x_1, ..., x_{|\mathcal{E}|} >\in \Gamma'$ with $x_i \in \{e_i, \overline{e_i}\}$ is well defined
  - $\mathbf{C}(s, a, e)$ deactivates some events after passing some states!

γ_1 : e_1 e_2 e_3 e_4

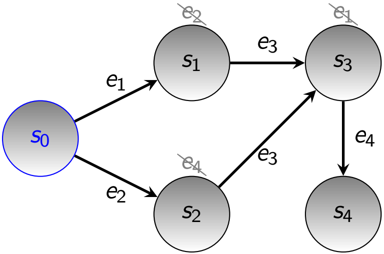


Figure: Different events lead to seperate paths

Exact proceeding - Influence of events

- **Zero-steps**
  - Events $e \in \mathcal{E}$ have a chance to be triggered
  - Build $\Gamma' \subseteq \Gamma$ with all paths for evaluation
  - $\gamma =< s_t, x_1, ..., x_{|\mathcal{E}|} >\in \Gamma'$ with $x_i \in \{e_i, \overline{e_i}\}$ is well defined
  - $\mathbf{C}(s, a, e)$ deactivates some events after passing some states!

γ_1 : e_1 e_2 e_3 e_4

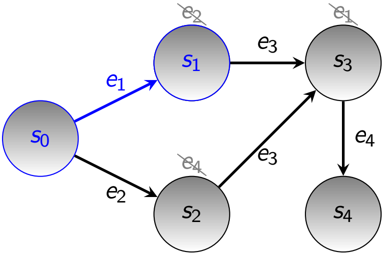


Figure: Different events lead to seperate paths

Exact proceeding - Influence of events

- **Zero-steps**
  - Events $e \in \mathcal{E}$ have a chance to be triggered
  - Build $\Gamma' \subseteq \Gamma$ with all paths for evaluation
  - $\gamma =< s_t, x_1, ..., x_{|\mathcal{E}|} >\in \Gamma'$ with $x_i \in \{e_i, \overline{e_i}\}$ is well defined
  - $\mathbf{C}(s, a, e)$ deactivates some events after passing some states!

γ_1 : e_1 e_2 e_3 e_4

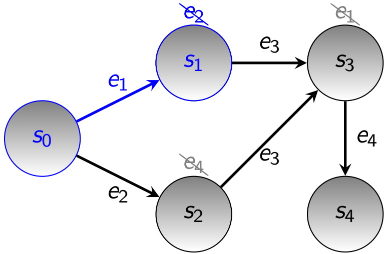


Figure: Different events lead to seperate paths

Exact proceeding - Influence of events

- **Zero-steps**
  - Events $e \in \mathcal{E}$ have a chance to be triggered
  - Build $\Gamma' \subseteq \Gamma$ with all paths for evaluation
  - $\gamma =< s_t, x_1, ..., x_{|\mathcal{E}|} >\in \Gamma'$ with $x_i \in \{e_i, \overline{e_i}\}$ is well defined
  - $\mathbf{C}(s, a, e)$ deactivates some events after passing some states!

γ_1 : e_1 e_2 e_3 e_4

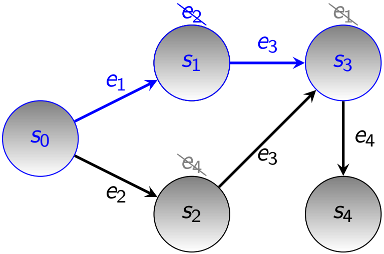


Figure: Different events lead to seperate paths

Exact proceeding - Influence of events

- **Zero-steps**
  - Events $e \in \mathcal{E}$ have a chance to be triggered
  - Build $\Gamma' \subseteq \Gamma$ with all paths for evaluation
  - $\gamma =< s_t, x_1, ..., x_{|\mathcal{E}|} >\in \Gamma'$ with $x_i \in \{e_i, \overline{e_i}\}$ is well defined
  - $\mathbf{C}(s, a, e)$ deactivates some events after passing some states!

