

Stopping Criteria for Value Iteration on Concurrent Stochastic Reachability and Safety Games

(LICS'25)

Marta Grobelna¹, Jan Křetinský^{2,1}, Maximilian Weininger^{3,4,1}

1) Technical University of Munich 2) Masaryk University 3) Ruhr-University Bochum 4) Institute of Science and Technology Austria

Highlights 2025

Saarbrücken



**Concurrent Stochastic
Games**

**Bounded Value
Iteration**

**Bloated End
Components**

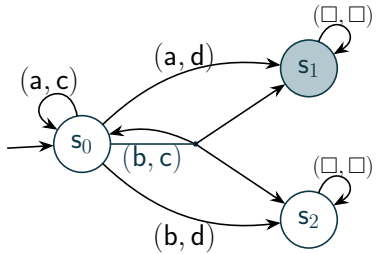
Deflating

Concurrent Stochastic Games

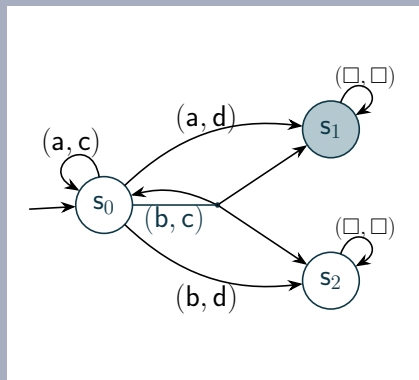
Bounded Value Iteration

Bloated End Components

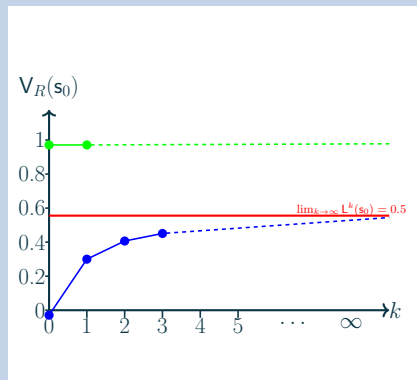
Deflating



Concurrent Stochastic Games



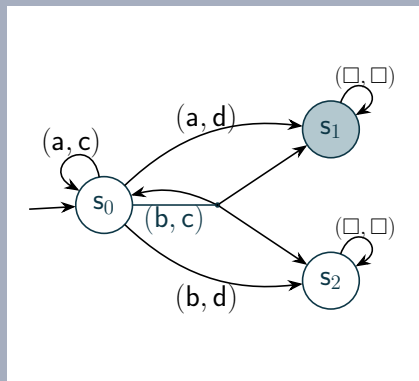
Bounded Value Iteration



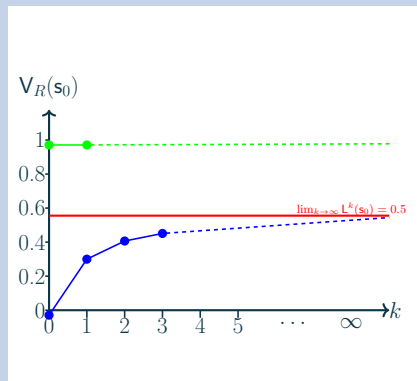
Bloated End Components

Deflating

Concurrent Stochastic Games



Bounded Value Iteration

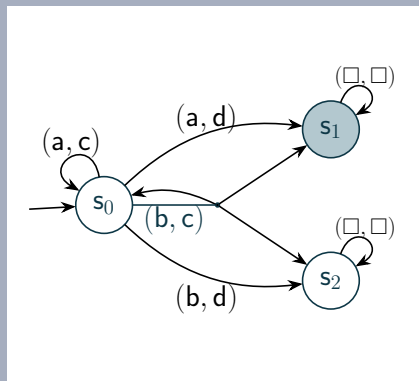


Bloated End Components

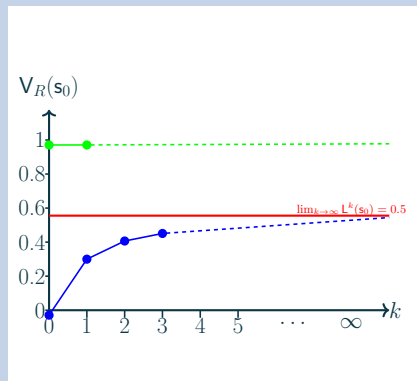
	c	d	
$\begin{pmatrix} 1 & 1 \end{pmatrix}$	a		
$\begin{pmatrix} \frac{1}{3} \cdot 1 + \frac{1}{3} & 0 \end{pmatrix}$	b		

Deflating

Concurrent Stochastic Games



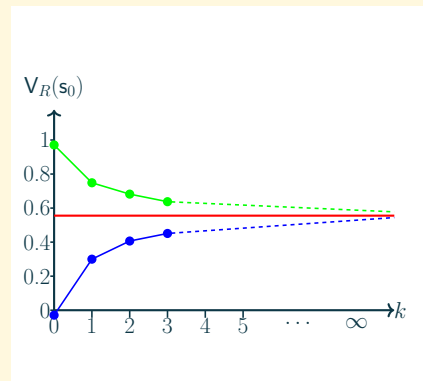
Bounded Value Iteration



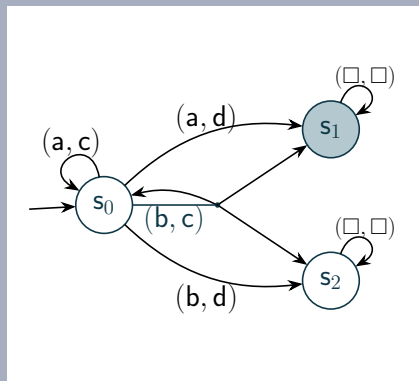
Bloated End Components

$$\begin{array}{cc}
 & \begin{array}{cc} c & d \end{array} \\
 \begin{pmatrix} 1 & 1 \\ \frac{1}{3} \cdot 1 + \frac{1}{3} & 0 \end{pmatrix} & \begin{array}{l} a \\ b \end{array}
 \end{array}$$

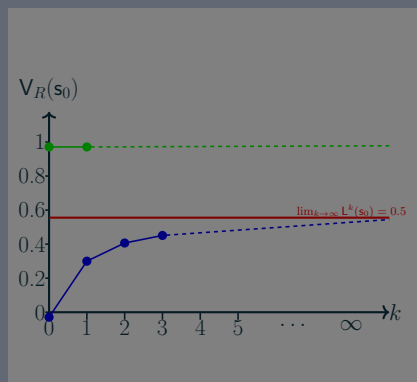
Deflating



Concurrent Stochastic Games



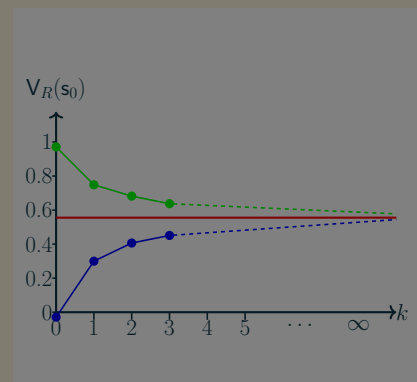
Bounded Value Iteration



Bloated End Components

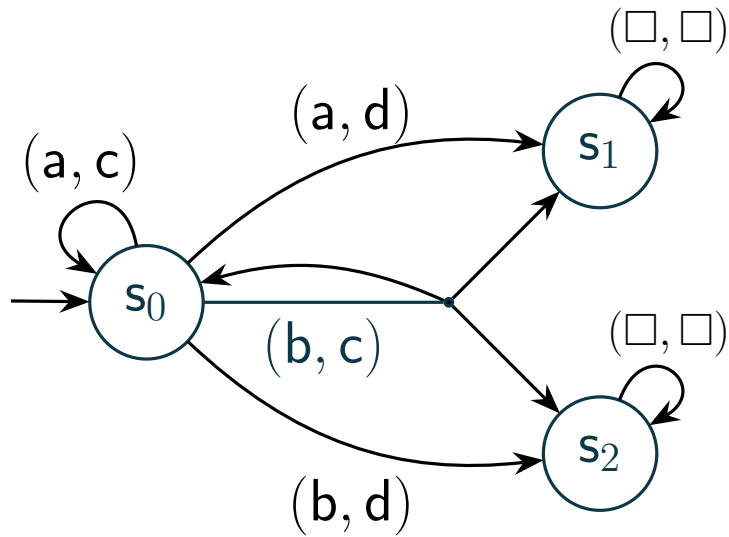
	c	d	
$\begin{pmatrix} 1 & 1 \end{pmatrix}$	a		
$\begin{pmatrix} \frac{1}{3} \cdot 1 + \frac{1}{3} & 0 \end{pmatrix}$	b		

Deflating

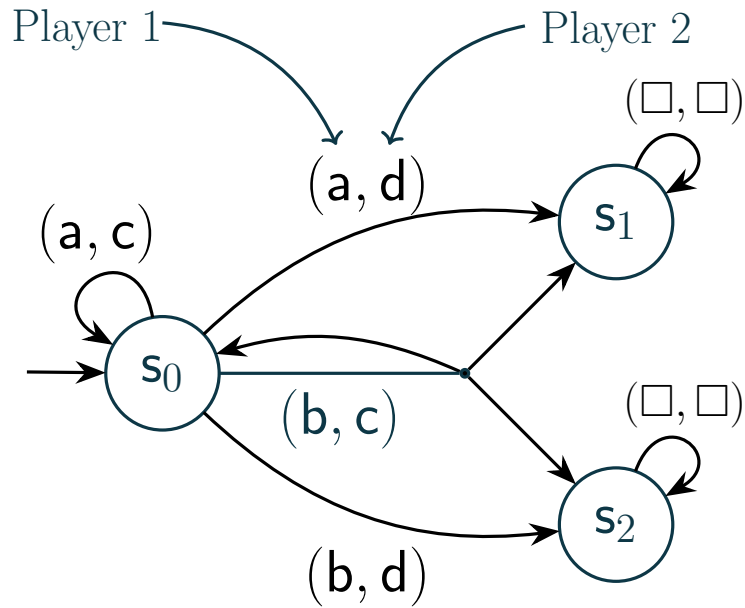


Concurrent Stochastic Reachability Games

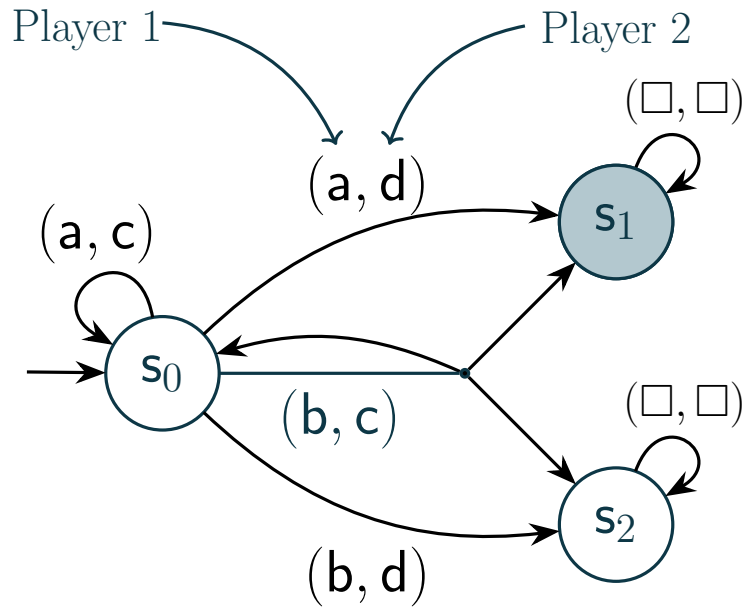
Concurrent Stochastic Reachability Games



