One-Counter Markov Decission Processes

T. Brázdil, V. Brožek, K. Etessami, A. Kučera, D. Wojtczak

13 Nov 2009 / Memics 2009

A presentation of a SODA 2010 paper.

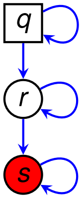
Outline

1. Model and Objectives
   - Finite MDP
   - Extensions
   - One-Counter Automata
   - Solvency Games
2. Problems
   - Questions
   - Long Run Properties
   - Non-selective Termination
   - Selective Termination
   - Solvency Games
3. Future Work

Outline

Finite MDP with Reachability

- A system with finitely many states (finite underlying graph)
- 2 types of states: random and controlled
- Objective: reach/avoid a subset F of states
- Deciding the following is in PTIME
  - "$\exists$ strategy: $\mathcal{P}(\text{reach } F) \geq \varrho$?" (reachability, MAX)
  - "$\exists$ strategy: $\mathcal{P}(\text{reach } F) \leq \varrho$?" (safety, MIN)



Outline

Pros and Cons of Finite MDP

Pros: Well studied model; low complexity of model-checking

Cons: Infinite space often required:

- Recursive procedures
- Unbounded counting

Extending the Finite MDP

Requirement: finite representability

- Finite Automaton → Pushdown Automaton
  - Natural model for recursive processes
  - Transition graph of PDA is infinite
  - Cons: Undecidability for very basic questions
- Finite Automaton → One-Counter Automaton
  - Interesting model from multiple points of view
  - Transition graph of OC is infinite
  - Cons: Weaker than PDA
  - Pros: Interesting questions decidable

Extending the Finite MDP

Requirement: finite representability

- Finite Automaton → Pushdown Automaton
  - Natural model for recursive processes
  - Transition graph of PDA is infinite
  - Cons: Undecidability for very basic questions
- Finite Automaton → One-Counter Automaton
  - Interesting model from multiple points of view
  - Transition graph of OC is infinite
  - Cons: Weaker than PDA
  - Pros: Interesting questions decidable

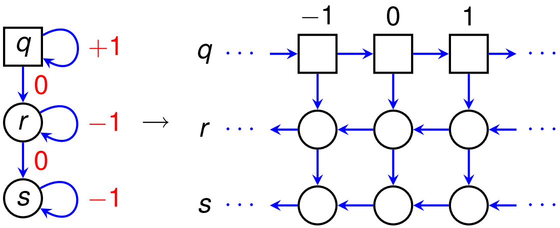
Outline

Extending with Counter

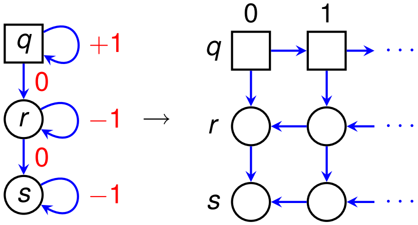
Motivation:

- A natural restriction of PDA – unary stack alphabet
- Enhancing the finite automaton inside MDP with a counter
- (Adversarial extension of) a discrete time Quasi Birth-Death Process, a model of a queue
- OC-MDP subsume *solvency games* [Berger et al., 2008]

OC-MDP in Pictures



OC-MDP in Pictures



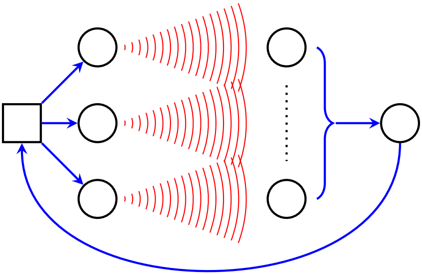
OC-MDP in Detail

- Game graph: vertices = $s(n)$ (state and counter value)
- Objectives – 2 kinds of termination
  - non-selective = reaching $s(0)$ in any state s
  - selective = reaching $s(0)$ for $s \in F$, a specified subset of states
- Reachability is equivalent to selective termination

Outline

Solvency Games

- Initial pot of money
- Finite # of wealth-changing distributions
- Repeated choices of the distributions
- Goal: avoid bankrupt (not maximise profit)



Only red changes counter

Outline

Interesting Questions

Is there a strategy for the OC-MDP so that the probability of the following events satisfies e.g. $= 1$, > 0, $\geq \varrho \ldots$?

