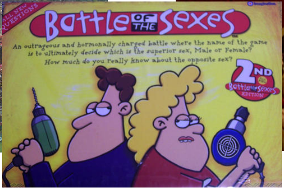
Nash Equilibria in Concurrent Priced Games

Miroslav Klimoš[1], Kim G. Larsen[2], **Filip Štefaňák**[1], Jeppe Thaarup[2]

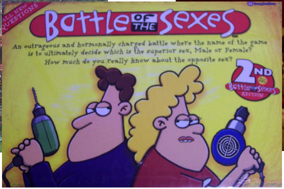
[1]Masaryk University
[2]Aalborg University

LATA
March 8, 2012

Battle of the Sexes



Battle of the Sexes



	F	**O**
F	$(1, 2)$	$(4, 4)$
O	$(6, 6)$	$(2, 1)$

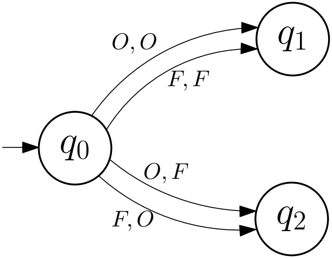
Game Characterization

- games on finite graphs with reachability objectives
- turn-based vs. **concurrent**
  - players take turns vs. *take actions simultaneously*
  - turn-based can be modelled by concurrent games
- zero-sum vs. **non-zero-sum**
  - opposite vs. *independent* objectives
- qualitative vs. **quantitative**
  - binary objectives vs. payoffs or *costs*
- object of study
  - who has a winning strategy vs. *(pure) Nash equilibria*

Overview

1. Priced Concurrent Game Structures
   Nash Equilibria
2. Algorithm for finding Nash equilibria
3. Complexity results

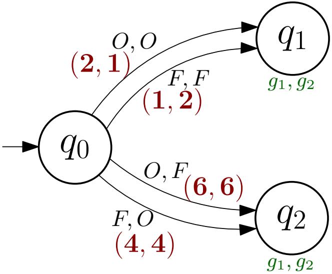
Concurrent Game Structure



Concurrent Game Structure

- K players, set of moves $\mathbb{M}$
- transition function $\delta : Q \times \mathbb{M}^K \to Q$
- δ total and deterministic for *enabled moves*
- a *computation* is a finite or infinite word over $\mathbb{M}^K$

Priced Concurrent Game Structure



Priced Concurrent Game Structure (PCGS)

- K-tuples of nonnegative *prices* on transitions
- goal states (independent for each player)

Preliminaries

Strategy

- $(\mathbb{M}^K)^* \to \mathbb{M}$
- history-dependent strategies, observing history of moves

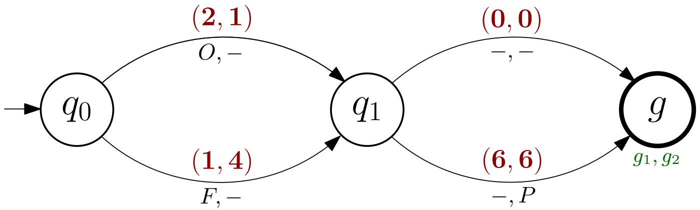
Strategy profile

- a *strategy profile* is a K-tuple of strategies (one for each player)

Cost

- $(\mathbb{M}^K)^* \times \{1 \dots K\} \to \mathbb{N} \cup \infty$
- cumulative price of transitions until the first goal state of the player
- if there is no goal state, the cost is ∞

Example

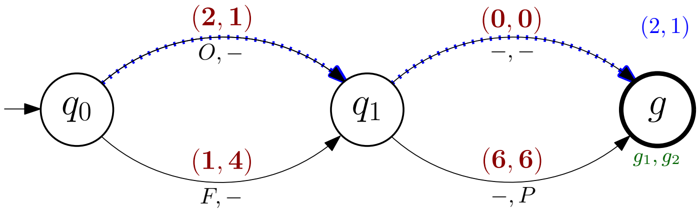


Strategy examples:

$$f_O(\Lambda) = \begin{cases} O & \text{if } \Lambda = \epsilon \\ - & \text{otherwise} \end{cases} \qquad g_P(\Lambda) = \begin{cases} P & \text{if } \Lambda = (F, -) \\ - & \text{otherwise} \end{cases}$$

$$f_F(\Lambda) = \begin{cases} F & \text{if } \Lambda = \epsilon \\ - & \text{otherwise} \end{cases} \qquad g_U(\Lambda) = \begin{cases} - & \text{always} \end{cases}$$