γ_1 : e_1, e_2, e_3, e_4

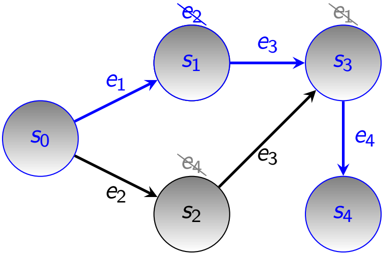


Figure: Different events lead to seperate paths

Exact proceeding - Influence of events

- **Zero-steps**
  - Events $e \in \mathcal{E}$ have a chance to be triggered
  - Build $\Gamma' \subseteq \Gamma$ with all paths for evaluation
  - $\gamma =< s_t, x_1, ..., x_{|\mathcal{E}|} > \in \Gamma'$ with $x_i \in \{e_i, \overline{e_i}\}$ is well defined
  - $\mathbf{C}(s, a, e)$ deactivates some events after passing some states!

γ_2 : $\overline{e_1}$, e_2, e_3, e_4

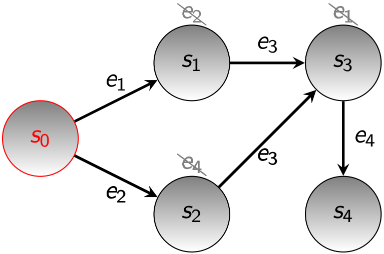


Figure: Different events lead to seperate paths

Exact proceeding - Influence of events

- **Zero-steps**
  - Events $e \in \mathcal{E}$ have a chance to be triggered
  - Build $\Gamma' \subseteq \Gamma$ with all paths for evaluation
  - $\gamma =< s_t, x_1, ..., x_{|\mathcal{E}|} >\in \Gamma'$ with $x_i \in \{e_i, \overline{e_i}\}$ is well defined
  - $\mathbf{C}(s, a, e)$ deactivates some events after passing some states!

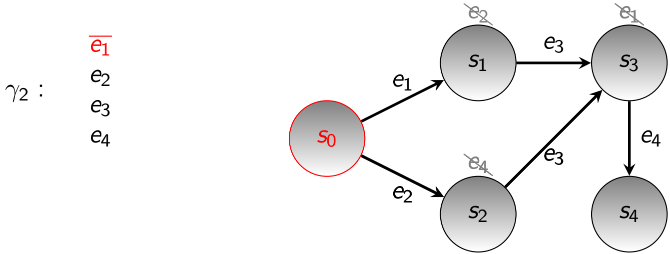


Figure: Different events lead to seperate paths

Exact proceeding - Influence of events

- **Zero-steps**
  - Events $e \in \mathcal{E}$ have a chance to be triggered
  - Build $\Gamma' \subseteq \Gamma$ with all paths for evaluation
  - $\gamma =< s_t, x_1, ..., x_{|\mathcal{E}|} > \in \Gamma'$ with $x_i \in \{e_i, \overline{e_i}\}$ is well defined
  - $\mathbf{C}(s, a, e)$ deactivates some events after passing some states!

γ_2 : $\overline{e_1}$ e_2 e_3 e_4

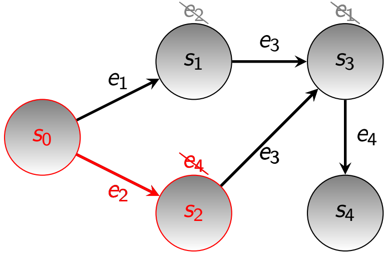


Figure: Different events lead to seperate paths

Exact proceeding - Influence of events

- **Zero-steps**
  - Events $e \in \mathcal{E}$ have a chance to be triggered
  - Build $\Gamma' \subseteq \Gamma$ with all paths for evaluation
  - $\gamma =< s_t, x_1, ..., x_{|\mathcal{E}|} >\in \Gamma'$ with $x_i \in \{e_i, \overline{e_i}\}$ is well defined
  - $\mathbf{C}(s, a, e)$ deactivates some events after passing some states!