- The stack is erased and any state, or a particular one is reached
- The queue represented by the QBD gets eventually empty
- The gambler is going bankrupt
- The value of the counter gets less than every bound
- …

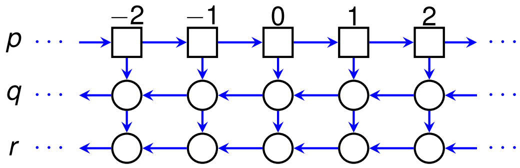
Outline

1. Model and Objectives
   - Finite MDP
   - Extensions
   - One-Counter Automata
   - Solvency Games
2. Problems
   - Questions
   - Long Run Properties
   - Non-selective Termination
   - Selective Termination
   - Solvency Games
3. Future Work

Accumulated Reward

Maximize $\mathcal{P}(\{\liminf_{n\to\infty}(\text{counter value at } n\text{-th step}) = -\infty\})$

- Optimal strategies always exist, are deterministic and only need to know the state (not history, nor counter value)
- Finding an optimal strategy is in polynomial time

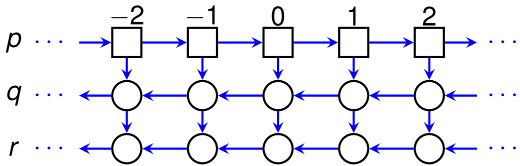


Minimize $\mathcal{P}(\{\liminf_{n\to\infty}(\text{counter value at } n\text{-th step}) = -\infty\})$?

Accumulated Reward

Maximize $\mathcal{P}(\{\liminf_{n\to\infty}(\text{counter value at } n\text{-th step}) = -\infty\})$

- Optimal strategies always exist, are deterministic and only need to know the state (not history, nor counter value)
- Finding an optimal strategy is in polynomial time

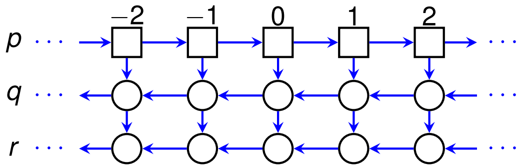


Minimize $\mathcal{P}(\{\liminf_{n\to\infty}(\text{counter value at } n\text{-th step}) = -\infty\})$?

Accumulated Reward

Maximize $\mathcal{P}(\{\liminf_{n\to\infty}(\text{counter value at } n\text{-th step}) = -\infty\})$

- Optimal strategies always exist, are deterministic and only need to know the state (not history, nor counter value)
- Finding an optimal strategy is in polynomial time



Maximize $\mathcal{P}(\{\limsup_{n\to\infty}(\text{counter value at } n\text{-th step}) = -\infty\})$?

Strategy Synthesis I.

How to find a strategy for maximizing

$$\mathcal{V} := \mathcal{P}(\{\liminf_{n\to\infty}(\text{counter at step } n) = -\infty\})?$$

It suffices to define it on vertices with $\mathcal{V} = 1$

- The rest is reaching these vertices
- The proof exploits the simple form of optimal strategies – plugging them in yields a finite Markov Chain

Strategy Synthesis I.

How to find a strategy for maximizing

$$\mathcal{V} := \mathcal{P}(\{\liminf_{n\to\infty}(\text{counter at step } n) = -\infty\})?$$

It suffices to define it on vertices with $\mathcal{V} = 1$

- The rest is reaching these vertices
- The proof exploits the simple form of optimal strategies – plugging them in yields a finite Markov Chain

Strategy Synthesis II.

How to find a strategy for

$$\mathcal{P}\left(\{\liminf_{n\to\infty}(\text{counter at step } n) = -\infty\}\right) = 1?$$

It is reducible to finding a strategy for

$$\mathcal{P}\left(\{\lim_{n\to\infty}\left(\frac{\text{counter at step } n}{n}\right) \leq 0\}\right) = 1$$

Strategy Synthesis II.