Example



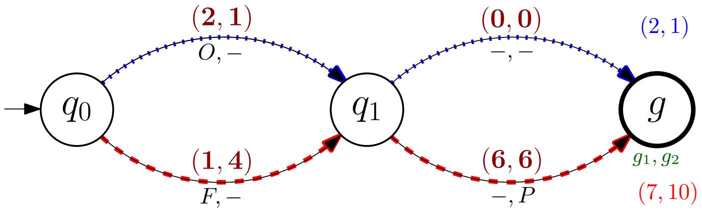
Strategy examples:

$$f_O(\Lambda) = \begin{cases} O & \text{if } \Lambda = \epsilon \\ - & \text{otherwise} \end{cases} \qquad g_P(\Lambda) = \begin{cases} P & \text{if } \Lambda = (F, -) \\ - & \text{otherwise} \end{cases}$$

$$f_F(\Lambda) = \begin{cases} F & \text{if } \Lambda = \epsilon \\ - & \text{otherwise} \end{cases} \qquad g_U(\Lambda) = \begin{cases} - & \text{always} \end{cases}$$

the outcome of (f_O,g_P) is $(O, -)(-, -)$ with costs **(2,1)**

Example



Strategy examples:

$$f_O(\Lambda) = \begin{cases} O & \text{if } \Lambda = \epsilon \\ - & \text{otherwise} \end{cases} \qquad g_P(\Lambda) = \begin{cases} P & \text{if } \Lambda = (F, -) \\ - & \text{otherwise} \end{cases}$$

$$f_F(\Lambda) = \begin{cases} F & \text{if } \Lambda = \epsilon \\ - & \text{otherwise} \end{cases} \qquad g_U(\Lambda) = \begin{cases} - & \text{always} \end{cases}$$

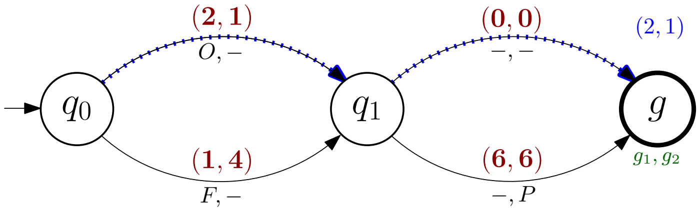
the outcome of (f_O,g_P) is $(O, -)(-, -)$ with costs **(2,1)**
the outcome of (f_F,g_P) is $(F, -)(-, P)$ with costs **(7,10)**

Nash Equilibrium

- a stable strategy profile: no player can lower her cost by changing her strategy
- not necessarily optimal
- may not exist, or there can be more

- bounds vector $\mathbb{B} \in (\mathbb{N} \cup \infty)^K$
- the decision problem: is there a Nash equilibrium satisfying bounds $\mathbb{B}$?
- main problem: find all Nash equilibria satisfying bounds $\mathbb{B}$

Equilibrium Example



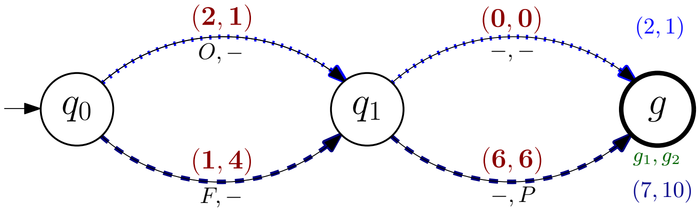
Strategy examples:

$$f_O(\Lambda) = \begin{cases} O & \text{if } \Lambda = \epsilon \\ - & \text{otherwise} \end{cases} \qquad g_P(\Lambda) = \begin{cases} P & \text{if } \Lambda = (F, -) \\ - & \text{otherwise} \end{cases}$$

$$f_F(\Lambda) = \begin{cases} F & \text{if } \Lambda = \epsilon \\ - & \text{otherwise} \end{cases} \qquad g_U(\Lambda) = \begin{cases} - & \text{always} \end{cases}$$

Profile (f_O, g_P) is a Nash equilibrium as no player can reduce their cost. If player 1 uses f_F he gets a lower cost on the first step, but suffers a penalty in the second.