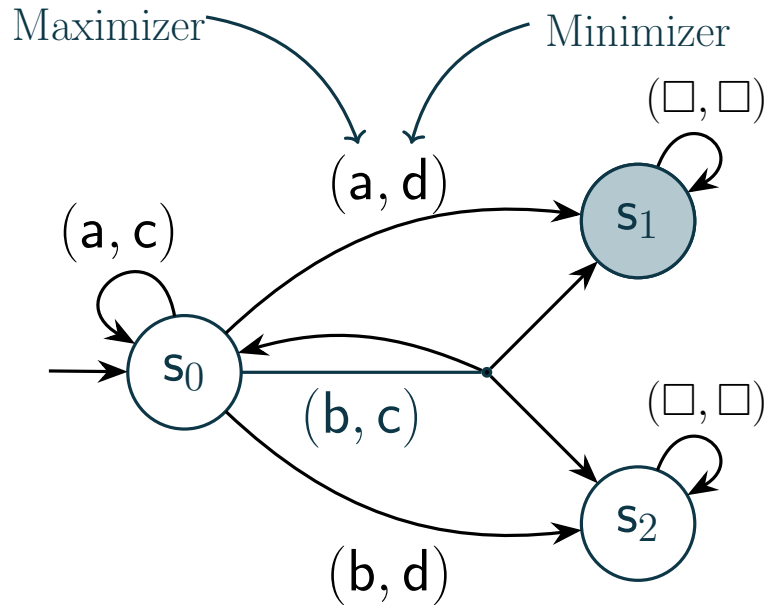
Concurrent Stochastic Reachability Games



Concurrent Stochastic Reachability Games

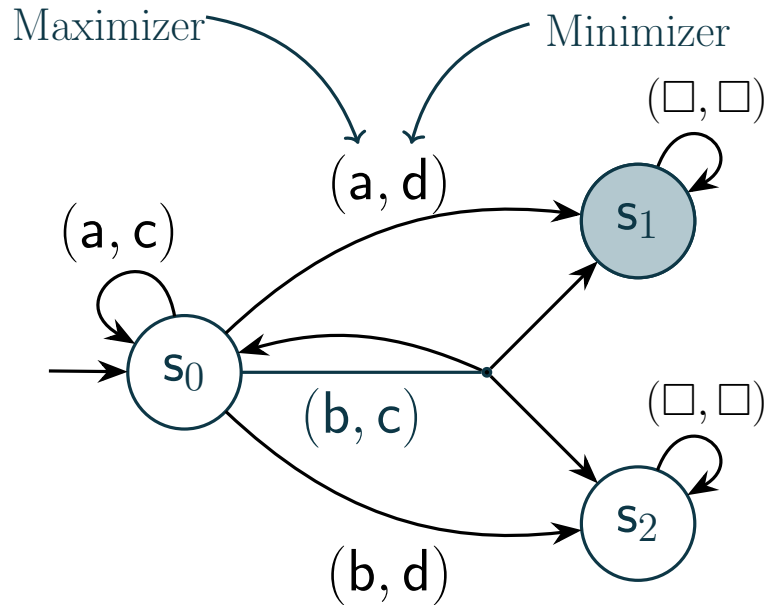


Concurrent Stochastic Reachability Games



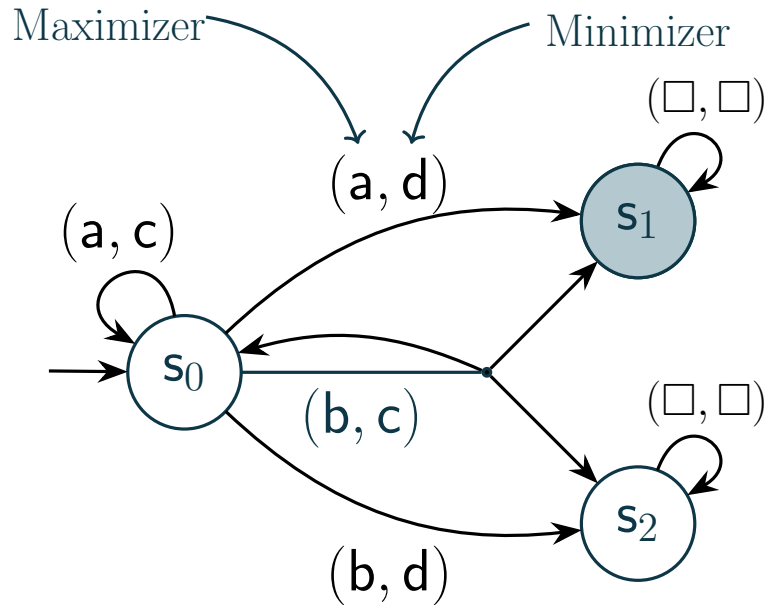
- **Maximizer** (Player 1) wants to **reach** s_1
- **Minimizer** (Player 2) wants to **avoid** s_1

Concurrent Stochastic Reachability Games



- **Maximizer** (Player 1) wants to **reach** s_1
- **Minimizer** (Player 2) wants to **avoid** s_1
- **Strategy**: Distribution over available actions

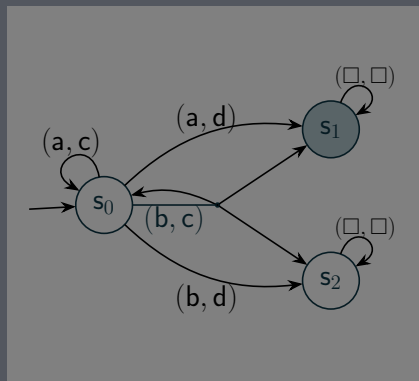
Concurrent Stochastic Reachability Games



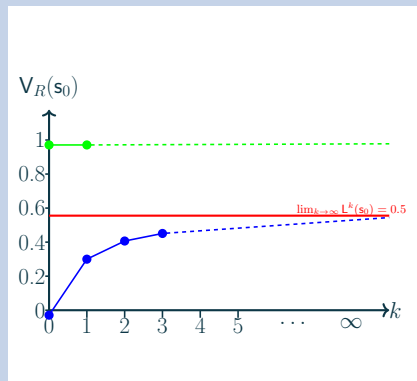
- **Maximizer** (Player 1) wants to **reach** s_1
- **Minimizer** (Player 2) wants to **avoid** s_1
- **Strategy**: Distribution over available actions

Value: Supremum probability of reaching s_1

Concurrent Stochastic Games



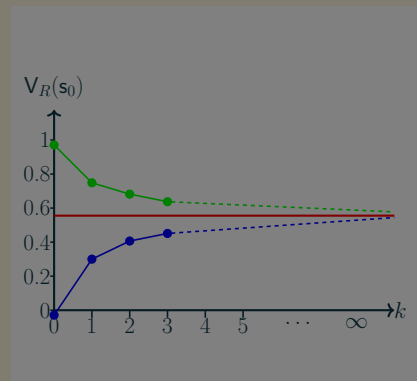
Bounded Value Iteration



Bloated End Components

$$\begin{array}{cc}
 & \begin{array}{cc} c & d \end{array} \\
 \begin{pmatrix} 1 & 1 \\ \frac{1}{3} \cdot 1 + \frac{1}{3} & 0 \end{pmatrix} & \begin{array}{l} a \\ b \end{array}
 \end{array}$$

Deflating



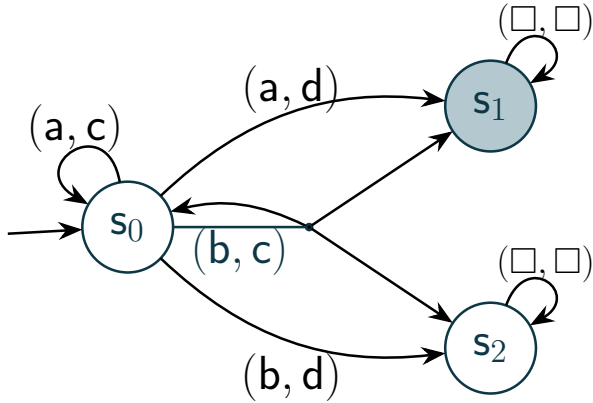
Bounded Value Iteration

Bounded Value Iteration

- **Initialize** with **lowest** possible value and **iteratively update** (e.g. [1])

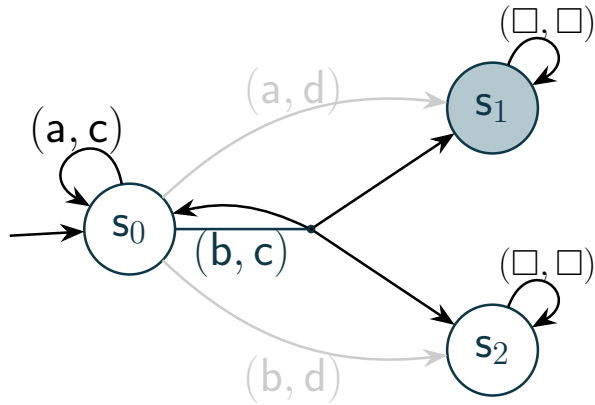
Bounded Value Iteration

- **Initialize** with **lowest** possible value and **iteratively update** (e.g. [1])



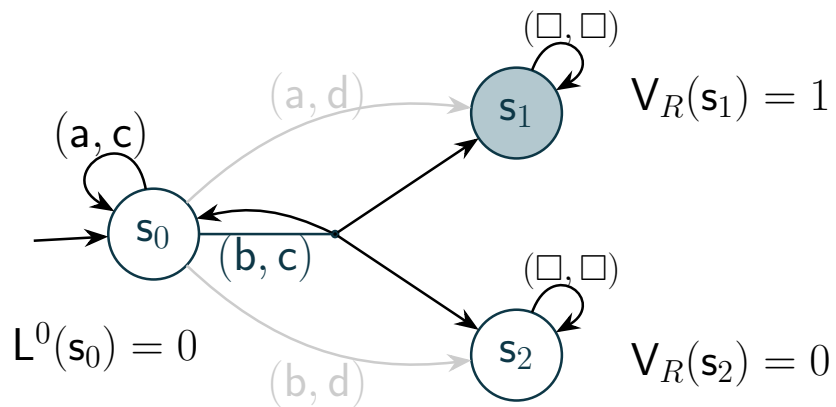
Bounded Value Iteration

- **Initialize** with **lowest** possible value and **iteratively update** (e.g. [1])



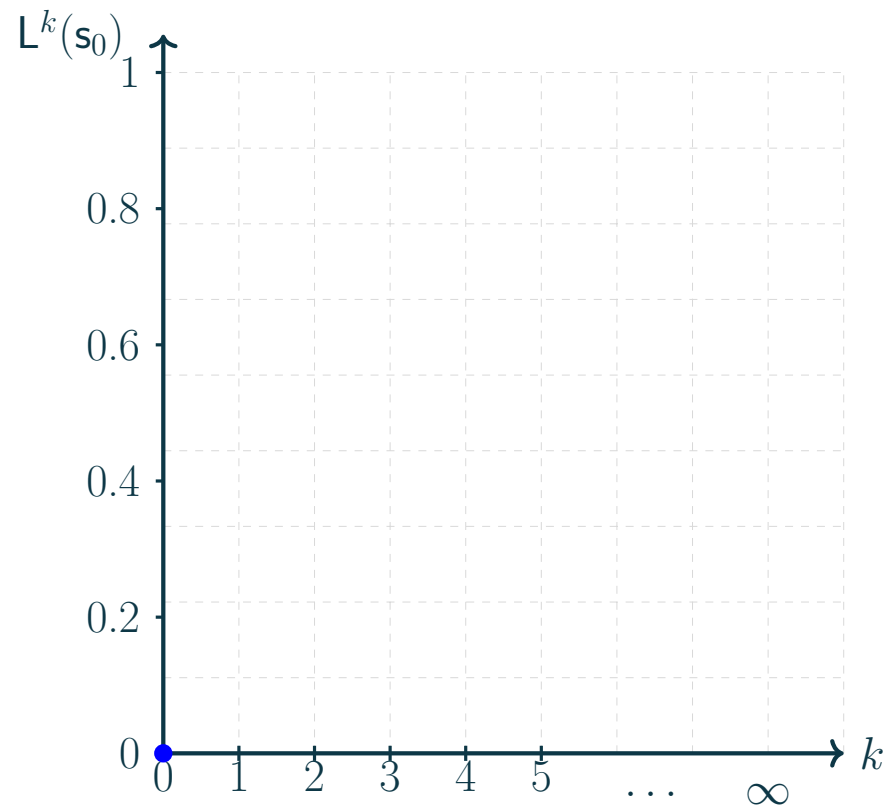
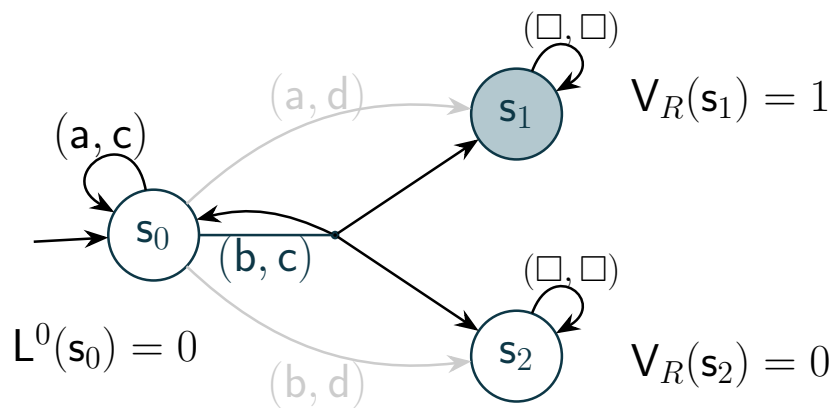
Bounded Value Iteration