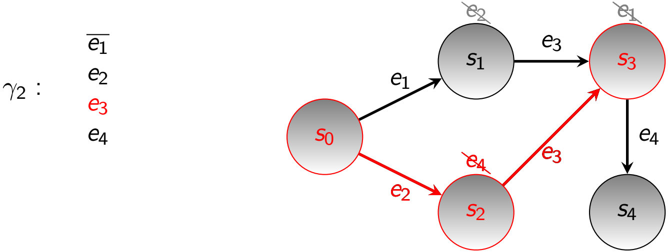


Figure: Different events lead to seperate paths

Exact proceeding - Influence of events

- **Zero-steps**
  - Events $e \in \mathcal{E}$ have a chance to be triggered
  - Build $\Gamma' \subseteq \Gamma$ with all paths for evaluation
  - $\gamma =< s_t, x_1, ..., x_{|\mathcal{E}|} >\in \Gamma'$ with $x_i \in \{e_i, \overline{e_i}\}$ is well defined
  - $\mathbf{C}(s, a, e)$ deactivates some events after passing some states!

γ_2 : $\overline{e_1}$ e_2 e_3 e_4

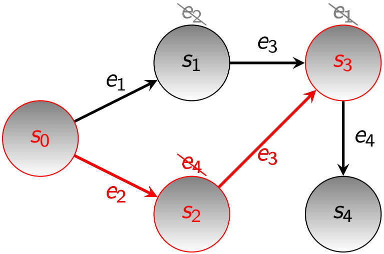


Figure: Different events lead to seperate paths

Exact proceeding - Building a MDP

- Build a MDP $(\mathbf{P}, \mathbf{R}, p)$
  - $\mathbf{P}$ and $\mathbf{R}$ are computed evaluating all zero-steps Γ
  - $\Leftarrow$ Sum up probabilities and rewards for paths γ with the same (s, a, s')
  - MDPs are solvable in polynomial time!
  - Approximative approaches build with an subset Γ'

Exact proceeding - Building a MDP

- Build a MDP $(\mathbf{P}, \mathbf{R}, p)$
  - $\mathbf{P}$ and $\mathbf{R}$ are computed evaluating all zero-steps Γ
  - $\Leftarrow$ Sum up probabilities and rewards for paths γ with the same (s, a, s')
  - MDPs are solvable in polynomial time!
  - Approximative approaches build with an subset Γ'

Exact proceeding - Building a MDP

- Build a MDP $(\mathbf{P}, \mathbf{R}, p)$
  - $\mathbf{P}$ and $\mathbf{R}$ are computed evaluating all zero-steps Γ
  - $\Leftarrow$ Sum up probabilities and rewards for paths γ with the same (s, a, s')
  - MDPs are solvable in polynomial time!
  - Approximative approaches build with an subset Γ'

Exact proceeding - Building a MDP

- Build a MDP $(\mathbf{P}, \mathbf{R}, p)$
  - $\mathbf{P}$ and $\mathbf{R}$ are computed evaluating all zero-steps Γ
  - $\Leftarrow$ Sum up probabilities and rewards for paths γ with the same (s, a, s')
  - MDPs are solvable in polynomial time!
  - Approximative approaches build with an subset Γ'

Γ-Method

The Γ-Method is more intuitive:

- Compute the event probabilities for triggering $\mathbb{P}(e_i)$
  Fix randomly a subset of $\mathcal{E}$ to 0 or 1 depending on $\mathbb{P}(e_i)$
  Unfixed events gives us a set of all paths $\Gamma' \subseteq \Gamma$
- Take $\gamma \in \Gamma'$ and check the path for regularity
- Evaluate all regular $\gamma \in \Gamma'$
  Combine results for $\mathbb{P}$ and $\mathbf{R}$ in a memory list

Γ-Method

The Γ-Method is more intuitive:

- Compute the event probabilities for triggering $\mathbb{P}(e_i)$

  Fix randomly a subset of $\mathcal{E}$ to 0 or 1 depending on $\mathbb{P}(e_i)$

  Unfixed events gives us a set of all paths $\Gamma' \subseteq \Gamma$
- Take $\gamma \in \Gamma'$ and check the path for regularity
- Evaluate all regular $\gamma \in \Gamma'$