Equilibrium Example



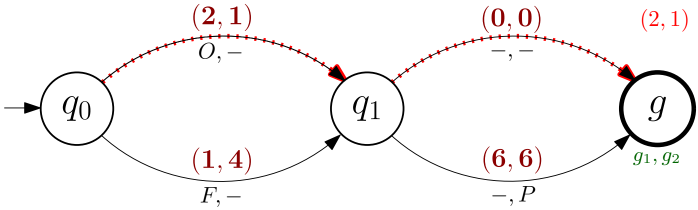
Strategy examples:

$$f_O(\Lambda) = \begin{cases} O & \text{if } \Lambda = \epsilon \\ - & \text{otherwise} \end{cases} \qquad g_P(\Lambda) = \begin{cases} P & \text{if } \Lambda = (F, -) \\ - & \text{otherwise} \end{cases}$$

$$f_F(\Lambda) = \begin{cases} F & \text{if } \Lambda = \epsilon \\ - & \text{otherwise} \end{cases} \qquad g_U(\Lambda) = \begin{cases} - & \text{always} \end{cases}$$

Profile (f_O, g_P) is a Nash equilibrium as no player can reduce their cost. If player 1 uses f_F he gets a lower cost on the first step, but suffers a penalty in the second.

Equilibrium Example



Strategy examples:

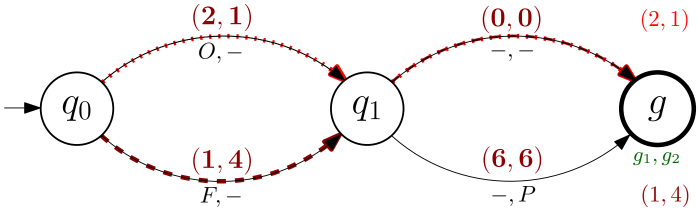
$$f_O(\Lambda) = \begin{cases} O & \text{if } \Lambda = \epsilon \\ - & \text{otherwise} \end{cases} \qquad g_P(\Lambda) = \begin{cases} P & \text{if } \Lambda = (F,-) \\ - & \text{otherwise} \end{cases}$$

$$f_F(\Lambda) = \begin{cases} F & \text{if } \Lambda = \epsilon \\ - & \text{otherwise} \end{cases} \qquad g_U(\Lambda) = \begin{cases} - & \text{always} \end{cases}$$

Profile (f_O, g_P) is a Nash equilibrium as no player can reduce their cost. If player 1 uses f_F he gets a lower cost on the first step, but suffers a penalty in the second.

Profile (f_O, g_U) is not a Nash equilibrium as player 1 can lower his cost by switching to f_F without penalty.

Equilibrium Example



Strategy examples:

$$f_O(\Lambda) = \begin{cases} O & \text{if } \Lambda = \epsilon \\ - & \text{otherwise} \end{cases} \qquad g_P(\Lambda) = \begin{cases} P & \text{if } \Lambda = (F,-) \\ - & \text{otherwise} \end{cases}$$

$$f_F(\Lambda) = \begin{cases} F & \text{if } \Lambda = \epsilon \\ - & \text{otherwise} \end{cases} \qquad g_U(\Lambda) = \begin{cases} - & \text{always} \end{cases}$$

Profile (f_O, g_P) is a Nash equilibrium as no player can reduce their cost. If player 1 uses f_F he gets a lower cost on the first step, but suffers a penalty in the second.

Profile (f_O, g_U) is not a Nash equilibrium as player 1 can lower his cost by switching to f_F without penalty.

Algorithm Overview

How to represent equilibria?