- **Initialize** with **lowest** possible value and **iteratively update** (e.g. [1])



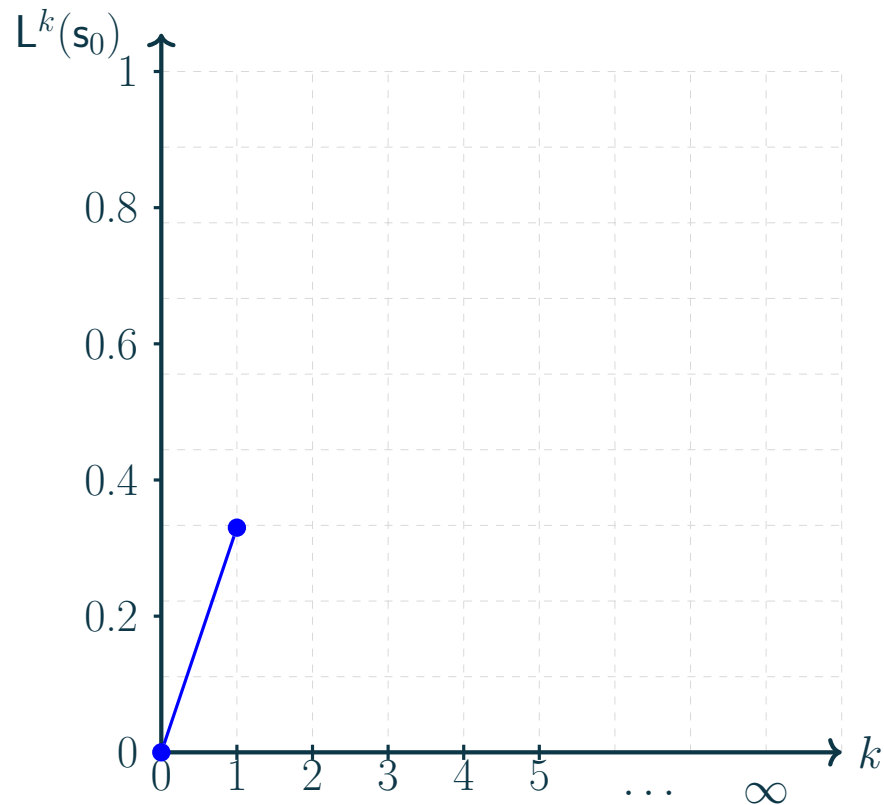
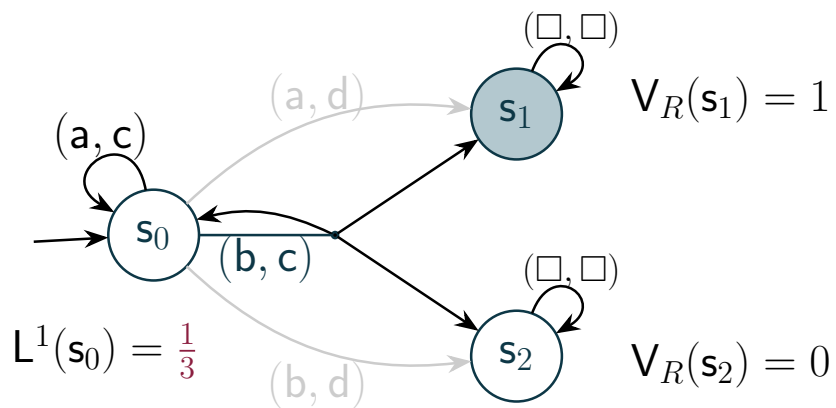
Bounded Value Iteration

- **Initialize** with **lowest** possible value and **iteratively update** (e.g. [1])



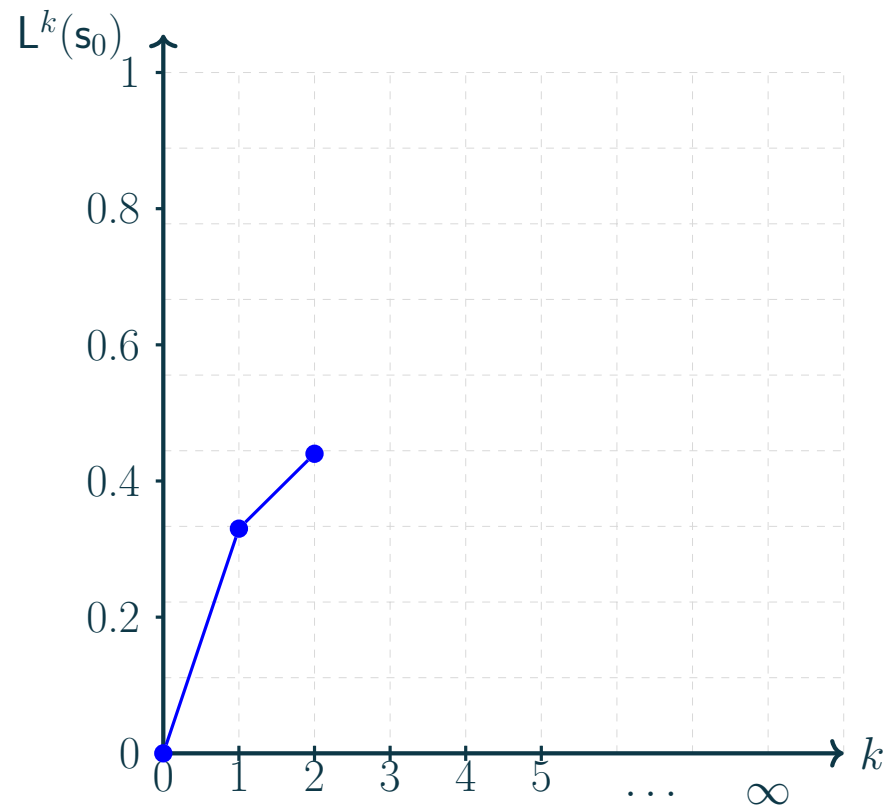
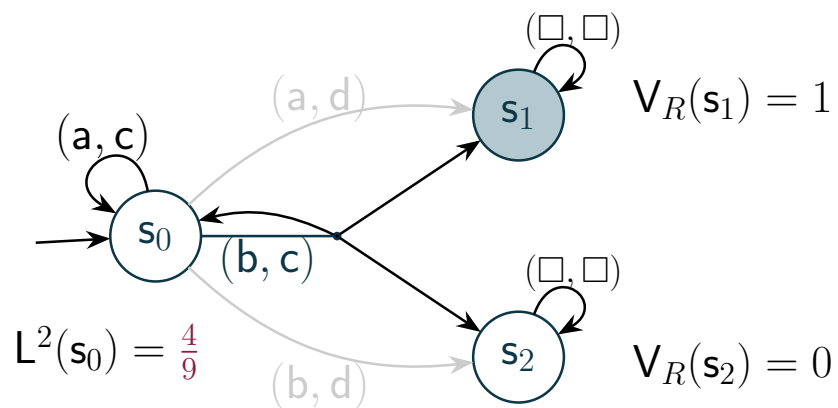
Bounded Value Iteration

- **Initialize** with **lowest** possible value and **iteratively update** (e.g. [1])



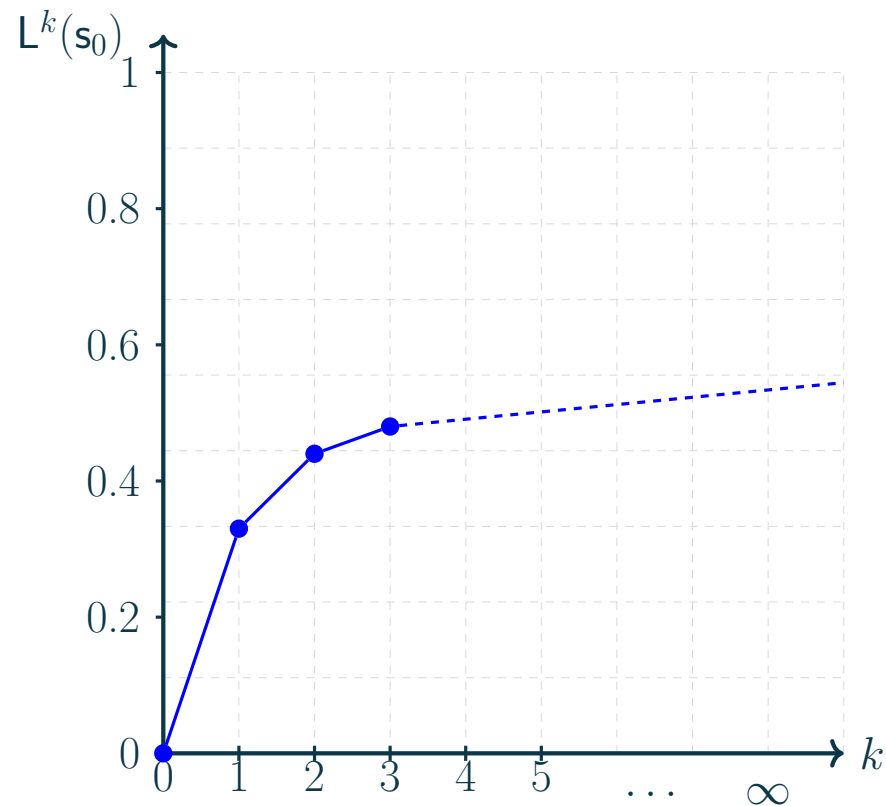
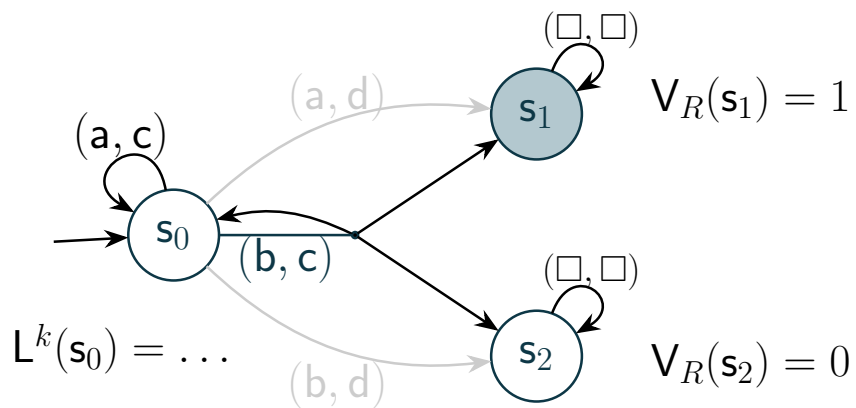
Bounded Value Iteration

- **Initialize** with **lowest** possible value and **iteratively update** (e.g. [1])



