Mean-payoff games with incomplete information

Paul Hunter, Guillermo Pérez, Jean-François Raskin

Université Libre de Bruxelles
COST Meeting @ Madrid

October, 2013
Outline

1 MPG variations
   - Mean-payoff games
   - Imperfect information

2 Tackling MPGs with imperfect information
   - Incomplete information
   - Observable determinacy
   - Decidable subclasses
   - Pure games with incomplete information

3 Conclusions
MPGs imperfect information: example

- The game involves nodes 1, 2, 3, and 4 with edges connecting them.

- The game's objective is to move a token to win by maximizing the average weight of the edges traversed.

- Example: Player ve chooses node a, while player dam chooses edge (1, a, 2) with a payoff of -1.
MPGs imperfect information: example

\[ \Sigma = \{a, b\} \]

and weights on the edges

Game to move token: \( \exists ve \) chooses \( \sigma \) and \( \forall dam \) chooses edge to win (\( \exists ve \)): maximize average weight of edges traversed

Example: \( \exists ve \) chooses \( a \), \( \forall dam \) chooses \( (1, a, 2) \); payoff = -1

\[
\begin{array}{ccc}
1 & 2 & 3 \\
2 & 4 & 1 \\
3 & 3 & 2 \\
4 & 4 & 3
\end{array}
\]

P. Hunter, G. Pérez, J.F. Raskin (ULB)
MPGs imperfect information: example

- \( \Sigma = \{a, b\} \) and weights on the edges

![Diagram]

- \( \Sigma, -1 \)
- \( a, -1 \) from 1 to 2
- \( b, -1 \) from 2 to 4
- \( b, -1 \) from 3 to 1
- \( a, -1 \) from 4 to 3
- \( \Sigma, +1 \)

Example: \( \exists \) ve chooses \( a \), \( \forall \) dam chooses \( (1, a, 2) \); payoff = -1
MPGs imperfect information: example

- $\Sigma = \{a, b\}$ and weights on the edges

**Game**
- to move token: $\exists$ve chooses $\sigma$ and $\forall$dam chooses edge
- to win ($\exists$ve): maximize average weight of edges traversed

```
\[\begin{array}{c}
\text{Node 1} \quad \text{Node 2} \quad \text{Node 3} \quad \text{Node 4} \\
\Sigma, -1 \quad a, -1 \quad b, -1 \quad \Sigma, +1 \\
\Sigma, -1 \quad \Sigma, -1 \quad a, -1 \quad \Sigma, +1 \\
\end{array}\]
```
MPGs imperfect information: example

- $\Sigma = \{a, b\}$ and weights on the edges
- Game
  - to move token: $\exists$ve chooses $\sigma$ and $\forall$dam chooses edge
  - to win ($\exists$ve): maximize average weight of edges traversed
- Example: $\exists$ve chooses $a$, $\forall$dam chooses $(1, a, 2)$; payoff = -1
MPGs imperfect information: example

- $\Sigma = \{a, b\}$ and weights on the edges
- Game
  - to move token: $\exists$ve chooses $\sigma$ and $\forall$dam chooses edge
  - to win ($\exists$ve): maximize average weight of edges traversed
- Example: $\exists$ve chooses $a$, $\forall$dam chooses $(1, a, 2)$; payoff = -1

![Graph]

$\Sigma$, -1

$\Sigma$, +1

1

3

4

$\Sigma$, -1

$\Sigma$, -1

$\Sigma$, -1
MPGs imperfect information: example

- $\Sigma = \{a, b\}$ and weights on the edges
- Game
  - to move token: $\exists v e$ chooses $\sigma$ and $\forall d a m$ chooses edge
  - to win ( $\exists v e$ ): maximize average weight of edges traversed

![Graph Diagram]
MPGs imperfect information: example

- $\Sigma = \{a, b\}$ and weights on the edges
- Game
  - to move token: $\exists$ve chooses $\sigma$ and $\forall$dam chooses edge
  - to win ( $\exists$ve ): maximize average weight of edges traversed
- $\exists$ve only sees colors, $\forall$dam sees everything

```
\begin{itemize}
  \item $\Sigma = \{a, b\}$ and weights on the edges
  \item Game
    \begin{itemize}
      \item to move token: $\exists$ve chooses $\sigma$ and $\forall$dam chooses edge
      \item to win ( $\exists$ve ): maximize average weight of edges traversed
    \end{itemize}
  \item $\exists$ve only sees colors, $\forall$dam sees everything
\end{itemize}
```
Mean-payoff game

Definition (MPGs)

- **Mean-payoff games** are 2-player games of infinite duration played on (directed) weighted graphs. **∃ve** chooses an action, and **∀dam** resolves non-determinism by choosing the next state.