- strategies are infinite - how to represent and characterize them?
- *outcomes* of *all* equilibria form an ω-regular set
- we represent those outcomes by a Büchi automaton

Outcomes are as good as strategies!

- for each equilibrium outcome, we can find a strategy profile

Construction outline

1. we calculate *temptation* and *punishment* values for the game
2. we construct a Büchi automaton accepting outcomes of Nash equilibria

Temptation and Punishment

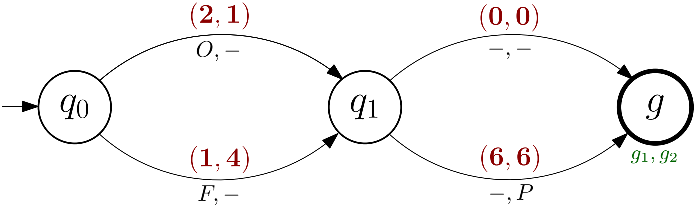
Why an outcome is not an equilibrium?

- a player betrays if she can do better - *temptation*

How to help players resist the temptation?

- strategies of other players try to make this payoff worse - *punishment*
- the punishment is announced in advance and its role is prevention
- strategies can start punishing one step after the defecting step

Punishment Values Example



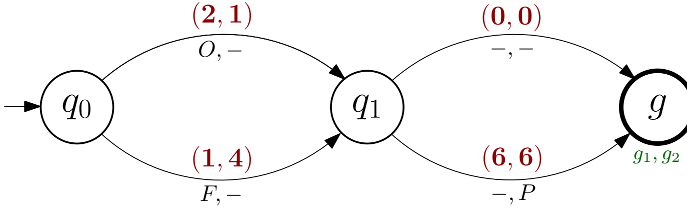
Punishment - $\pi : Q \times \{1 \dots K\} \to \mathbb{N} \cup \infty$

Given a state and a defecting player, *punishment* π is the worst cost the remaining players can enforce for her.

Temptation - $\tau : Q \times \mathbb{M}^K \times \{1 \dots K\} \to \mathbb{N} \cup \infty$

Given a move vector from a state and a player, *temptation* τ for that player is the best cost she can achieve if she defects and chooses another move.

Punishment Values Example



Punishment - $\pi : Q \times \{1 \dots K\} \to \mathbb{N} \cup \infty$

Given a state and a defecting player, *punishment* π is the worst cost the remaining players can enforce for her.

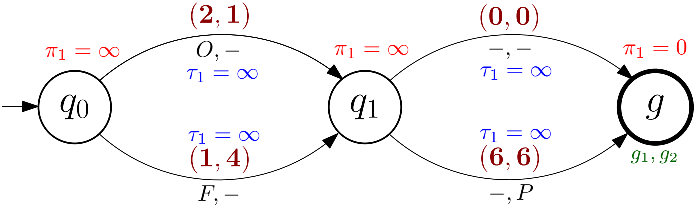
the maximum of temptations from this state

Temptation - $\tau : Q \times \mathbb{M}^K \times \{1 \dots K\} \to \mathbb{N} \cup \infty$

Given a move vector from a state and a player, *temptation* τ for that player is the best cost she can achieve if she defects and chooses another move.

the minimum of punishments to which the player can defect from this transition

Punishment Values Example



Punishment - $\pi : Q \times \{1 \ldots K\} \to \mathbb{N} \cup \infty$

Given a state and a defecting player, *punishment* π is the worst cost the remaining players can enforce for her.

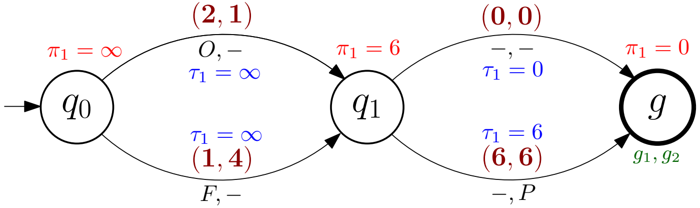
the maximum of temptations from this state

Temptation - $\tau : Q \times \mathbb{M}^K \times \{1 \ldots K\} \to \mathbb{N} \cup \infty$

Given a move vector from a state and a player, *temptation* τ for that player is the best cost she can achieve if she defects and chooses another move.

the minimum of punishments to which the player can defect from this transition

Punishment Values Example



Punishment - $\pi : Q \times \{1 \dots K\} \to \mathbb{N} \cup \infty$

Given a state and a defecting player, *punishment* π is the worst cost the remaining players can enforce for her.