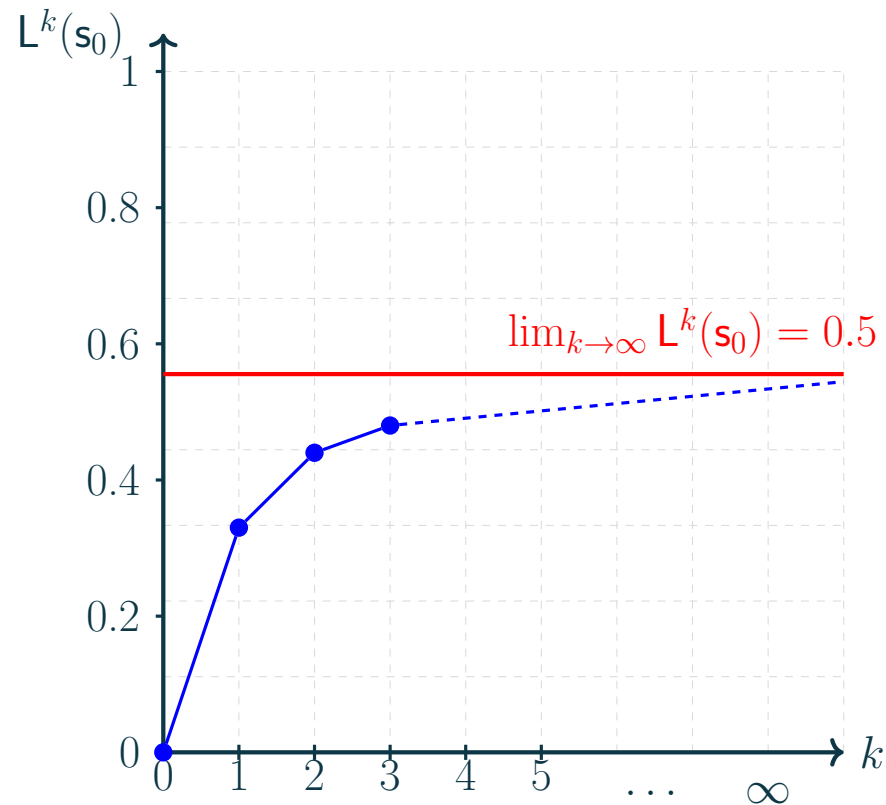
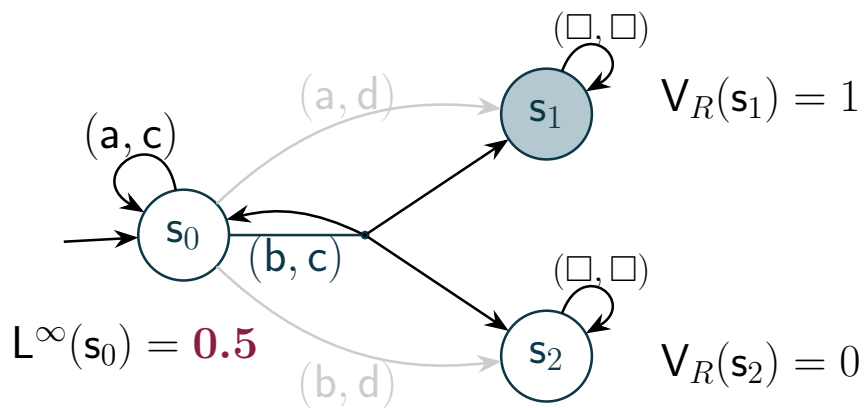
Bounded Value Iteration

- **Initialize** with **lowest** possible value and **iteratively update** (e.g. [1])



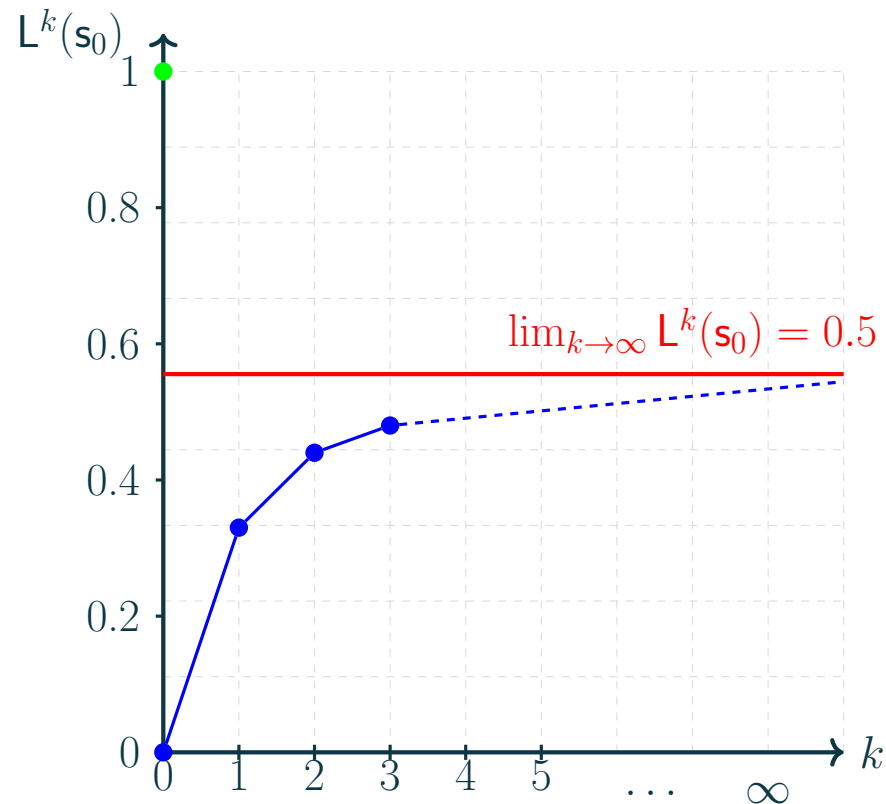
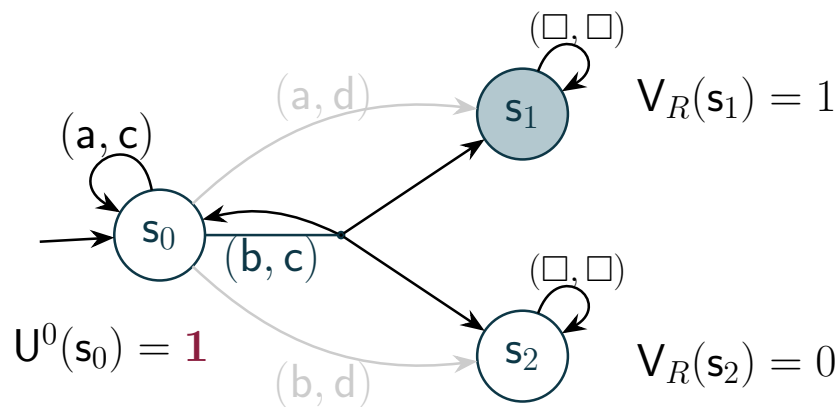
Bounded Value Iteration

- **Initialize** with **lowest** possible value and **iteratively update** (e.g. [1])



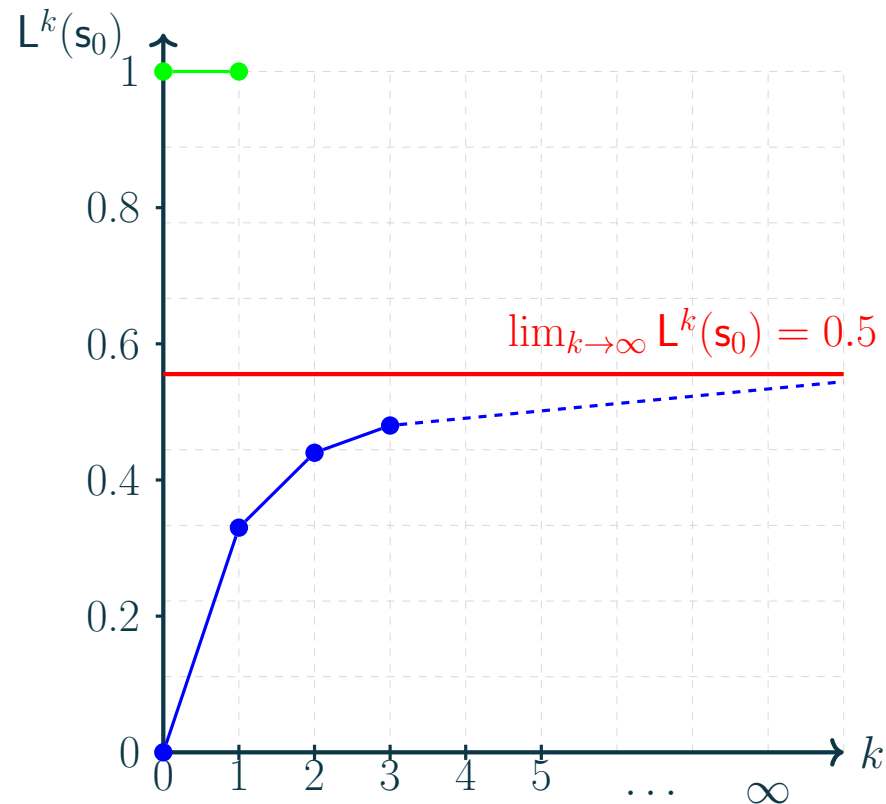
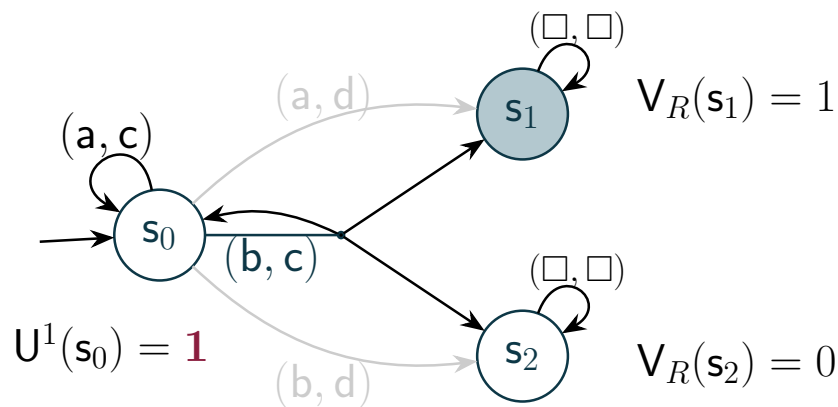
Bounded Value Iteration

- **Initialize** with **lowest** possible value and **iteratively update** (e.g. [1])
- **Initialize** with **highest** possible value and **iteratively update** (MDPs: [2], TSGs: [3])



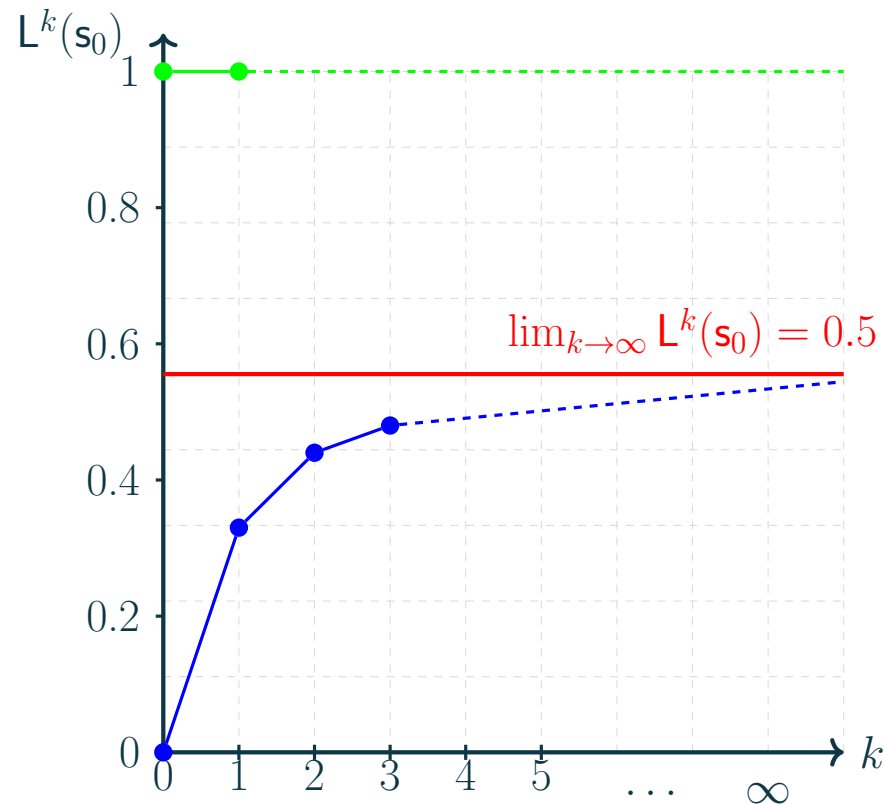
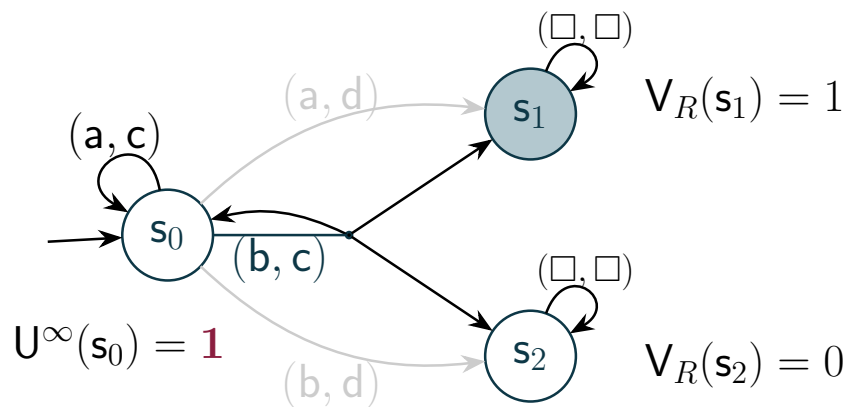
Bounded Value Iteration