- **∃ve** wants to maximize the average weight of the edges traversed (the **MP value**).

- **∀dam** wants to minimize the same value.
Definition (Strategies for $\exists v_e$)

An observable strategy for $\exists v_e$ is a function from finite sequences $(Obs \cdot \Sigma)^* Obs$ to the next action.
Definition (Strategies for $\exists$ve)

An observable strategy for $\exists$ve is a function from finite sequences $(\text{Obs} \cdot \Sigma)^* \text{Obs}$ to the next action.

Definition (MP value)

Given the transition relation $\Delta$ and the weight function $w : \Delta \to \mathbb{Z}$ of a MPG, the MP value is $\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} w(q_i, \sigma_i, q_{i+1})$. 
Strategies, Mean-payoff value

Definition (Strategies for $\exists$ve )
An observable strategy for $\exists$ve is a function from finite sequences $(Obs \cdot \Sigma)^* Obs$ to the next action.

Definition (MP value)
Given the transition relation $\Delta$ and the weight function $w : \Delta \rightarrow \mathbb{Z}$ of a MPG, the MP value is $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} w(q_i, \sigma_i, q_{i+1})$.

Problem (Winner of a MPG)
Given a threshold $\nu \in \mathbb{N}$, the MPG is won by $\exists$ve iff $MP \geq \nu$. W.l.o.g assume $\nu = 0$. 
Theorem (Ehrenfeucht and Mycielski [1979])

- **MPGs are determined**, i.e. if $\exists \text{eve}$ doesn’t have a winning strategy then $\forall \text{dam}$ does (and vice versa).

- **Positional strategies suffice for either $\forall \text{dam}$ or $\exists \text{eve}$ to win a MPG.**
Theorem (Ehrenfeucht and Mycielski [1979])

- **MPGs are determined**, i.e. if $\exists$ve doesn’t have a winning strategy then $\forall$dam does (and vice versa).
- **Positional strategies suffice for either $\forall$dam or $\exists$ve to win a MPG.**

$\Sigma = \{a, b\}$

![Diagram](image-url)
Theorem (Ehrenfeucht and Mycielski [1979])

- **MPGs are determined**, i.e. if $\exists ve$ doesn’t have a winning strategy then $\forall dam$ does (and vice versa).
- Positional strategies suffice for either $\forall dam$ or $\exists ve$ to win a MPG.

$\Sigma = \{a, b\}$ $\exists ve$ has a winning strat: play $b$ in 2 and $a$ in 3
Outline

1. **MPG variations**
   - Mean-payoff games
   - Imperfect information

2. **Tackling MPGs with imperfect information**
   - Incomplete information
   - Observable determinacy
   - Decidable subclasses
   - Pure games with incomplete information

3. **Conclusions**
Definition (MPGs with imperfect info.)

A MPG with imperfect information is played on a weighted graph given with a coloring of the state space that defines equivalence classes of indistinguishable states (observations).
A MPG with imperfect information is played on a weighted graph given with a coloring of the state space that defines equivalence classes of indistinguishable states (observations).

\[ \Sigma = \{a, b\} \]
Definition (MPGs with imperfect info.)

A MPG with imperfect information is played on a weighted graph given with a coloring of the state space that defines equivalence classes of indistinguishable states (observations).

\[ \Sigma = \{ a, b \} \]

Neither \( \exists \text{eve} \) nor \( \forall \text{dam} \) have a winning strategy anymore.
Motivation and properties

Why consider such a model?

- MPGs are natural models for systems where we want to optimize the limit-average usage of a resource.
- Imperfect information arises from the fact that most systems have a limited amount of sensors and input data.
Motivation and properties

Why consider such a model?
- MPGs are natural models for systems where we want to optimize the limit-average usage of a resource.
- Imperfect information arises from the fact that most systems have a limited amount of sensors and input data.