  Combine results for $\mathbb{P}$ and $\mathbf{R}$ in a memory list

Γ-Method

The Γ-Method is more intuitive:

- Compute the event probabilities for triggering $\mathbb{P}(e_i)$

  Fix randomly a subset of $\mathcal{E}$ to 0 or 1 depending on $\mathbb{P}(e_i)$

  Unfixed events gives us a set of all paths $\Gamma' \subseteq \Gamma$
- Take $\gamma \in \Gamma'$ and check the path for regularity
- Evaluate all regular $\gamma \in \Gamma'$

  Combine results for $\mathbb{P}$ and $\mathbf{R}$ in a memory list

Γ-Method

The Γ-Method is more intuitive:

- Compute the event probabilities for triggering $\mathbb{P}(e_i)$
  Fix randomly a subset of $\mathcal{E}$ to 0 or 1 depending on $\mathbb{P}(e_i)$
  Unfixed events gives us a set of all paths $\Gamma' \subseteq \Gamma$
- Take $\gamma \in \Gamma'$ and check the path for regularity
- Evaluate all regular $\gamma \in \Gamma'$
  Combine results for $\mathbb{P}$ and $\mathbf{R}$ in a memory list

Γ-Method

Fixing Probabilities	e_1	e_2	e_3
$\mathbb{P}(e_i)$	0.8	0.6	0.4
$\mathbb{P}(e_i \leftarrow \{0,1\})$	0.6	0.2	0.2
$\mathbb{P}(e_i \leftarrow 0 \mid e_i \leftarrow \{0,1\})$	0.12	0.12	0.08
$\mathbb{P}(e_i \leftarrow 1 \mid e_i \leftarrow \{0,1\})$	0.48	0.08	0.12

$$\Gamma'_{example} = \{< 1,0,0 >,\ < 1,0,1 >,\ < 1,1,0 >,\ < 1,1,1 >\} \subsetneq \Gamma$$

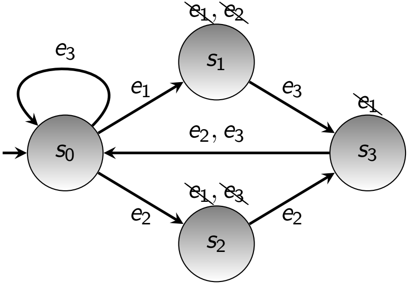


Figure: Check for regularity and evaluating $\gamma \in \Gamma'$

Γ-Method

Fixing Probabilities	e_1	e_2	e_3
$\mathbb{P}(e_i)$	0.8	0.6	0.4
$\mathbb{P}(e_i \leftarrow \{0,1\})$	0.6	0.2	0.2
$\mathbb{P}(e_i \leftarrow 0 \mid e_i \leftarrow \{0,1\})$	0.12	0.12	0.08
$\mathbb{P}(e_i \leftarrow 1 \mid e_i \leftarrow \{0,1\})$	0.48	0.08	0.12

$$\Gamma'_{example} = \{< 1, 0, 0 >, \ < 1, 0, 1 >, \ < 1, 1, 0 >, \ < 1, 1, 1 >\} \subsetneq \Gamma$$

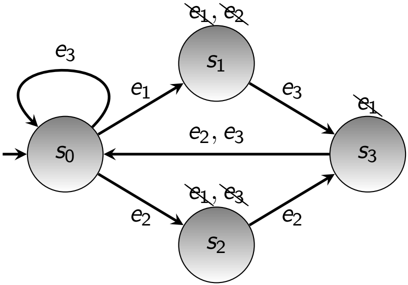


Figure: Check for regularity and evaluating $\gamma \in \Gamma'$

Γ-Method

Fixing Probabilities	e_1	e_2	e_3
$\mathbb{P}(e_i)$	0.8	0.6	0.4
$\mathbb{P}(e_i \leftarrow \{0,1\})$	0.6	0.2	0.2
$\mathbb{P}(e_i \leftarrow 0 \mid e_i \leftarrow \{0,1\})$	0.12	0.12	0.08
$\mathbb{P}(e_i \leftarrow 1 \mid e_i \leftarrow \{0,1\})$	0.48	0.08	0.12

$$\Gamma'_{example} = \{< 1, 0, 0 >,\ < 1, 0, 1 >,\ < 1, 1, 0 >,\ < 1, 1, 1 >\} \subsetneq \Gamma$$