- **Initialize** with **lowest** possible value and **iteratively update** (e.g. [1])
- **Initialize** with **highest** possible value and **iteratively update** (MDPs: [2], TSGs: [3])

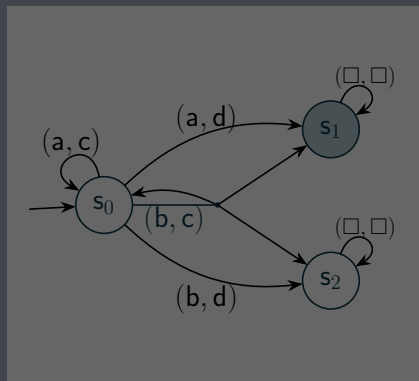


Bounded Value Iteration

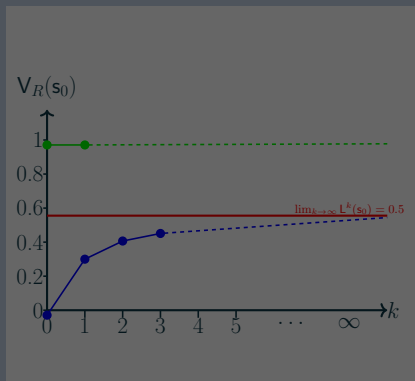
- **Initialize** with **lowest** possible value and **iteratively update** (e.g. [1])
- **Initialize** with **highest** possible value and **iteratively update** (MDPs: [2], TSGs: [3])
- **Why does it not converge?**



Concurrent Stochastic Games



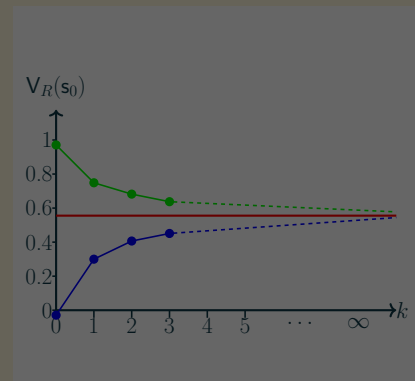
Bounded Value Iteration



Bloated End Components

	c	d	
a	1	1	
b	$\frac{1}{3} \cdot 1 + \frac{1}{3}$	0	

Deflating



Bloated End Components

Bloated End Components

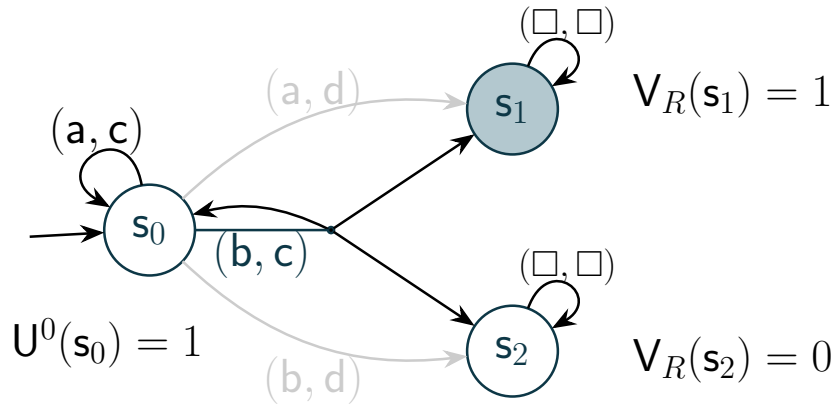
- **End component:** Set of states where staying for ever is possible

Bloated End Components

- **End component:** Set of states where staying for ever is possible
- **Bloated end component (BEC):** End component where staying pays off (TSGs: [3])

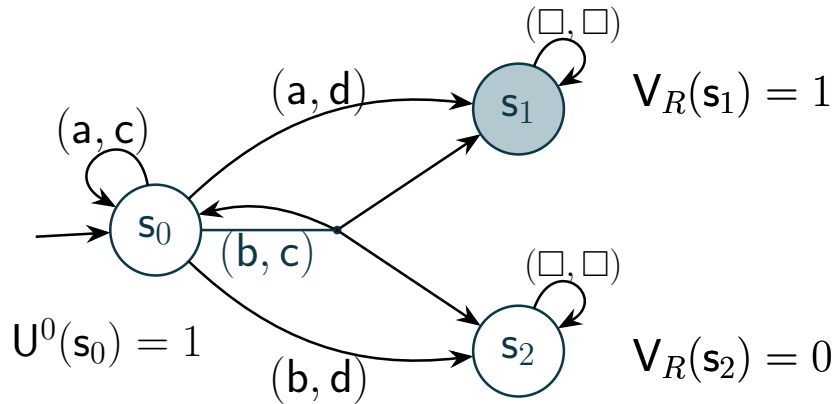
Bloated End Components

- **End component:** Set of states where staying for ever is possible
- **Bloated end component (BEC):** End component where staying pays off (TSGs: [3])
 - Example TSG: $\{s_0\}$ is BEC



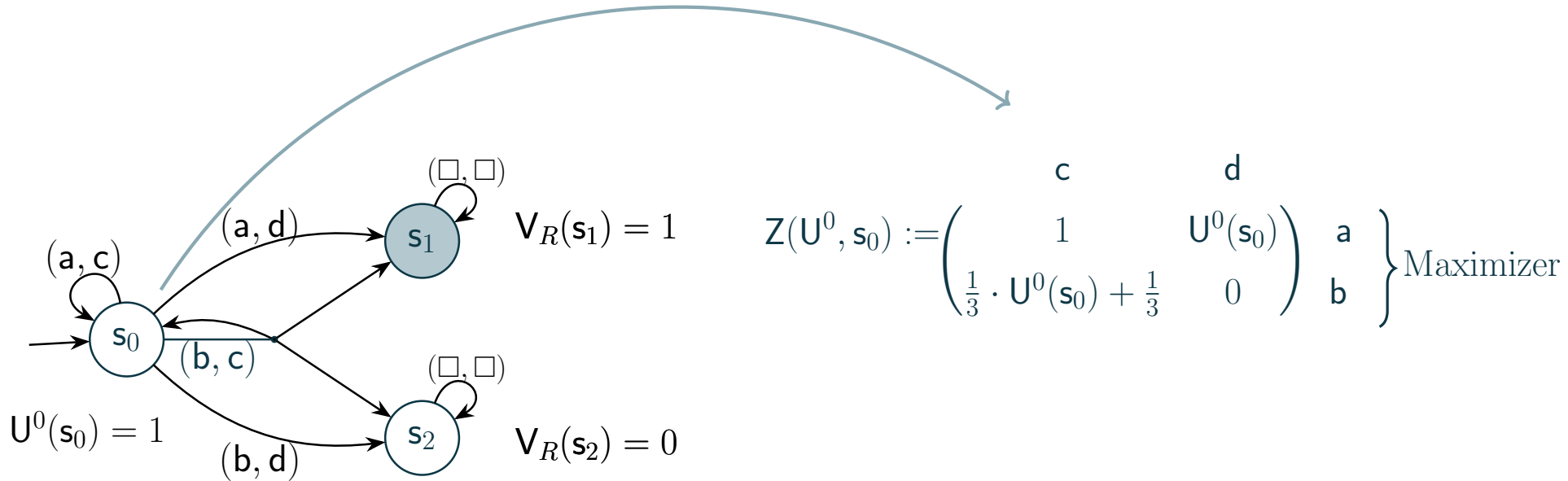
Bloated End Components

- **End component:** Set of states where staying for ever is possible
- **Bloated end component (BEC):** End component where staying pays off (TSGs: [3])
 - Example **CSG**: ???



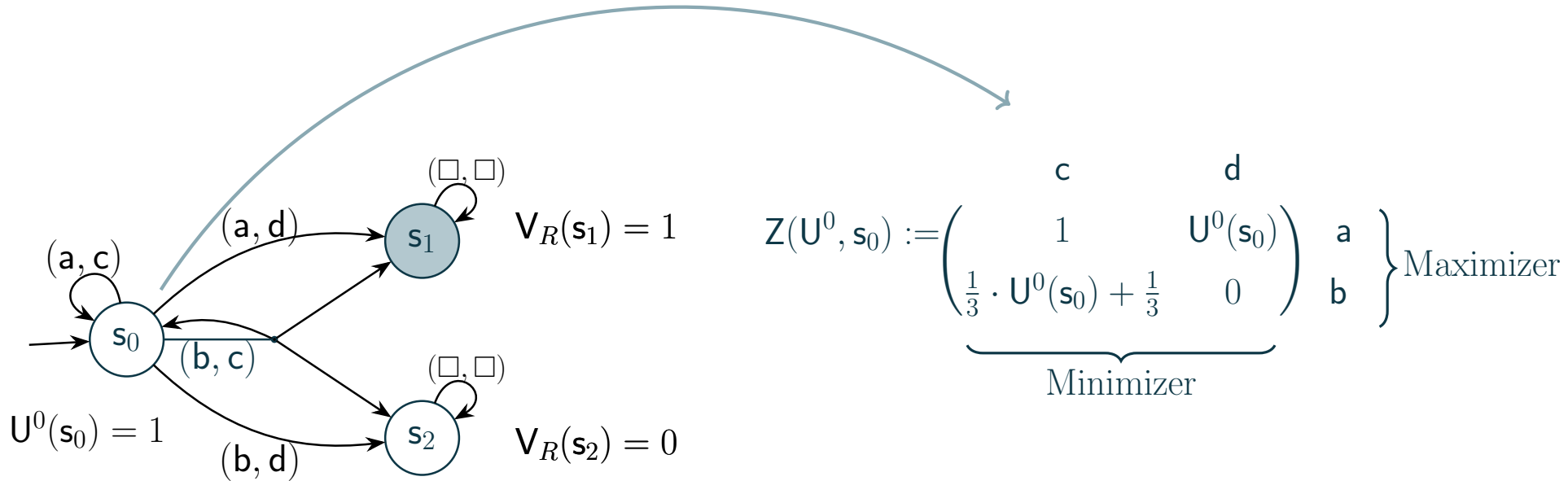
Bloated End Components

- **End component:** Set of states where staying for ever is possible
- **Bloated end component (BEC):** End component where staying pays off (TSGs: [3])



