# Measuring Permissiveness in Parity Games: Mean-Payoff Parity Games Revisited

Michael Ummels

ummels@lsv.ens-cachan.fr

LSV, CNRS & ENS Cachan

(Joint work with Patricia Bouyer, Nicolas Markey and Jörg Olschewski)

ATVA 2011

12 October 2011

# Motivation

Two-player parity games for verification/synthesis of reactive systems

## Setting (old)

- System (Player 1) vs. Environment (Player 2)
- Strategy for Player 1 ≡ Controller for the system
- Controller's behaviour is deterministic

## Setting (new)

- nondeterministic / multi strategies
- nondeterministic strategy ≡ abstract implementation
- measure of permissiveness

# Mean-Penalty Parity Games

## Definition

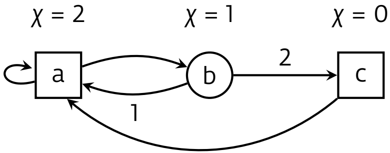
weighted game graph $G = (Q_1, Q_2, E, \text{weight})$

- $Q_1 \cup Q_2 = Q$ nodes of Player 1 (circle) and 2 (square)
- $E \subseteq Q \times Q$ edge relation (non-terminating)
- $\text{weight}\colon E \to \mathbb{Q}$ assigns a weight to every edge

priority function $\chi\colon Q \to \mathbb{N}$ (min-parity condition)

mean-penalty parity game $\mathcal{G} = (G, \chi)$

## Example

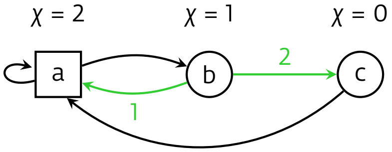


# Strategies and Multi-Strategies

## Definition

- (deterministic) strategy $\sigma\colon Q^* Q_1 \to Q$ such that $\sigma(\gamma q) \in qE$ maps a history to a successor node
- multi-strategy $\sigma\colon Q^* Q_1 \to 2^Q$ such that $\varnothing \neq \sigma(\gamma q) \subseteq qE$ maps a history to a set of successor nodes

## Example



# Penalties

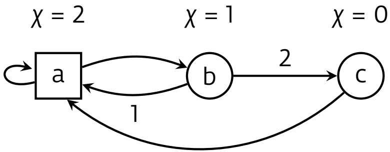
Idea: Every blocked edge adds a certain amount of penalty.
Penalties are accumulated along a play.

Given a multi-strategy $\sigma$ for Player 1, define the total penalty for $\gamma \in Q^*$:

- $\text{penalty}_\sigma(\varepsilon) \coloneqq 0$
- if $\gamma q \in Q^* Q_1$:
  - $B \coloneqq \{(q, q') \in E : q' \notin \sigma(\gamma q)\}$
  - $\text{penalty}_\sigma(\gamma q) \coloneqq \text{penalty}_\sigma(\gamma) + \sum_{e \in B} \text{weight}(e)$
- if $\gamma q \in Q^* Q_2$:
  - $\text{penalty}_\sigma(\gamma q) \coloneqq \text{penalty}_\sigma(\gamma)$

# Total Penalty

## Example




Consider multi-strategy $\sigma$ defined by $\sigma(\gamma b) = \begin{cases} \{c\} & \text{if } |\gamma|_b \text{ is even,} \\ \{a, c\} & \text{if } |\gamma|_b \text{ is odd.} \end{cases}$

- $\text{penalty}_\sigma(\varepsilon) = 0$
- $\text{penalty}_\sigma(b) = 1$
- $\text{penalty}_\sigma(bcab) = 1$
- $\text{penalty}_\sigma(bcabab) = 2$
- $\text{penalty}_\sigma(bcababcab) = 2$

- $\frac{1}{1}\text{penalty}_\sigma(b) = 1$
- $\frac{1}{4}\text{penalty}_\sigma(bcab) = 0.25$
- $\frac{1}{6}\text{penalty}_\sigma(bcabab) = 0.333\ldots$
- $\frac{1}{9}\text{penalty}_\sigma(bcababcab) = 0.222\ldots$
- $\cdots$