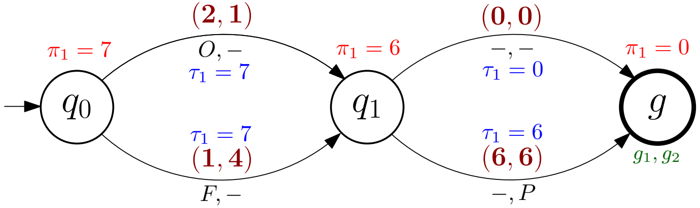
the maximum of temptations from this state

Temptation - $\tau : Q \times \mathbb{M}^K \times \{1 \dots K\} \to \mathbb{N} \cup \infty$

Given a move vector from a state and a player, *temptation* τ for that player is the best cost she can achieve if she defects and chooses another move.

the minimum of punishments to which the player can defect from this transition

Punishment Values Example



Punishment - $\pi : Q \times \{1 \dots K\} \to \mathbb{N} \cup \infty$

Given a state and a defecting player, *punishment* π is the worst cost the remaining players can enforce for her.

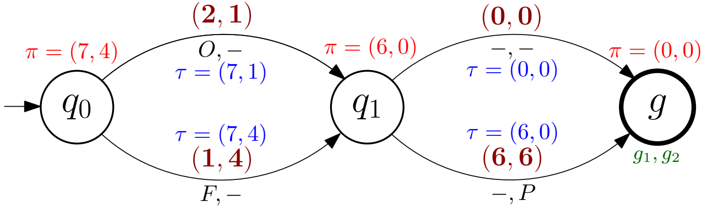
the maximum of temptations from this state

Temptation - $\tau : Q \times \mathbb{M}^K \times \{1 \dots K\} \to \mathbb{N} \cup \infty$

Given a move vector from a state and a player, *temptation* τ for that player is the best cost she can achieve if she defects and chooses another move.

the minimum of punishments to which the player can defect from this transition

Punishment Values Example



Punishment - $\pi : Q \times \{1 \ldots K\} \to \mathbb{N} \cup \infty$

Given a state and a defecting player, *punishment* π is the worst cost the remaining players can enforce for her.

the maximum of temptations from this state

Temptation - $\tau : Q \times \mathbb{M}^K \times \{1 \ldots K\} \to \mathbb{N} \cup \infty$

Given a move vector from a state and a player, *temptation* τ for that player is the best cost she can achieve if she defects and chooses another move.

the minimum of punishments to which the player can defect from this transition

Equilibrium Automaton

Local bounds construction

- with each state, we remember the remaining possible costs
- transitions reduce these costs with their costs and temptations

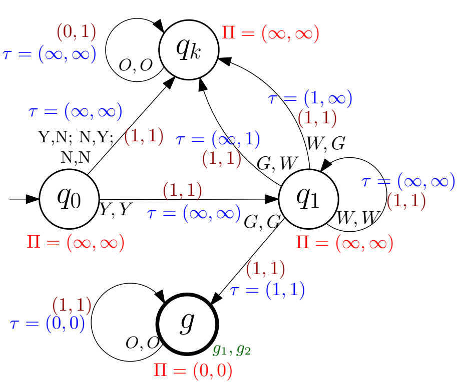
Büchi automaton

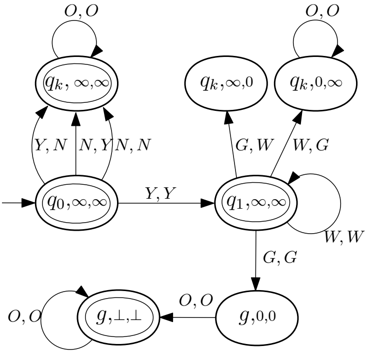
- accepting all equilibria outcomes
- number of states might be exponential

From equilibrium outcomes to strategy profiles

- strategies follow the outcome until someone betrays
- if that happens everybody starts only punishing

Equilibrium Automaton Example





Complexity of the Decision Variant

Recall the decision problem:

- Is there a Nash equilibrium satisfying bounds?