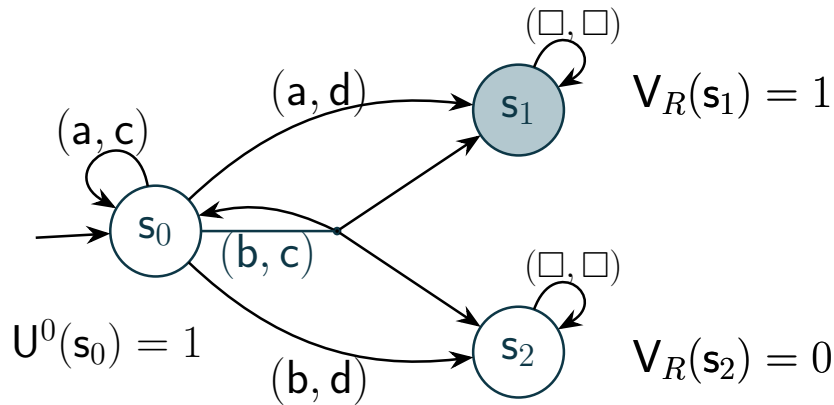
Bloated End Components

- **End component:** Set of states where staying for ever is possible
- **Bloated end component (BEC):** End component where staying pays off (TSGs: [3])



Bloated End Components

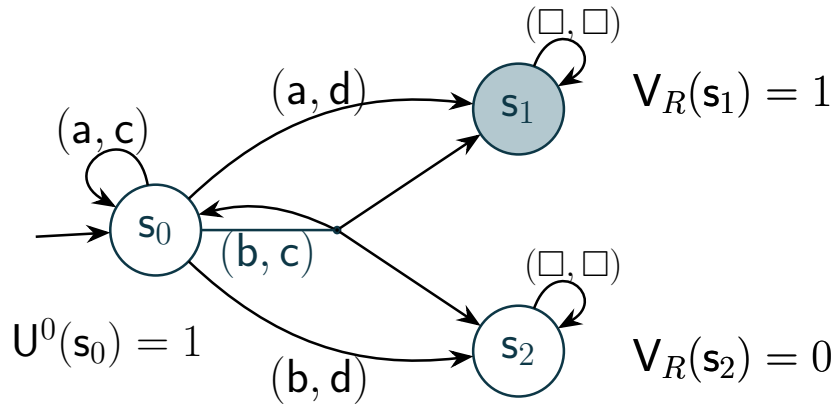
- **End component:** Set of states where staying for ever is possible
- **Bloated end component (BEC):** End component where staying pays off (TSGs: [3])



$$Z(U^0, s_0) := \left(\begin{array}{cc} & \begin{matrix} c & d \end{matrix} \\ \begin{matrix} 1 & 1 \\ \frac{1}{3} \cdot 1 + \frac{1}{3} & 0 \end{matrix} & \begin{matrix} a \\ b \end{matrix} \end{array} \right) \left. \vphantom{\begin{matrix} 1 & 1 \\ \frac{1}{3} \cdot 1 + \frac{1}{3} & 0 \end{matrix}} \right\} \begin{matrix} \text{Maximizer} \\ \text{Minimizer} \end{matrix}$$

Bloated End Components

- **End component:** Set of states where staying for ever is possible
- **Bloated end component (BEC):** End component where staying pays off (TSGs: [3])

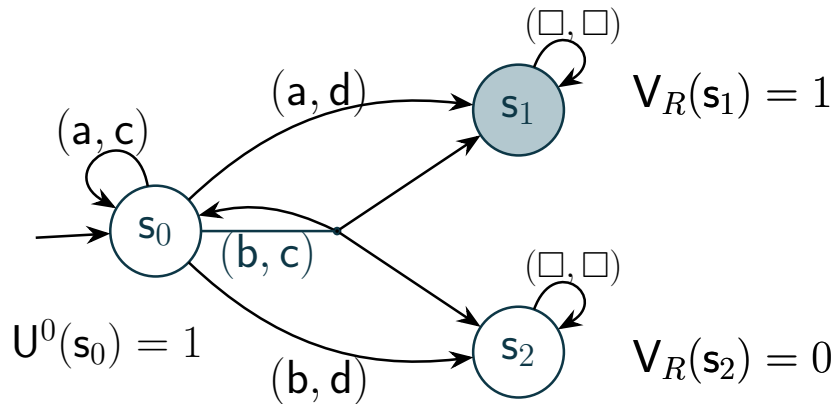


$$Z(U^0, s_0) := \left(\begin{array}{cc|c} & c & d & \\ \hline 1 & 1 & a \\ \frac{1}{3} \cdot 1 + \frac{1}{3} & 0 & b \end{array} \right) \left. \vphantom{\begin{array}{cc|c} & c & d & \\ \hline 1 & 1 & a \\ \frac{1}{3} \cdot 1 + \frac{1}{3} & 0 & b \end{array}} \right\} \begin{array}{l} \text{Maximizer} \\ \text{Minimizer} \end{array}$$

- $\{a \mapsto 1\}$ is a **hazardous strategy** [NEW!]:

Bloated End Components

- **End component:** Set of states where staying for ever is possible
- **Bloated end component (BEC):** End component where staying pays off (TSGs: [3])

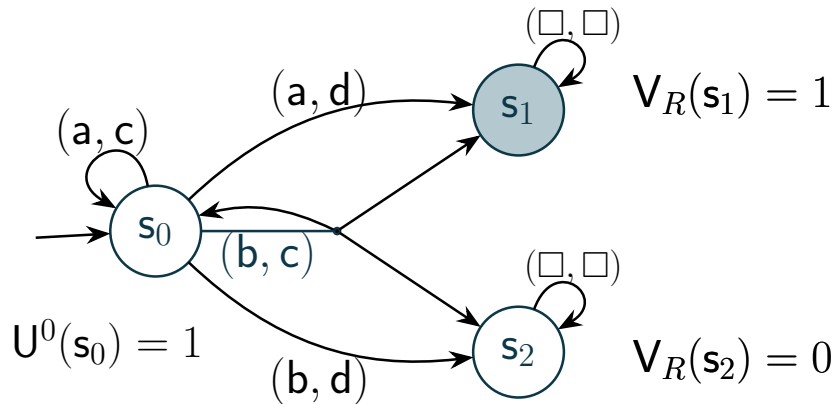


$$Z(U^0, s_0) := \left(\begin{array}{cc|c} & c & d & \\ \hline 1 & 1 & a \\ \frac{1}{3} \cdot 1 + \frac{1}{3} & 0 & b \end{array} \right) \left. \vphantom{\begin{array}{cc|c} & c & d & \\ \hline 1 & 1 & a \\ \frac{1}{3} \cdot 1 + \frac{1}{3} & 0 & b \end{array}} \right\} \begin{array}{l} \text{Maximizer} \\ \text{Minimizer} \end{array}$$

- $\{a \mapsto 1\}$ is a **hazardous strategy** [NEW!]:
- Optimal

Bloated End Components

- **End component:** Set of states where staying for ever is possible
- **Bloated end component (BEC):** End component where staying pays off (TSGs: [3])

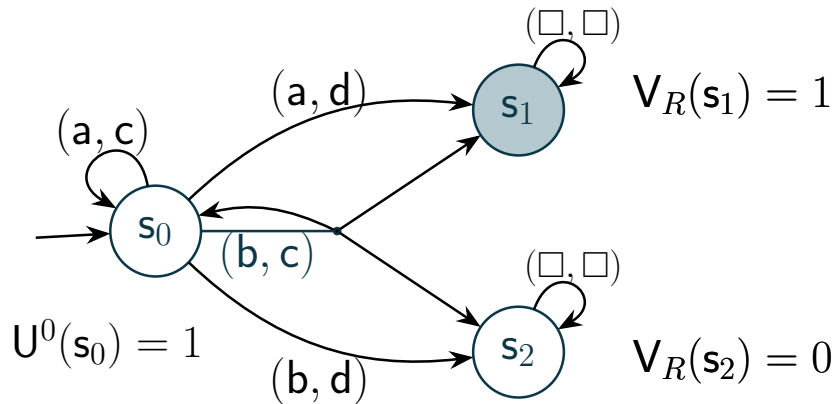


$$Z(U^0, s_0) := \left(\begin{array}{cc|c} & c & d & \\ \hline 1 & 1 & a \\ \frac{1}{3} \cdot 1 + \frac{1}{3} & 0 & b \end{array} \right) \left. \vphantom{\begin{array}{cc|c} & c & d & \\ \hline 1 & 1 & a \\ \frac{1}{3} \cdot 1 + \frac{1}{3} & 0 & b \end{array}} \right\} \begin{array}{l} \text{Maximizer} \\ \text{Minimizer} \end{array}$$

- $\{a \mapsto 1\}$ is a **hazardous strategy** [NEW!]:
 - Optimal
 - Staying

Bloated End Components

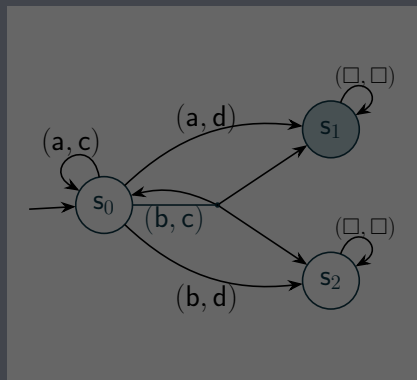
- **End component:** Set of states where staying for ever is possible
- **Bloated end component (BEC):** End component where staying pays off (TSGs: [3])



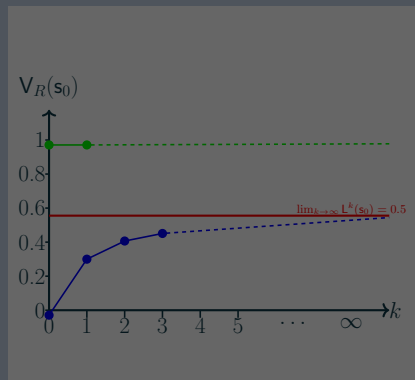
$$Z(U^0, s_0) := \left(\begin{array}{cc|c} & c & d & \\ \hline 1 & 1 & a \\ \frac{1}{3} \cdot 1 + \frac{1}{3} & 0 & b \end{array} \right) \left. \vphantom{\begin{array}{cc|c} & c & d & \\ \hline 1 & 1 & a \\ \frac{1}{3} \cdot 1 + \frac{1}{3} & 0 & b \end{array}} \right\} \begin{array}{l} \text{Maximizer} \\ \text{Minimizer} \end{array}$$

- $\{a \mapsto 1\}$ is a **hazardous strategy** [NEW!]:
 - Optimal
 - Staying
 - All leaving strategies are worse

Concurrent Stochastic Games



Bounded Value Iteration

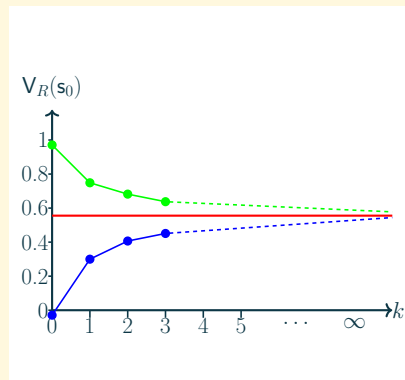


Bloated End Components

Hazardous strategy at s_0

$$\begin{array}{cc} & \begin{matrix} c & d \end{matrix} \\ \left(\begin{array}{cc} 1 & 1 \\ \frac{1}{3} \cdot 1 + \frac{1}{3} & 0 \end{array} \right) & \begin{matrix} a \\ b \end{matrix} \end{array}$$

Deflating



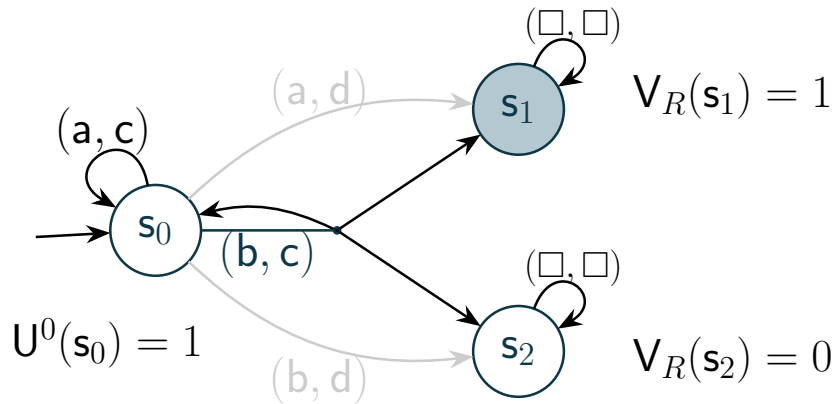
Deflating

Deflating

- **Deflating:** Reduce the upper bound to the value attainable “outside the BEC” (TSGs: [3])

