Stopping Criteria for Value Iteration on Stochastic Games with Quantitative Objectives

Jan Kretinsky, Tobias Meggendorfer, Maximilian Weininger
Talk in one slide

- **Probabilistic systems**: Best algorithm (usually) is *Value Iteration (VI)*
- But: Requires a **stopping criterion**
  For *Stochastic Games (SG)* with most **infinite-horizon, quantitative objectives** there is **none**!

- This paper: **Uniform** solution for **large class of quantitative objectives**
  (including total reward, mean payoff, …)
Talk in one slide

- **Probabilistic systems**: Best algorithm (usually) is **Value Iteration (VI)**
- But: Requires a **stopping criterion** for correctness
  For **Stochastic Games (SG)** with most **infinite-horizon, quantitative objectives** there is **none!**

- This paper: **Uniform** solution for **large class of quantitative objectives** (including total reward, mean payoff, …)

1-player SG: separate papers giving stopping criteria for each objective [BCC+14, HM14, BKL+17, ACD+17].
Talk in one slide

- **Probabilistic systems**: Best algorithm (usually) is **Value Iteration (VI)**
- But: Requires a **stopping criterion** for correctness
  For **Stochastic Games (SG)** with most **infinite-horizon, quantitative objectives** there is **none!**

- This paper: **Uniform** solution for **large class of quantitative objectives**
  (including total reward, mean payoff, …)

1-player SG: separate papers giving stopping criteria for each objective [BCC+14, HM14, BKL+17, ACD+17].

Unifies all previous ones and is more broadly applicable.
Stochastic Games

Stochastic: vs.

Games:

\[ s \quad \text{vs.} \quad t \]
Stochastic Games

Games: \( S \) vs. \( t \)

Goal: \( \text{Diamond} \)
Stochastic Games and Value Iteration

<table>
<thead>
<tr>
<th>Iteration</th>
<th>L(s)</th>
<th>L(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Stochastic Games and Value Iteration

Bellman update:
\[ x_i(s) = \text{opt}_a x_{i-1}(s, a) \]

<table>
<thead>
<tr>
<th>Iteration</th>
<th>L(s)</th>
<th>L(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Stochastic Games and Value Iteration

Bellman update:

\[ x_i(s) = \max_a x_{i-1}(s, a) \]
Stochastic Games and Value Iteration

Bellman update:

\[ x_i(s) = \text{opt}_a x_{i-1}(s, a) \]

<table>
<thead>
<tr>
<th>Iteration</th>
<th>L(s)</th>
<th>L(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Stochastic Games and Value Iteration

Bellman update:

\[ x_i(s) = \text{opt}_a x_{i-1}(s, a) \]

<table>
<thead>
<tr>
<th>Iteration</th>
<th>L(s)</th>
<th>L(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1/3</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Stochastic Games and Value Iteration**

**Bellman update:**

\[ x_i(s) = \text{opt}_a x_{i-1}(s, a) \]

<table>
<thead>
<tr>
<th>Iteration</th>
<th>L(s)</th>
<th>L(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1/3</td>
</tr>
<tr>
<td>2</td>
<td>1/3</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Stochastic Games and Value Iteration

Bellman update:
\[ x_i(s) = \text{opt}_a x_{i-1}(s, a) \]

<table>
<thead>
<tr>
<th>Iteration</th>
<th>L(s)</th>
<th>L(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1/3</td>
</tr>
<tr>
<td>2</td>
<td>1/3</td>
<td>4/9</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Stochastic Games and Value Iteration

Bellman update:

\[ x_i(s) = \max_a x_{i-1}(s, a) \]

<table>
<thead>
<tr>
<th>Iteration</th>
<th>L(s)</th>
<th>L(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1/3</td>
</tr>
<tr>
<td>2</td>
<td>1/3</td>
<td>4/9</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Stochastic Games and Value Iteration

Bellman update:

\[ x_i(s) = \text{opt}_a x_{i-1}(s, a) \]

<table>
<thead>
<tr>
<th>Iteration</th>
<th>L(s)</th>
<th>L(t)</th>
<th>U(s)</th>
<th>U(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1/3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1/3</td>
<td>4/9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Stochastic Games and Value Iteration

Bellman update:

$$x_i(s) = \text{opt}_a x_{i-1}(s, a)$$

<table>
<thead>
<tr>
<th>Iteration</th>
<th>L(s)</th>
<th>L(t)</th>
<th>U(s)</th>
<th>U(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1/3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1/3</td>
<td>4/9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Stochastic Games and Value Iteration

Bellman update:

\[ x_i(s) = \text{opt}_a x_{i-1}(s, a) \]

<table>
<thead>
<tr>
<th>Iteration</th>
<th>L(s)</th>
<th>L(t)</th>
<th>U(s)</th>
<th>U(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1/3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1/3</td>
<td>4/9</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Stochastic Games and Value Iteration

Bellman update:

\[ x_i(s) = \operatorname{opt}_a x_{i-1}(s, a) \]

<table>
<thead>
<tr>
<th>Iteration</th>
<th>L(s)</th>
<th>L(t)</th>
<th>U(s)</th>
<th>U(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1/3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1/3</td>
<td>4/9</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Stochastic Games and Value Iteration

Bellman update:
\[ x_i(s) = \text{opt}_a x_{i-1}(s, a) \]

<table>
<thead>
<tr>
<th>Iteration</th>
<th>L(s)</th>
<th>L(t)</th>
<th>U(s)</th>
<th>U(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1/3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1/3</td>
<td>4/9</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Stochastic Games and Value Iteration

Bellman update:
\[ x_i(s) = \operatorname{opt}_a x_{i-1}(s, a) \]