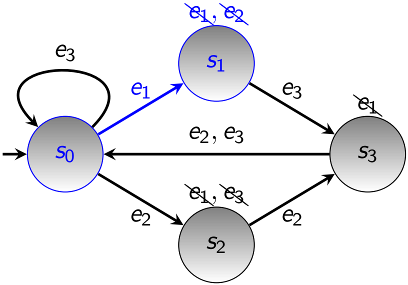


Figure: Check for regularity and evaluating $\gamma \in \Gamma'$

Γ-Method

Fixing Probabilities	e_1	e_2	e_3
$\mathbb{P}(e_i)$	0.8	0.6	0.4
$\mathbb{P}(e_i \leftarrow \{0,1\})$	0.6	0.2	0.2
$\mathbb{P}(e_i \leftarrow 0 \mid e_i \leftarrow \{0,1\})$	0.12	0.12	0.08
$\mathbb{P}(e_i \leftarrow 1 \mid e_i \leftarrow \{0,1\})$	0.48	0.08	0.12

$$\Gamma'_{example} = \{<1,0,0>,\ <1,0,1>,\ <1,1,0>,\ <1,1,1>\} \subsetneq \Gamma$$

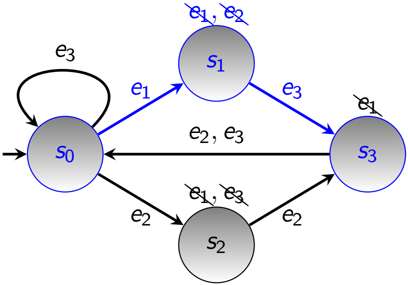


Figure: Check for regularity and evaluating $\gamma \in \Gamma'$

Γ-Method

Fixing Probabilities	e_1	e_2	e_3
$\mathbb{P}(e_i)$	0.8	0.6	0.4
$\mathbb{P}(e_i \leftarrow \{0,1\})$	0.6	0.2	0.2
$\mathbb{P}(e_i \leftarrow 0 \mid e_i \leftarrow \{0,1\})$	0.12	0.12	0.08
$\mathbb{P}(e_i \leftarrow 1 \mid e_i \leftarrow \{0,1\})$	0.48	0.08	0.12

$$\Gamma'_{example} = \{<1,0,0>,\ <1,0,1>,\ <1,1,0>,\ <1,1,1>\} \subsetneq \Gamma$$

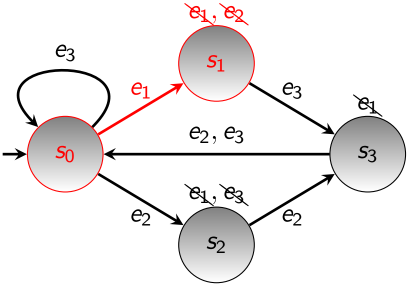


Figure: Check for regularity and evaluating $\gamma \in \Gamma'$

Γ-Method

Fixing Probabilities	e_1	e_2	e_3
$\mathbb{P}(e_i)$	0.8	0.6	0.4
$\mathbb{P}(e_i \leftarrow \{0,1\})$	0.6	0.2	0.2
$\mathbb{P}(e_i \leftarrow 0 \mid e_i \leftarrow \{0,1\})$	0.12	0.12	0.08
$\mathbb{P}(e_i \leftarrow 1 \mid e_i \leftarrow \{0,1\})$	0.48	0.08	0.12

$$\Gamma'_{example} = \{<1,0,0>,\ <1,0,1>,\ <1,1,0>,\ <1,1,1>\} \subsetneq \Gamma$$