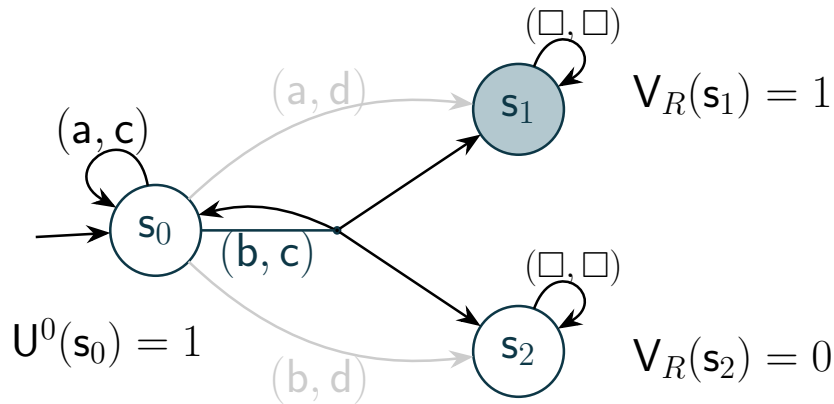
Deflating

- **Deflating:** Reduce the upper bound to the value attainable “outside the BEC” (TSGs: [3])
- $\{s_0\}$ is bloated end component



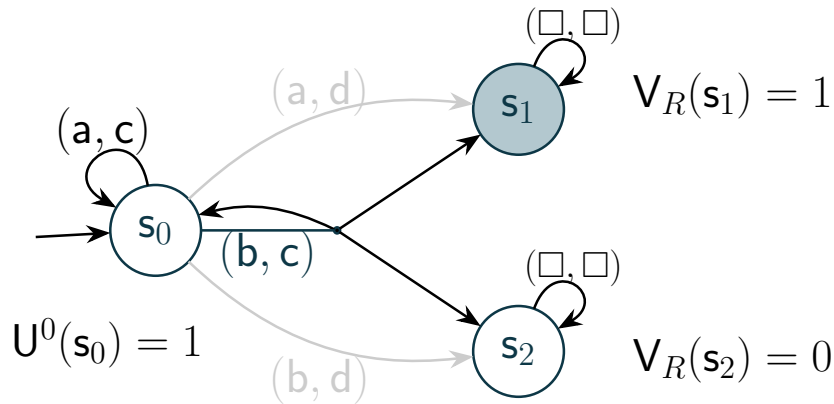
Deflating

- **Deflating:** Reduce the upper bound to the value attainable “outside the BEC” (TSGs: [3])
- $\{s_0\}$ is bloated end component
- Value attainable upon leaving $\{s_0\}$:



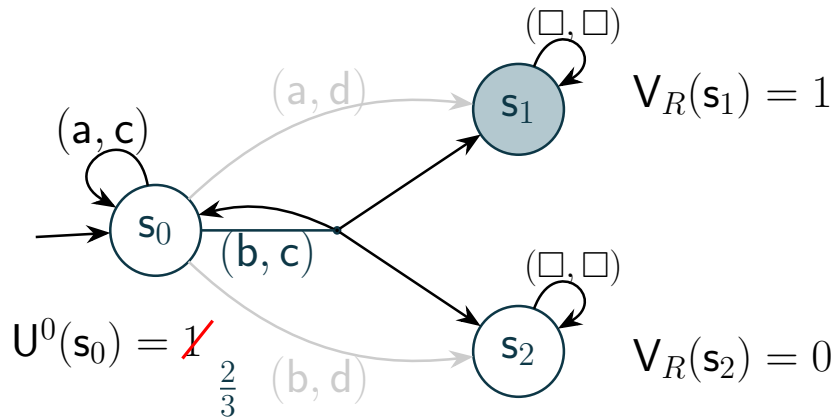
Deflating

- **Deflating:** Reduce the upper bound to the value attainable “outside the BEC” (TSGs: [3])
- $\{s_0\}$ is bloated end component
- Value attainable upon leaving $\{s_0\}$:
 $\frac{2}{3}$ using $\{b \mapsto 1\}$



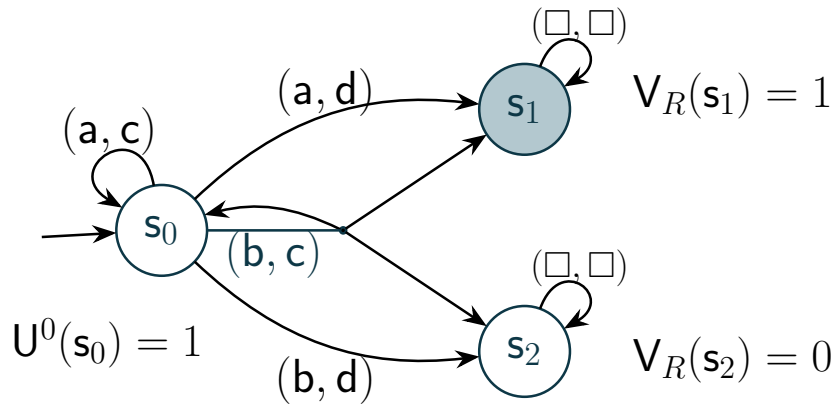
Deflating

- **Deflating:** Reduce the upper bound to the value attainable “outside the BEC” (TSGs: [3])
- $\{s_0\}$ is bloated end component
- Value attainable upon leaving $\{s_0\}$:
 $\frac{2}{3}$ using $\{b \mapsto 1\}$



Deflating

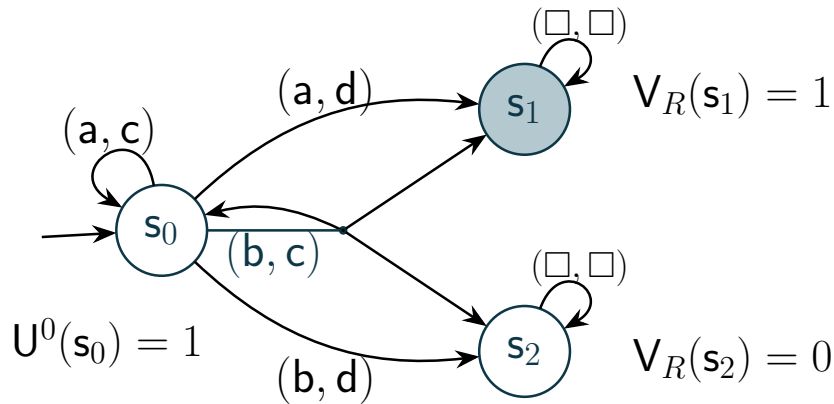
- **Deflating:** Reduce the upper bound to the value attainable “outside the BEC” (TSGs: [3])
- $\{s_0\}$ is bloated end component
- Value attainable upon leaving $\{s_0\}$:



Deflating

- **Deflating:** Reduce the upper bound to the value attainable “outside the BEC” (TSGs: [3])
- $\{s_0\}$ is bloated end component
- Value attainable upon leaving $\{s_0\}$:

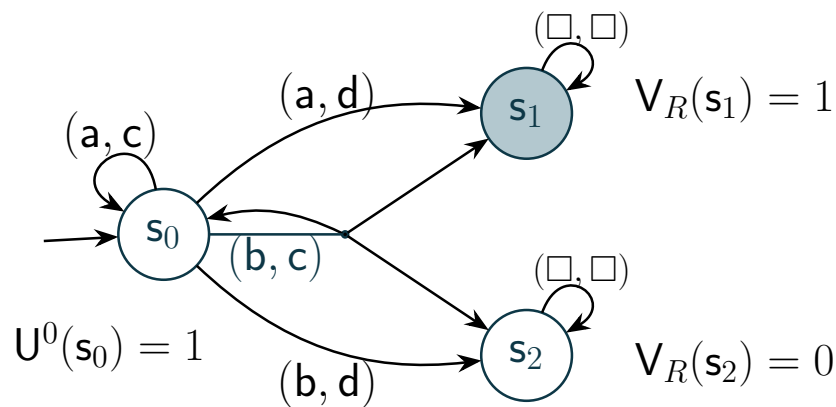
Optimal leaving strategy?



Deflating

- **Deflating**: Reduce the upper bound to the value attainable “outside the BEC” (TSGs: [3])
- $\{s_0\}$ is bloated end component
- Value attainable upon leaving $\{s_0\}$:

Optimal leaving strategy?

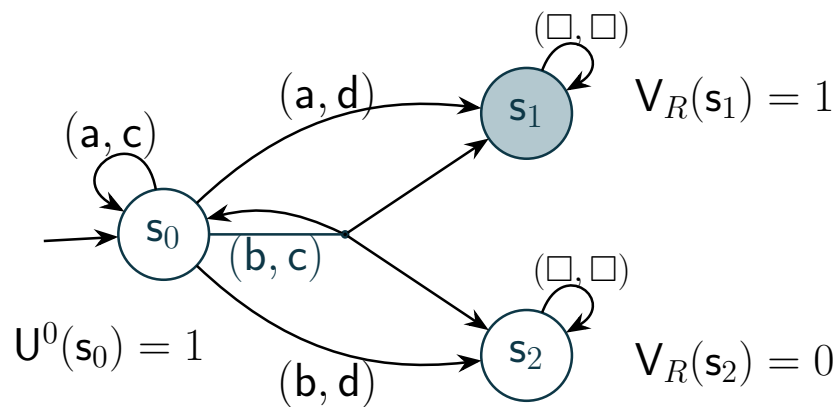


	c	d	
$\left(\begin{array}{cc} 1 & 1 \end{array} \right)$	a		
$\left(\begin{array}{cc} \frac{1}{3} \cdot 1 + \frac{1}{3} & 0 \end{array} \right)$	b		

Deflating

- **Deflating**: Reduce the upper bound to the value attainable “outside the BEC” (TSGs: [3])
- $\{s_0\}$ is bloated end component
- Value attainable upon leaving $\{s_0\}$:

Optimal leaving strategy?



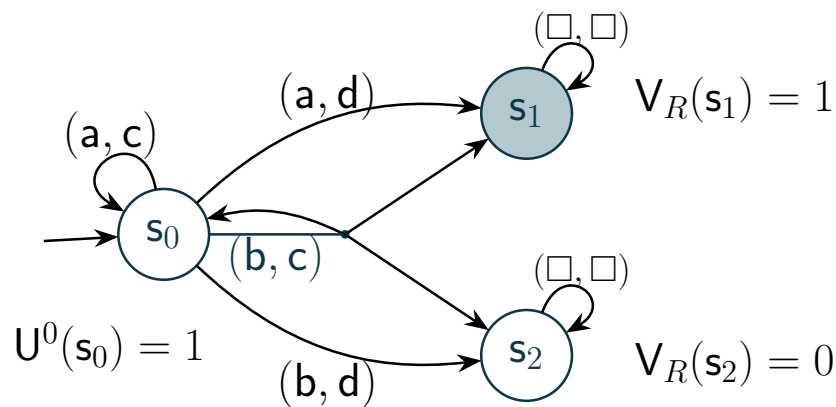
$$\begin{array}{cc} & \begin{array}{cc} c & d \end{array} \\ \begin{pmatrix} 1 & 1 \\ \frac{1}{3} \cdot 1 + \frac{1}{3} & 0 \end{pmatrix} & \begin{array}{c} a \\ b \end{array} \end{array}$$

- $\{c \mapsto 1\}$ is a **trapping strategy** [NEW!]:
Staying for **all** hazardous strategies

Deflating

- **Deflating**: Reduce the upper bound to the value attainable “outside the BEC” (TSGs: [3])
- $\{s_0\}$ is bloated end component
- Value attainable upon leaving $\{s_0\}$:

Optimal leaving strategy?