We prove that the decision problem is NP-complete

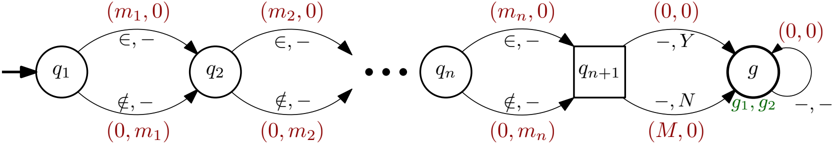
- in NP: Guessing an accepting lasso of polynomial length
- NP-hardness: Reduction from the subset sum problem

turn-based PCGS without bounds

- a Nash equilibrium always exists

NP-hardness

- Reduction from the subset sum problem.
- For input instance $(\{m_1 \ldots m_n\}, m)$ of the Subset sum problem, construct *two-player turn-based* game:

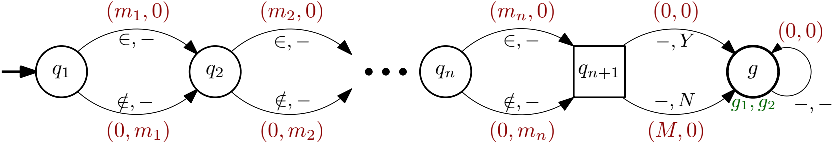


Lemma

There is a Nash Equilibrium satisfying the bounds $(m, M - m)$ if and only if there is a solution to the subset sum problem. $(M = \sum m_i)$

NP-hardness

- Reduction from the subset sum problem.
- For input instance $(\{m_1 ... m_n\}, m)$ of the Subset sum problem, construct *two-player turn-based* game:



Lemma

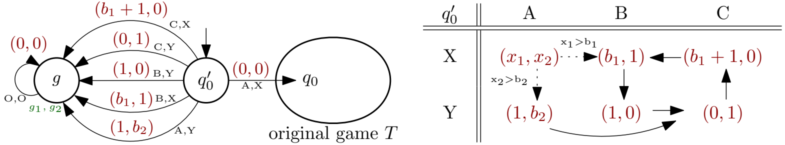
There is a Nash Equilibrium satisfying the bounds $(m, M - m)$ if and only if there is a solution to the subset sum problem. ($M = \sum m_i$)

Player 1 In q_i, choose $\in$ if $m_i \in S'$, otherwise choose $\notin$.

Player 2 In q_{n+1}, choose Y if the accumulated costs so far are $(m, M - m)$, otherwise choose N.

Omitting Bounds

- we reduce the decision problem with bounds to the problem with no bounds
- for two-player PCGS and bounds (b_1, b_2), construct new PCGS



- equilibria satisfying the bounds are preserved
- equilibria not satisfying the bounds are suppressed by added edges
- no new added edge from q'_0 to g is a part of an equilibrium outcome

Complexity overview

Complexity results for the problem of deciding an existence of Nash equilibrium:

- PCGS and with/without bounds $\in$ NP
- Subset sum $\leq$ Two-player turn-based games with bounds $\leq$ PCGS with bounds
- Two-player PCGS with bounds $\leq$ PCGS without bounds

	full PCGS	**turn-based**
with bounds	NP-complete	NP-complete
without bounds	NP-complete	Trivial

Conclusion

1. Priced CGS with individual reachability objectives
   - Non-negative integer prices on transitions for each player
   - Cost for a player is accumulated sum of prices before reaching their goal state

2. Characterization of Nash Equilibria in PCGS
   - Set of Nash equilibrium outcomes is ω-regular language
   - We can extract strategies for these outcomes

3. Complexity of the decision variant of the problem
   - NP-complete problem

4. Ideas for future work
   - Mixed (probabilistic) strategies
   - Partial observability
   - Negative costs