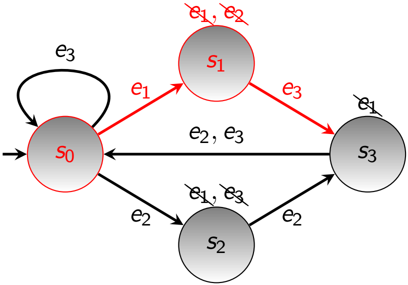


Figure: Check for regularity and evaluating $\gamma \in \Gamma'$

Γ-Method

Fixing Probabilities	e_1	e_2	e_3
$\mathbb{P}(e_i)$	0.8	0.6	0.4
$\mathbb{P}(e_i \leftarrow \{0,1\})$	0.6	0.2	0.2
$\mathbb{P}(e_i \leftarrow 0 \mid e_i \leftarrow \{0,1\})$	0.12	0.12	0.08
$\mathbb{P}(e_i \leftarrow 1 \mid e_i \leftarrow \{0,1\})$	0.48	0.08	0.12

$$\Gamma'_{example} = \{<1,0,0>,\ <1,0,1>,\ <1,1,0>,\ <1,1,1>\} \subsetneq \Gamma$$

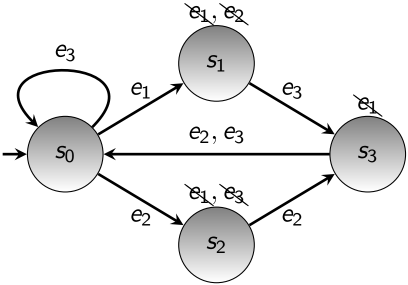


Figure: Check for regularity and evaluating $\gamma \in \Gamma'$

Γ-Method

Facts

- Exact for no fixing!
- Running time $\mathbf{\Theta}(\Omega \cdot |\mathcal{E}|^2 + |\mathcal{S}|)$, exact one with $\Omega = 2^{|\mathcal{E}|}$
- For a given Ω the algorithm fixes

$$\lceil |\mathcal{E}| - \log_2 \Omega \rceil$$

$\mathcal{E}$-Method

The $\mathcal{E}$-Method is more constructive and provident:

- Set an upper limit to the memory list Ω
- Evaluate every list element event by event (subpaths)
  Combine list elements with same future behavior (state, blockings,...)
  If the limit is exceeded: delete a random list element
- At the end there are no more than $\mathcal{S}$ list elements with expected rewards and probabilities

Facts

- Exact for no list limit!
- Running time $\Theta(\Omega^3 \cdot |\mathcal{E}| + |\mathcal{S}|)$, exact one with $\Omega = 2^{|\mathcal{E}|/2}$

$\mathcal{E}$-Method

The $\mathcal{E}$-Method is more constructive and provident:

- Set an upper limit to the memory list Ω
- Evaluate every list element event by event (subpaths)
  - Combine list elements with same future behavior (state, blockings,...)
  - If the limit is exceeded: delete a random list element
- At the end there are no more than $\mathcal{S}$ list elements with expected rewards and probabilities

Facts

- Exact for no list limit!
- Running time $\Theta(\Omega^3 \cdot |\mathcal{E}| + |\mathcal{S}|)$, exact one with $\Omega = 2^{|\mathcal{E}|/2}$

$\mathcal{E}$-Method

The $\mathcal{E}$-Method is more constructive and provident:

- Set an upper limit to the memory list Ω
- Evaluate every list element event by event (subpaths)
  Combine list elements with same future behavior (state, blockings,...)
  If the limit is exceeded: delete a random list element
- At the end there are no more than $\mathcal{S}$ list elements with expected rewards and probabilities

Facts

- Exact for no list limit!
- Running time $\mathbf{\Theta}(\Omega^3 \cdot |\mathcal{E}| + |\mathcal{S}|)$, exact one with $\Omega = 2^{|\mathcal{E}|/2}$

$\mathcal{E}$-Method

The $\mathcal{E}$-Method is more constructive and provident:

- Set an upper limit to the memory list Ω
- Evaluate every list element event by event (subpaths)
  Combine list elements with same future behavior (state, blockings,...)
  If the limit is exceeded: delete a random list element
- At the end there are no more than $\mathcal{S}$ list elements with expected rewards and probabilities

Facts

- Exact for no list limit!
- Running time $\Theta(\Omega^3 \cdot |\mathcal{E}| + |\mathcal{S}|)$, exact one with $\Omega = 2^{|\mathcal{E}|/2}$

$\mathcal{E}$-Method

The $\mathcal{E}$-Method is more constructive and provident:

- Set an upper limit to the memory list Ω
- Evaluate every list element event by event (subpaths)

  Combine list elements with same future behavior (state, blockings,...)

  If the limit is exceeded: delete a random list element
- At the end there are no more than $\mathcal{S}$ list elements with expected rewards and probabilities

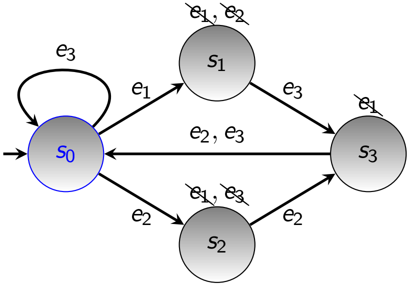
Facts

- Exact for no list limit!
- Running time $\mathbf{\Theta}(\Omega^3 \cdot |\mathcal{E}| + |\mathcal{S}|)$, exact one with $\Omega = 2^{|\mathcal{E}|/2}$

$\mathcal{E}$-Method

Fixing Probabilities	e_1	e_2	e_3
$\mathbb{P}(e_i)$	0.8	0.6	0.4

$$L_i = [\mathcal{S}, \mathcal{E}_{act}, \mathbb{P}, \mathbb{R}]$$



Initializing L

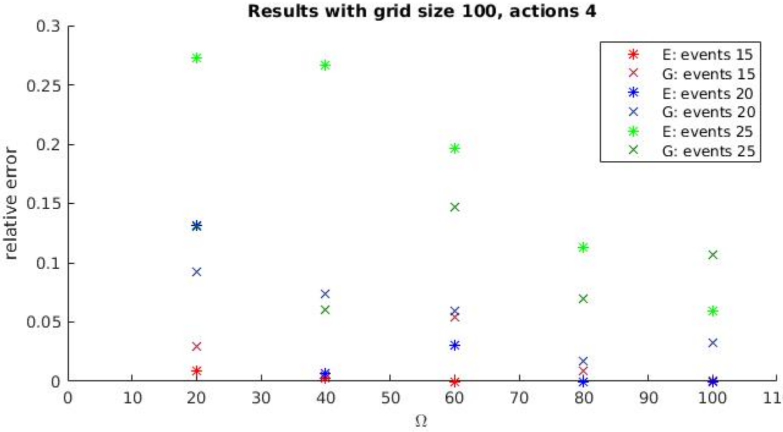
$[s_0, <e_1, e_2, e_3>, 1.00, 0]$

After Event e_3

$[s_3, < \quad >, 0.32, r_1 + r_3]$
$[s_1, < \quad >, 0.48, r_1]$
$[s_2, < \quad >, 0.12, r_2]$
$[s_0, < \quad >, 0.032, r_4]$
$[s_0, < \quad >, 0.048, r_0]$

Experimental results

Results - Relative Error



Results - Running Time

Table: Relative run times for $|\mathcal{S}| = 100$ and $|\mathcal{A}| = 4$

method and $\lvert\mathcal{E}\rvert$	$\Omega = 20$	$\Omega = 40$	$\Omega = 60$	$\Omega = 80$	$\Omega = 100$
$\mathcal{E}$-method, $\lvert\mathcal{E}\rvert = 15$	0.985	0.973	0.981	0.985	1.004
Γ-method, $\lvert\mathcal{E}\rvert = 15$	0.791	1.379	1.385	2.543	2.531
$\mathcal{E}$-method, $\lvert\mathcal{E}\rvert = 20$	0.840	0.953	0.971	0.988	0.992
Γ-method, $\lvert\mathcal{E}\rvert = 20$	0.368	0.689	0.689	1.364	1.365
$\mathcal{E}$-method, $\lvert\mathcal{E}\rvert = 25$	0.607	0.863	0.929	0.946	0.957
Γ-method, $\lvert\mathcal{E}\rvert = 25$	0.204	0.303	0.305	0.613	0.610