Theorem (Degorre et al. [2010])
- **MPGs with imperfect info. are no longer “determined”**.
- **∃Eve learns about the game by using memory.**
- **Determining who wins is undecidable.**
- **May require infinite memory to be won by ∃Eve.**
1. MPG variations
   - Mean-payoff games
   - Imperfect information

2. Tackling MPG with imperfect information
   - Incomplete information
   - Observable determinacy
   - Decidable subclasses
   - Pure games with incomplete information

3. Conclusions
Don’t lie to Eve

Definition

A game of imperfect information is of **incomplete information** if for every \((q, \sigma, q') \in \Delta\), then for every \(s'\) in the same observation as \(q'\) there is a transition \((s, \sigma, s') \in \Delta\) where \(s\) is in the same observation as \(q\).
Definition

A game of imperfect information is of incomplete information if for every \((q, \sigma, q') \in \Delta\), then for every \(s'\) in the same observation as \(q'\) there is a transition \((s, \sigma, s') \in \Delta\) where \(s\) is in the same observation as \(q\).
Don’t lie to Eve

Lemma (imperfect to incomplete info.)

*imperfect information can be turned into incomplete information with a possible exponential blow-up (via its knowledge-based subset construction).*
Outline

1. MPG variations
   - Mean-payoff games
   - Imperfect information

2. Tackling MPGs with imperfect information
   - Incomplete information
   - Observable determinacy
   - Decidable subclasses
   - Pure games with incomplete information

3. Conclusions
Observe that in an MPG of incomplete information:

1. the view $\exists v e$ has of a play in the game is $o_0 \sigma_0 o_1 \sigma_1 \ldots$, 

2. given current $o_i$ the game could be in any $q \in o_i$ (not true in imperfect information), 

3. $\forall d a m$ can have a two step strategy: choose observations first, 

4. "delay" the specific choice of states for later!
Observe that in an MPG of incomplete information:

1. the view $\exists v_e$ has of a play in the game is $o_0\sigma_0o_1\sigma_1\ldots$,
2. given current $o_i$ the game could be in any $q \in o_i$ (not true in imperfect information),
Observe that in an MPG of incomplete information:

1. the view $\exists v_e$ has of a play in the game is $o_0 \sigma_0 o_1 \sigma_1 \ldots$,
2. given current $o_i$ the game could be in any $q \in o_i$ (not true in imperfect information),
3. $\forall d$ can have a two step strategy: choose observations first,
Observe that in an MPG of incomplete information:

1. the view $\exists v e$ has of a play in the game is $o_0 o_0 o_1 o_1 \ldots$,
2. given current $o_i$ the game could be in any $q \in o_i$ (not true in imperfect information),
3. $\forall \text{adam}$ can have a two step strategy: choose observations first,
4. “delay” the specific choice of states for later!
**Observable strategies:** we let ∀dam reveal to ∃ve only the \((Obs \times \Sigma)^+ \mapsto Obs\) version of his strategy.

Let \(\gamma\) be a function mapping observation-action sequences to concrete state-action ones.
∀dam and determinacy

**Definition**

- **Observable strategies**: we let ∀dam reveal to ∃ve only the \((\text{Obs} \times \Sigma)^+ \mapsto \text{Obs}\) version of his strategy.
- Let \(\gamma\) be a function mapping observation-action sequences to concrete state-action ones.

**Definition (New winning condition)**

Let \(\psi\) be a play in the game. ∃ve wins if all paths in \(\gamma(\psi)\) are winning for her. ∀dam wins if there is some path which is winning for him.
∀dam and determinacy

Definition
- Observable strategies: we let ∀dam reveal to ∃ve only the $(Obs \times \Sigma)^+ \mapsto Obs$ version of his strategy.
- Let $\gamma$ be a function mapping observation-action sequences to concrete state-action ones.

Definition (New winning condition)
Let $\psi$ be a play in the game. ∃ve wins if all paths in $\gamma(\psi)$ are winning for her. ∀dam wins if there is some path which is winning for him.

Theorem (Observable determinacy)
The new winning condition is a projection of the perfect information game winning condition (via $\gamma$). The new winning condition is coSuslin and hence determined*. 
Outline

1 MPG variations
   - Mean-payoff games
   - Imperfect information

2 Tackling MPGs with imperfect information
   - Incomplete information
   - Observable determinacy
   - Decidable subclasses
   - Pure games with incomplete information

3 Conclusions
Definition (Function sequence classification)

A function sequence is **good** (**bad**) if a function is pointwise bigger or equal (smaller) than a previous one – same observation.

**obs:** blue

**play:** $f_i$

**cur. f:** $f_i(1) = 0$
Definition (Function sequence classification)

A function sequence is **good** (bad) if a function is pointwise bigger or equal (smaller) than a previous one – same observation.