<table>
<thead>
<tr>
<th>Iteration</th>
<th>L(s)</th>
<th>L(t)</th>
<th>U(s)</th>
<th>U(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1/3</td>
<td>1</td>
<td>2/3</td>
</tr>
<tr>
<td>2</td>
<td>1/3</td>
<td>4/9</td>
<td>2/3</td>
<td>5/9</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Stochastic Games and Value Iteration
Stochastic Games and Value Iteration

Bellman update:
\[ x_i(s) = \text{opt}_a x_{i-1}(s, a) \]

<table>
<thead>
<tr>
<th>Iteration</th>
<th>L(s)</th>
<th>L(t)</th>
<th>L(u)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Stochastic Games and Value Iteration

Bellman update:

\[ x_i(s) = \max_a x_{i-1}(s, a) \]

<table>
<thead>
<tr>
<th>Iteration</th>
<th>L(s)</th>
<th>L(t)</th>
<th>L(u)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Stochastic Games and Value Iteration

Bellman update:
\[ x_i(s) = \text{opt}_a x_{i-1}(s, a) \]

<table>
<thead>
<tr>
<th>Iteration</th>
<th>L(s)</th>
<th>L(t)</th>
<th>L(u)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Bellman update:

\[ x_i(s) = \text{opt}_a x_{i-1}(s, a) \]
Stochastic Games and Value Iteration

Bellman update:
\[ x_i(s) = \text{opt}_a x_{i-1}(s, a) \]

<table>
<thead>
<tr>
<th>Iteration</th>
<th>L(s)</th>
<th>L(t)</th>
<th>L(u)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Stochastic Games and Value Iteration

Bellman update:
\[ x_i(s) = \text{opt}_a x_{i-1}(s, a) \]

<table>
<thead>
<tr>
<th>Iteration</th>
<th>L(s)</th>
<th>L(t)</th>
<th>L(u)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Stochastic Games and Value Iteration

Bellman update:

\[ x_i(s) = \text{opt}_a x_{i-1}(s, a) \]

<table>
<thead>
<tr>
<th>Iteration</th>
<th>L(s)</th>
<th>L(t)</th>
<th>L(u)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1/3</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1/3</td>
<td>4/9</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Stochastic Games and Value Iteration

Inform the algorithm about the consequences of staying forever.

Bellman update:

\[ x_i(s) = \text{opt}_a x_{i-1}(s, a) \]

<table>
<thead>
<tr>
<th>Iteration</th>
<th>L(s)</th>
<th>L(t)</th>
<th>L(u)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1/3</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1/3</td>
<td>4/9</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Should I stay or should I go now?
Should I stay or should I go now?

Reachability:
   stay=0

Safety:
   stay=1
Should I stay or should I go now?

**Reachability:**

\[
\text{stay}=0 \quad \text{max(stay,exit)} = \text{exit}
\]

**Safety:**

\[
\text{stay}=1
\]
Should I stay or should I go now?

Reachability:

\[
\text{stay} = 0 \quad \text{max}(\text{stay}, \text{exit}) = \text{exit}
\]

Safety:

\[
\text{stay} = 1 \quad \text{min}(\text{stay}, \text{exit}) = \text{exit}
\]
Should I stay or should I go now?

Reachability:

\[
\text{stay}=0 \quad \text{max}(\text{stay}, \text{exit}) = \text{exit} \\
\text{min}(\text{stay}, \text{exit}) = 0
\]

Safety:

\[
\text{stay}=1 \quad \text{max}(\text{stay}, \text{exit}) = 1 \\
\text{min}(\text{stay}, \text{exit}) = \text{exit}
\]
Deflating and Inflating for Mean Payoff

\[ a : 10 \]
Deflating and Inflating for Mean Payoff

\[
a : 10 \\
S \xrightarrow{b: 5} 5
\]
Deflating and Inflating for Mean Payoff

a: 10

b: 5

Inflate from 0 to **exit** value 5
Deflating and Inflating for Mean Payoff

a: 10

Inflate from 0 to exit value 5

b: 5

a: 5

b: 10

10
Deflating and Inflating for Mean Payoff

a: 10

b: 5

Inflate from 0 to exit value 5

s

5

a: 5

b: 10

Inflate from 0 to stay value 5

s

10
Deflating and Inflating for Mean Payoff

Inflate from 0 to exit value 5

And dually for Maximizer states

Inflate from 0 to stay value 5
Where can I stay and go?

- Idea: If opponent thinks staying here is good, what happens if we really do stay?
Where can I stay and go?

- Idea: If opponent thinks staying here is good, what happens if we really do stay?
- Thus: Fix opponent’s strategy and analyse remaining cycles.
Where can I stay and go?

- Idea: If opponent thinks staying here is good, what happens if we really do stay?
- Thus: Fix opponent’s strategy and analyse remaining cycles

SOUND
Where can I stay and go?

- Idea: If opponent thinks staying here is good, what happens if we really do stay?

- Thus: Fix opponent's strategy and analyse remaining cycles
Where can I stay and go?

- Idea: If opponent thinks staying here is good, what happens if we really do stay?
- Thus: Fix opponent's strategy and analyse remaining cycles
Conclusion

- Given: **Stochastic Games with quantitative objectives**
  (including reachability, safety, mean payoff, expected total reward, ...),
- Goal: Solving them **quickly** and with **precision-guarantees**
- Approach: **Value Iteration** with our **new stopping criterion**

  Idea: Inform the algorithm about the consequences of staying forever:
  Should I stay or should I go now?

Unifies previous work [BCC+14, HM14, BKL+17, ACD+17, KKKW18, PTHH20] in an **elegant, objective-independent way**