Best-Effort Synthesis for MDPs

work in progress

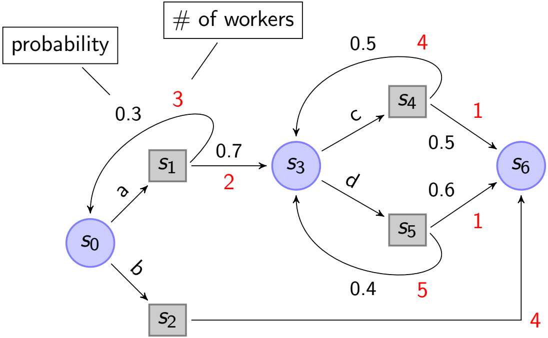
Marco Faella Laurent Doyen

Università di Napoli "Federico II", Italy

LSV, ENS Cachan & CNRS, France

September 19th, 2013

Choosing a Production Plan



- Different ways of producing the same product
- Cost of a path is the maximum # of workers along the path
- What's the *best* production plan?

Choosing a Production Plan

What is the best plan if...

1. ...I can hire workers by the day
   minimum **expected** number of workers

plan	expected		
a,c	3.15		
a,d	3.38		
b	4		

Choosing a Production Plan

What is the best plan if...

1. ...I can hire workers by the day
   minimum **expected** number of workers

plan	expected		
a,c	3.15		
a,d	3.38		
b	4		

Choosing a Production Plan

What is the best plan if...

1. ...I can hire workers by the day
   minimum **expected** number of workers
2. ...I must hire for the whole year
   minimum **worst-case** number of workers

plan	expected	worst	
a,c	3.15	4	
a,d	3.38	5	
b	4	4	

Choosing a Production Plan

What is the best plan if...

1. ...I can hire workers by the day
   minimum **expected** number of workers
2. ...I must hire for the whole year
   minimum **worst-case** number of workers

plan	expected	worst	
a,c	3.15	4	
a,d	3.38	5	
b	4	4	

Choosing a Production Plan

What is the best plan if...

1. ...I can hire workers by the day
   minimum **expected** number of workers
2. ...I must hire for the whole year
   minimum **worst-case** number of workers
3. ...today I have only two workers
   minimum **best-case** number of workers

plan	expected	worst	best
a,c	3.15	4	2
a,d	3.38	5	2
b	4	4	4

Choosing a Production Plan

What is the best plan if...

1. ...I can hire workers by the day
   minimum **expected** number of workers
2. ...I must hire for the whole year
   minimum **worst-case** number of workers
3. ...today I have only two workers
   minimum **best-case** number of workers

plan	expected	worst	best
a,c	3.15	4	2
a,d	3.38	5	2
b	4	4	4

Key Idea

Resolve ties by lexicographically combining different semantics

The 3 Basic Solution Concepts

- **w** for **w**orst case — two-player game
- **e** for **e**xpected case — standard criterion for MDPs
- **b** for **b**est case — one-player game, i.e., graph

- Optimal strategies exist
- For **w** and **b**, deterministic and memoryless optimal strategies exist

Composite Solution Concepts

- Basic solution concepts (SCs) can be lexicographically composed
- For instance, SC "**we**":

  first optimize **w**orst-case, then optimize **e**xpected case

A Taxonomy of Solution Concepts

$$
\begin{tikzcd}
 & \mathbf{w} & & & \mathbf{e} & & & \mathbf{b} & \\
\mathbf{we} \arrow[ur] & & \mathbf{wb} \arrow[ul] & \mathbf{ew} \arrow[ur] & & \mathbf{eb} \arrow[ul] & \mathbf{bw} \arrow[ur] & & \mathbf{be} \arrow[ul] \\
\mathbf{web} \arrow[u] & & \mathbf{wbe} \arrow[u] & \mathbf{ewb} \arrow[u] & & \mathbf{ebw} \arrow[u] & \mathbf{bwe} \arrow[u] & & \mathbf{bew} \arrow[u]
\end{tikzcd}
$$

An arrow from α to β signifies that α-optimal implies β-optimal.

Fact

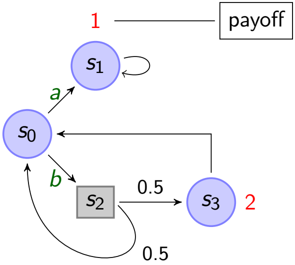
All pairs of distinct SCs induce distinct notions of optimality.

Interesting Problems

For an SC α:

- What is the complexity of computing the α-value of a state?
- Do α-optimal strategies always exist?
- If they do, are they deterministic and how much memory do they need?

Existence of Optimal Strategies: SC "**web**"



Goal: **maximize** the maximum payoff over trace

Value of strategy b^ω: $(0, 2, 2)$

$\downarrow$ *improved by*

Value of strategy a: $(1, 1, 1)$

$\downarrow$ *improved by*

Value of strategy $b^n a$: $(1, 2 - \frac{1}{2^n}, 2)$

Strategies get arbitrarily close to $(1, 2, 2)$

No **web**-optimal strategy exists!

Notion of Value of a State

- What's the **web**-value of the previous game?

- Strategies can achieve $(1, 2 - \varepsilon, 2)$ for all $\varepsilon > 0$
- Classic notion of value of a state: $\mathit{val}(s) = \sup_{\sigma}^{\text{lex}} \mathit{val}(s, \sigma)$
- Here: $\mathit{val}(s_0) = (1, 2, -\infty)$

- However, $(1, 2, 2)$ would be a more interesting "value"
- We need a different notion of value...

Principles of Best-Effort Control

Find a strategy that is as “smart” as possible.

1. As specific as possible
   - i.e., resolves as many ties as possible
2. Exploits environment’s “errors”
   - i.e., takes advantage of non strictly adversarial environment
   - a.k.a. subgame-perfection, sequential rationality

Other Applications

Compute smarter strategies for standard 2-player games.

Given a 2-player game (no probabilities),
compute a **web**-optimal strategy,
assuming uniform transition probabilities from all states

Conclusions

Working Hypothesis

Computing value is in **PTIME** for all solution concepts.

stay tuned...