# Mean Penalty

- For an infinite play $\rho = q_0 q_1 \ldots$, the mean penalty is defined by

$$\text{penalty}_\sigma(\rho) \coloneqq \begin{cases} \limsup_{n\to\infty} \frac{1}{n} \text{penalty}_\sigma(q_0 \ldots q_{n-1}) & \text{if } \rho \text{ is winning,} \\ +\infty & \text{otherwise.} \end{cases}$$

- Then the mean penalty of $\sigma$ at $q$ is defined as

$$\text{penalty}(\sigma, q) \coloneqq \sup\{\text{penalty}_\sigma(\rho) : \rho \text{ is outcome of } \sigma \text{ from } q\}.$$

- The value of a node $q$ is

$$\text{val}(q) = \inf_\sigma \text{penalty}(\sigma, q).$$

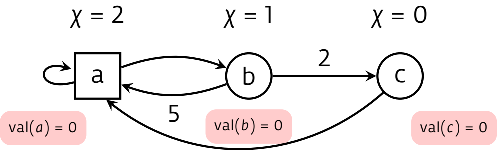
# Values and optimal strategies

## Definition

The value problem is the decision problem:

Given $\mathcal{G} = (G, \chi)$, $q \in Q$, and $x \in \mathbb{Q}$, decide whether $\mathrm{val}(q) \le x$.

## Example




Questions:

- Complexity of the value problem
- Optimal strategies

# Solving Mean-Penalty Parity Games

How to solve penalty games? Reduction to mean-payoff parity games.

## Definition

A mean-payoff parity game is a pair $\mathcal{G} = (G, \chi)$. The payoff of an infinite play $\rho = q_0 q_1 \ldots$ (for Player 1) is

$$\text{payoff}(\rho) = \begin{cases} \liminf_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \text{weight}(q_i, q_{i+1}) & \text{if } \rho \text{ is winning,} \\ -\infty & \text{otherwise.} \end{cases}$$

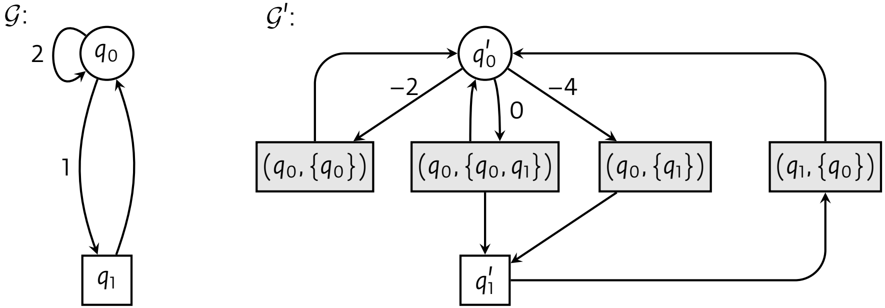
The value of a node $q \in Q$ is $\text{val}(q) = \sup_\sigma \inf_\tau(\text{payoff}(\rho(\sigma, \tau, q)))$, where $\sigma$ ranges over Player 1 strategies and $\tau$ ranges over Player 2 strategies.

## Theorem (Chatterjee-Doyen 2010)

The value problem for mean-payoff parity games is in NP ∩ coNP.

# Reduction

From a mean-penalty parity game $\mathcal{G}$, we can construct a mean-payoff parity game $\mathcal{G}'$ such that $\text{val}(q) = -\text{val}(q')$ for every $q \in Q$.




For every $q \in Q$ and every nonempty subset $P$ of successors, introduce a new node $(q, P)$.

Note: The new game graph is of exponential size.

# Optimal Strategies

## Theorem (Chatterjee-Henzinger-Jurdziński, BMOU)

Mean-payoff parity games are *determined* with optimal strategies.

- Optimal strategies for Player 1 require infinite memory.
- Player 2 has a memoryless optimal strategy.