How to find a strategy for

$$\mathcal{P}\left(\{\liminf_{n\to\infty}(\text{counter at step } n) = -\infty\}\right) = 1?$$

It is reducible to finding a strategy for

$$\mathcal{P}\left(\{\lim_{n\to\infty}\left(\frac{\text{counter at step } n}{n}\right) \leq 0\}\right) = 1$$

Strategy Synthesis III.

How to find a strategy for

$$\mathcal{P}\left(\left\{\lim_{n\to\infty}\left(\frac{\text{counter at step } n}{n}\right) \leq 0\right\}\right) = 1?$$

It is reducible to finding a strategy for minimizing

$$\mathbb{E}\left(\lim_{n\to\infty}\left(\frac{\text{counter at step } n}{n}\right)\right)$$

This is solvable with linear programming

Strategy Synthesis III.

How to find a strategy for

$$\mathcal{P}\left(\{\lim_{n\to\infty}\left(\frac{\text{counter at step } n}{n}\right) \leq 0\}\right) = 1?$$

It is reducible to finding a strategy for minimizing

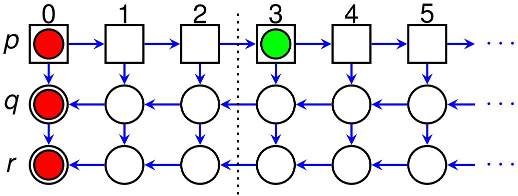
$$\mathbb{E}\left(\lim_{n\to\infty}\left(\frac{\text{counter at step } n}{n}\right)\right)$$

This is solvable with linear programming

Outline

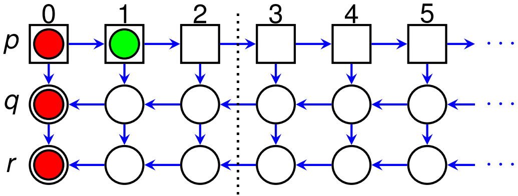
Non-selective Termination

- If $\mathcal{V} := \sup_\sigma \mathcal{P}^\sigma(\text{●}) = 1$ then there is an optimal strategy σ which is deterministic and only needs to know the state (not history, nor counter value)
- This strategy can be computed in polynomial time



Non-selective Termination

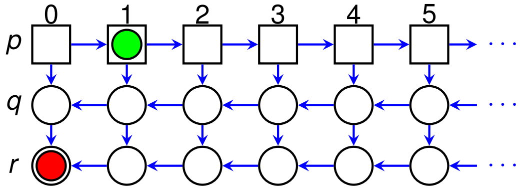
- If $\mathcal{V} := \sup_\sigma \mathcal{P}^\sigma(\bullet) = 1$ then there is an optimal strategy σ which is deterministic and only needs to know the state (not history, nor counter value)
- This strategy can be computed in polynomial time



Outline

Selective Termination

- Even if $\mathcal{V} := \sup_\sigma \mathcal{P}^\sigma(●) = 1$ then there may not be an optimal strategy σ
- If there is then there is a deterministic one and only need to know the state and a regular information about the counter
- Deciding its existence is PSPACE-hard
- It can be computed in exponential time



Outline

Solvency Games

- Solvency games $\subset$ OC-MDP with one controlled state
- The following questions related to non-selective termination are decidable in polynomial time:
  "$\exists$ a strategy for $\mathcal{P}(\{\text{bankrupt}\})$ (i.e. termination)
  - $= 0$?
  - > 0?
  - < 1?
  - $= 1$?

Proof Idea for Case < 1

When $\exists$ a strategy for $\mathcal{P}(\{\text{bankrupt}\}) < 1$?

Easy part – sufficient conditions:

- If $\exists$ a strategy for $\mathcal{P}(\{\text{bankrupt}\}) = 0$
- If one of the distributions has positive expected counter change

Hard part is showing they are necessary

Proof Idea for Case < 1

When $\exists$ a strategy for $\mathcal{P}(\{\text{bankrupt}\}) < 1$?

Easy part – sufficient conditions:

- If $\exists$ a strategy for $\mathcal{P}(\{\text{bankrupt}\}) = 0$
- If one of the distributions has positive expected counter change

Hard part is showing they are necessary

Future Work

- Minimizing questions
- Games
- Refined analysis of the complexity jump
    - Selective Termination: EXPTIME or PSPACE?
    - 1-selective Termination?
- Quantitative case