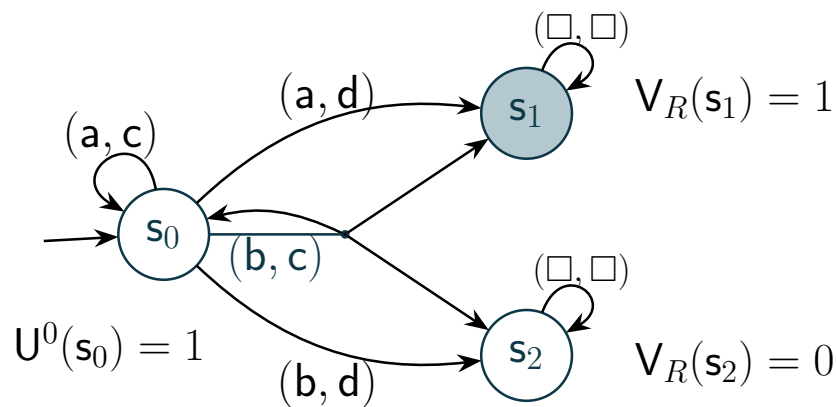
$$\begin{array}{cc} & c & d \\ \left(\begin{array}{cc} 1 & 1 \\ \frac{1}{3} \cdot 1 + \frac{1}{3} & 0 \end{array} \right) & \begin{array}{l} a \\ b \end{array} \end{array}$$

- $\{c \mapsto 1\}$ is a **trapping strategy** [NEW!]:
Staying for **all** hazardous strategies

Deflating

- **Deflating**: Reduce the upper bound to the value attainable “outside the BEC” (TSGs: [3])
- $\{s_0\}$ is bloated end component
- Value attainable upon leaving $\{s_0\}$:

Optimal leaving strategy?



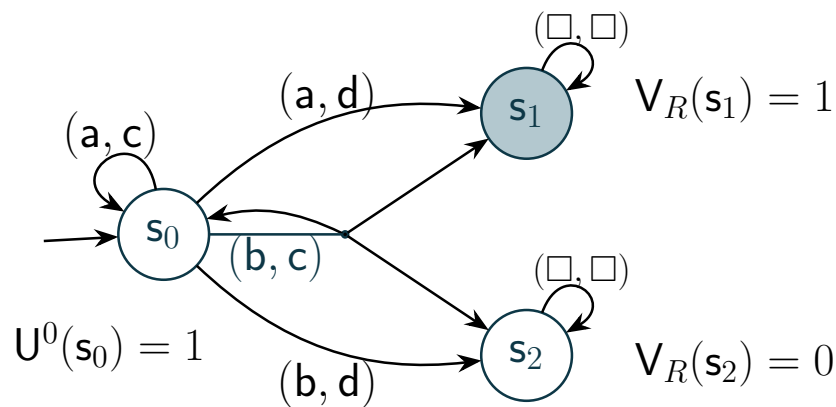
$$\begin{pmatrix} c & d \\ 1 & 1 \\ \frac{1}{3} \cdot 1 + \frac{1}{3} & 0 \end{pmatrix} \begin{matrix} a \\ b \end{matrix}$$

- $\{c \mapsto 1\}$ is a **trapping strategy** [NEW!]:
Staying for **all** hazardous strategies
- $\{b \mapsto 1\}$ is a **deflating strategy** [NEW!]:
Leaving for **a** trapping strategy

Deflating

- **Deflating**: Reduce the upper bound to the value attainable “outside the BEC” (TSGs: [3])
- $\{s_0\}$ is bloated end component
- Value attainable upon leaving $\{s_0\}$:

Optimal leaving strategy?



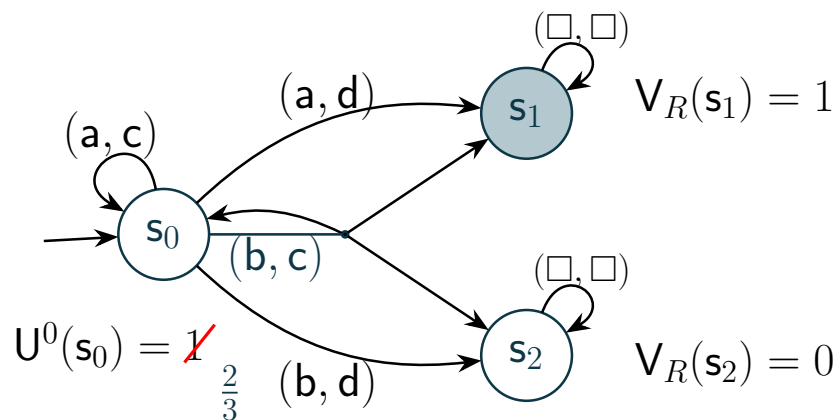
	c	d	
$\left(\begin{array}{cc} 1 & 1 \end{array} \right)$			a
$\left(\frac{1}{3} \cdot 1 + \frac{1}{3} \right)$		0	b

- $\{c \mapsto 1\}$ is a **trapping strategy** [NEW!]:
Staying for **all** hazardous strategies
- $\{b \mapsto 1\}$ is a **deflating strategy** [NEW!]:
Leaving for **a** trapping strategy
- Define **exiting sub-game** [NEW!]

Deflating

- **Deflating**: Reduce the upper bound to the value attainable “outside the BEC” (TSGs: [3])
- $\{s_0\}$ is bloated end component
- Value attainable upon leaving $\{s_0\}$:

Optimal leaving strategy?



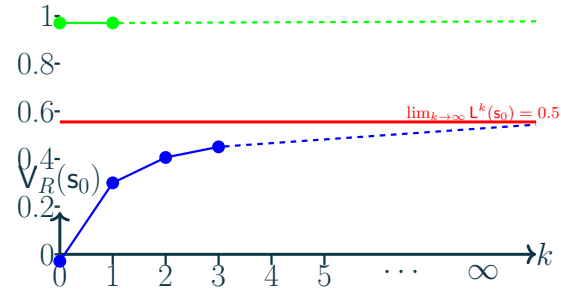
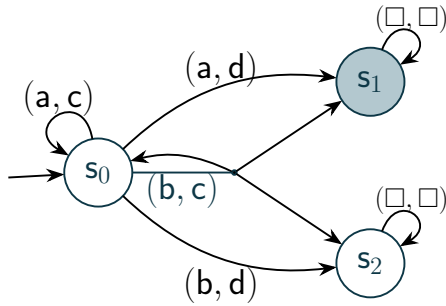
$$\begin{matrix} & c \\ \left(\frac{1}{3} \cdot 1 + \frac{1}{3}\right) & b \end{matrix}$$

- $\{c \mapsto 1\}$ is a **trapping strategy** [NEW!]:
Staying for **all** hazardous strategies
- $\{b \mapsto 1\}$ is a **deflating strategy** [NEW!]:
Leaving for **a** trapping strategy
- Define **exiting sub-game** [NEW!]

Summary

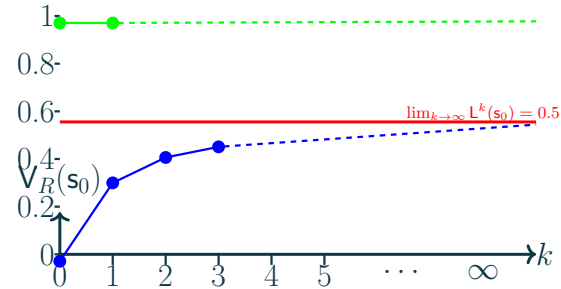
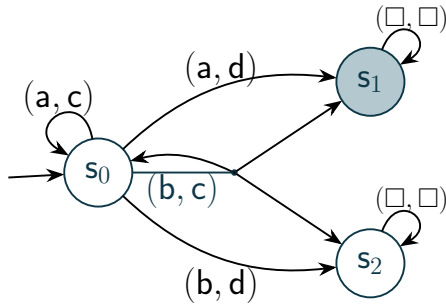
Summary

- Bounded value iteration **does not converge** from above



Summary

- Bounded value iteration **does not converge** from above

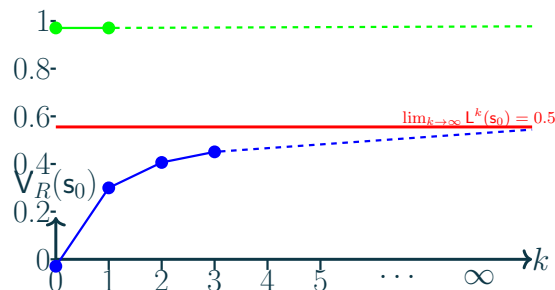
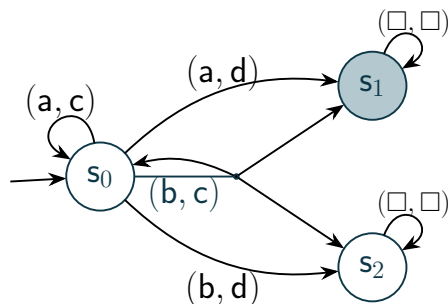


- Bloated end components** hinder the convergence

$$\begin{array}{cc} & \begin{array}{cc} c & d \end{array} \\ \begin{pmatrix} 1 & 1 \\ \frac{1}{3} \cdot 1 + \frac{1}{3} & 0 \end{pmatrix} & \begin{array}{l} a \\ b \end{array} \end{array}$$

Summary

- Bounded value iteration **does not converge** from above



- Bloated end components** hinder the convergence

$$\begin{array}{cc}
 & \begin{array}{cc} c & d \end{array} \\
 \begin{pmatrix} 1 & 1 \\ \frac{1}{3} \cdot 1 + \frac{1}{3} & 0 \end{pmatrix} & \begin{array}{c} a \\ b \end{array}
 \end{array}
 \longrightarrow
 \begin{array}{cc}
 & c \\
 \left(\frac{1}{3} \cdot 1 + \frac{1}{3} \right) & b
 \end{array}$$

- Deflating:** **trapping** and **deflating** strategies to obtain **exiting** sub-game

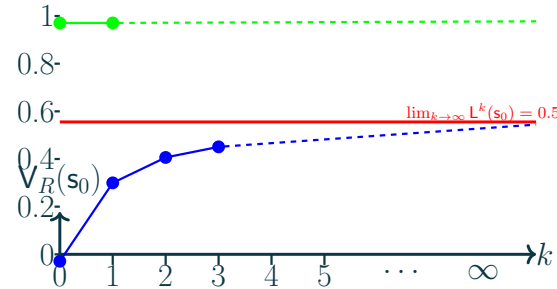
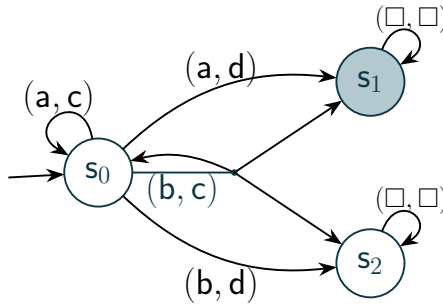
Summary

Scan for the paper:



<https://arxiv.org/pdf/2505.21087>

- Bounded value iteration **does not converge** from above



- Bloated end components** hinder the convergence

$$\begin{array}{cc} & \begin{array}{cc} c & d \end{array} \\ \begin{pmatrix} 1 & 1 \\ \frac{1}{3} \cdot 1 + \frac{1}{3} & 0 \end{pmatrix} & \begin{array}{c} a \\ b \end{array} \end{array} \longrightarrow \begin{array}{cc} & c \\ \left(\frac{1}{3} \cdot 1 + \frac{1}{3} \right) & b \end{array}$$

- Deflating:** **trapping** and **deflating** strategies to obtain **exiting** sub-game

Thank You!

References

- [1] K. Chatterjee and T. A. Henzinger. Value Iteration. In Orna Grumberg and Helmut Veith, 25 Years of Model Checking: History, Achievements, Perspectives. Springer, 2008.
- [2] P. Ashok, K. Chatterjee, P. Daca, J. Kretinsky, and T. Meggendorfer. Value Iteration for Long-Run Average Reward in Markov Decision Processes. In Rupak Majumdar and Viktor Kuncak, ~ editors, Computer Aided Verification. Springer International Publishing, 2017.
- [3] J. Eisentraut, E. Kelmendi, J. Kretinsky, and M. Weininger. Value iteration for simple stochastic games: Stopping criterion and learning algorithm. Information and Computation, 2022.