Conclusion and future work

Conclusion and future work

- $\mathcal{E}$-Method is almost better to approximate optimal solutions than Γ-Method
- $\mathcal{E}$-Method $\Theta(\Omega^3 \cdot |\mathcal{E}| + |\mathcal{S}|)$, exact one with $\Omega = 2^{|\mathcal{E}|/2}$
- Γ-Method $\Theta(\Omega \cdot |\mathcal{E}|^2 + |\mathcal{S}|)$, exact one with $\Omega = 2^{|\mathcal{E}|}$
- Expand the solutions to the more general problem (without special subclass)
- With a precalculation Δ the results and running times can be improved

$$\Omega^* \leq 2^{\Delta} \leq 2^{|\mathcal{E}_{act}|/2} < 2^{|\mathcal{E}|} = |\Gamma|$$

- Also the relative error can be estimated by Δ with no exact solution

Conclusion and future work

- $\mathcal{E}$-Method is almost better to approximate optimal solutions than Γ-Method
- $\mathcal{E}$-Method $\mathbf{\Theta}(\Omega^3 \cdot |\mathcal{E}| + |\mathcal{S}|)$, exact one with $\Omega = 2^{|\mathcal{E}|/2}$
- Γ-Method $\mathbf{\Theta}(\Omega \cdot |\mathcal{E}|^2 + |\mathcal{S}|)$, exact one with $\Omega = 2^{|\mathcal{E}|}$
- Expand the solutions to the more general problem (without special subclass)
- With a precalculation Δ the results and running times can be improved

$$\Omega^* \leq 2^{\Delta} \leq 2^{|\mathcal{E}_{act}|/2} < 2^{|\mathcal{E}|} = |\Gamma|$$

- Also the relative error can be estimated by Δ with no exact solution

Conclusion and future work

- $\mathcal{E}$-Method is almost better to approximate optimal solutions than Γ-Method
- $\mathcal{E}$-Method $\mathbf{\Theta}(\Omega^3 \cdot |\mathcal{E}| + |\mathcal{S}|)$, exact one with $\Omega = 2^{|\mathcal{E}|/2}$
- Γ-Method $\mathbf{\Theta}(\Omega \cdot |\mathcal{E}|^2 + |\mathcal{S}|)$, exact one with $\Omega = 2^{|\mathcal{E}|}$
- Expand the solutions to the more general problem (without special subclass)
- With a precalculation Δ the results and running times can be improved

$$\Omega^* \leq 2^{\Delta} \leq 2^{|\mathcal{E}_{act}|/2} < 2^{|\mathcal{E}|} = |\Gamma|$$

- Also the relative error can be estimated by Δ with no exact solution

Conclusion and future work

- $\mathcal{E}$-Method is almost better to approximate optimal solutions than Γ-Method
- $\mathcal{E}$-Method $\mathbf{\Theta}(\Omega^3 \cdot |\mathcal{E}| + |\mathcal{S}|)$, exact one with $\Omega = 2^{|\mathcal{E}|/2}$
- Γ-Method $\mathbf{\Theta}(\Omega \cdot |\mathcal{E}|^2 + |\mathcal{S}|)$, exact one with $\Omega = 2^{|\mathcal{E}|}$
- Expand the solutions to the more general problem (without special subclass)
- With a precalculation Δ the results and running times can be improved

$$\Omega^* \leq 2^{\Delta} \leq 2^{|\mathcal{E}_{act}|/2} < 2^{|\mathcal{E}|} = |\Gamma|$$

- Also the relative error can be estimated by Δ with no exact solution

Questions

The algorithms are able to compute randomized solutions (near to the exact one with more running times). Start solutions are relevant for many planing problems and need also human expertise. So the results can support planing management with initial computations.

Questions?