The reduction leads directly to:

## Corollary

In every mean-penalty parity game, Player 1 has an optimal multi-strategy.
Optimal multi-strategies require infinite memory.

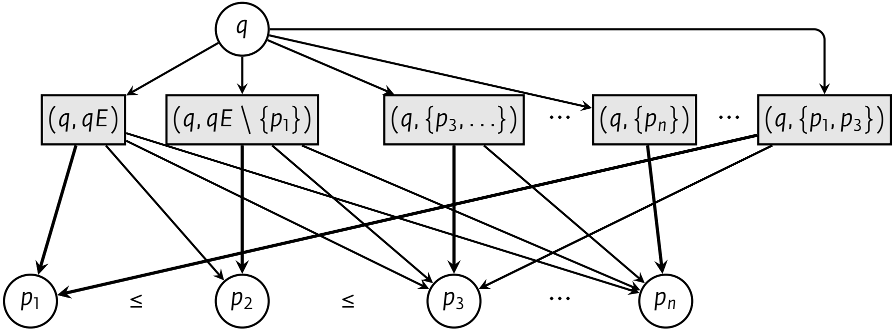
But for the value problem, we need some more work.

# Simple Strategies for Player 2

The new game graph is of exponential size, but:

**Lemma**

Player 2 has an optimal simple strategy $\tau'$ in $\mathcal{G}'$, i.e. it is memoryless and for every $q \in Q$ there is a total order $\leq_q$ on $qE$ with $\tau'(q, P) = \min_{\leq_q} P$.



# Membership in coNP

Given a simple strategy $\tau$, we only need to construct a small fragment of $\mathcal{G}'$ to determine the maximal payoff for Player 1 (states $(q, P)$ where $P$ is not closed under greater elements wrt. $<_q$ can be omitted).

## Theorem

The value problem for mean-penalty parity games is in coNP.

## Proof.

1. Guess a simple strategy for Player 2.
2. Construct $\mathcal{G}'$ from $\mathcal{G}$ but omit all unnecessary states $\rightsquigarrow \mathcal{G}''$
3. $\mathcal{G}''$ is a mean-payoff parity game without Player 2 states.
4. Check in $\mathcal{G}''$ whether Player 1 has a strategy $\sigma$ with $\mathrm{val}(\sigma, q) \geq -x$ (can be done in polynomial time). $\square$

# Membership in NP

Actually, we can do better…

**Lemma**

There is a polynomial-time reduction from mean-penalty parity games to mean-payoff parity games.

Idea: For each transition,

- Player 1 has to decide whether to enable transition;
- Player 2 has to decide whether he wishes to explore transition.

Game proceeds along the *last* enabled transition Player 2 wants to explore.

**Theorem**

The value problem for mean-penalty parity games is in NP.

# A Deterministic Algorithm

We propose a *McNaughton style* deterministic algorithm for solving mean-payoff parity games.

**Theorem**

The values of a mean-payoff parity game with $d$ priorities can be computed in time $O(|Q|^{d+2} \cdot |E| \cdot W)$.

Faster than the previously best algorithm, which has running time $O(|Q|^{d+4} \cdot |E| \cdot d \cdot W)$ and can only compare values to a threshold.

By applying the algorithm *symbolically* on the (exponential) mean-payoff parity game derived from a mean-penalty parity game, we get:

**Theorem**

The values of a mean-penalty parity game with $d$ priorities can be computed in time $O(|Q|^{d+3} \cdot |E| \cdot W)$.

# Conclusion

Summary:

- New model for measuring permissiveness in games
- Reduction from mean-penalty parity to mean-payoff parity games
- Existence of optimal multi-strategies
- The value problem for mean-penalty parity games is in NP ∩ coNP
- Mean-penalty parity games can be solved in $O(|Q|^{d+3} \cdot |E| \cdot W)$
- New algorithm for solving mean-payoff parity games

Outlook:

- Synthesis of almost optimal strategies
- Extension to stochastic games