**obs:** blue-a-yellow

**play:** $f_1 \; a \; f_1$

**cur. f:** $f_1(2) = -3, \; f_1(3) = -1$
Definition (Function sequence classification)

A function sequence is **good** (**bad**) if a function is pointwise bigger or equal (smaller) than a previous one – same observation.

**Diagram:**

- Node 1: $\Sigma, -3$
  - Edge: $a, -1$
  - Edge: $b, -1$
  - Edge: $\Sigma, -1$

- Node 2: $\Sigma, -1$

- Node 3: $\Sigma, -1$

- Node 4: $\Sigma, +1$

**Observation:** blue-a-yellow-b-green

**Play:** $f_1 \ a \ f_1 \ b \ f_2$

**Current Function:** $f_2(4) = -4$
Function-Reachability game

Definition (Function sequence classification)

A function sequence is good (bad) if a function is pointwise bigger or equal (smaller) then a previous one – same observation.

**obs:** blue-a-yellow-b-green-a-green

**play:** $f_1 \ a \ f_1 \ b \ f_2 \ a \ f_3$ **good**

**cur. f:** $f_3(4) = -3$
Unfolding a MPG with incomplete information

“Unfold” $G$, stop when a good or bad sequence is reached.

- We are left with a new reachability game
- Not all branches will be labelled...
Let $H$ be the reachability game played on the unfolding of $G$,

**Theorem (Strategy transfer for $\exists$ve )**

$\exists$ve has a finite memory winning strategy in $G$ if and only if she has a winning strategy in $H$.

**Theorem (Strat. transfer for $\forall$dam )**

If $\forall$dam has a winning observable strategy in $H$ then he also has a winning strategy in $G$. 
Finite memory, Adeq. Pure, Pure games

All based on function sequences (branches) of the associated reachability game $H$.

**Definition**

1. **Finite memory games**: $\exists \text{eve}$ can force good leaves or $\forall \text{adam}$ can force bad leaves.

2. **Adequately pure games**: $\exists \text{eve}$ ($\forall \text{adam}$) can force good (bad) branches where all but 2 functions have different support.

3. **Pure games [structural]**: the unfolding of $G$ is finite and in all branches, all but 2 functions have different support.
Finite memory, Adeq. Pure, Pure games

All based on function sequences (branches) of the associated reachability game $H$.

**Definition**

1. **Finite memory games**: $\exists \text{eve}$ can force good leaves or $\forall \text{dam}$ can force bad leaves.

2. **Adequately pure games**: $\exists \text{eve} \ (\forall \text{dam})$ can force good (bad) branches where all but 2 functions have different support.
Finite memory, Adeq. Pure, Pure games

All based on function sequences (branches) of the associated reachability game $H$.

**Definition**

1. **Finite memory games**: $\exists$ve can force good leaves or $\forall$dam can force bad leaves.

2. **Adequately pure games**: $\exists$ve ($\forall$dam) can force good (bad) branches where all but 2 functions have different support.

3. **Pure games [structural]**: the unfolding of $G$ is finite and in all branches, all but 2 functions have different support.
Let $A$ be a class of MPGs with incomplete (or imperfect) information. Given MPG with incomplete (imperfect) information $G$,

**Problem (Class membership)**

*Is $G$ a member of $A$?*

**Problem (Winner determination)**

*Does $\exists$ve have a winning strategy in $G$?*
### Summary

<table>
<thead>
<tr>
<th>Information</th>
<th>Finite memory</th>
<th>Adequately pure</th>
<th>Pure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class-membership</td>
<td>Undec(^1)</td>
<td>PSPACE-complete</td>
<td>NEXP-hard, in EXPSPACE</td>
</tr>
<tr>
<td>Winner-det.</td>
<td>R-c</td>
<td>PSPACE-complete</td>
<td>EXP-complete</td>
</tr>
</tbody>
</table>

\(^1\)gray=Degorre et al. [2010], other colors are new results
Outline

1. MPG variations
   - Mean-payoff games
   - Imperfect information

2. Tackling MPG with imperfect information
   - Incomplete information
   - Observable determinacy
   - Decidable subclasses
   - Pure games with incomplete information

3. Conclusions
Does $\exists$ve win pure $G$?

Theorem

Deciding if $\exists$ve has a winning strategy in a given pure MPG with incomplete information is in $\text{NP} \cap \text{coNP}$.

Based on Björklund et al. [2004].

Observe* that positional strategies suffice for $\exists$ve to win pure games with incomplete information.
The class membership problem for pure games with incomplete information is coNP-complete.

**Proof.**

- One can “guess” a branch in $H$ (of size at most $|\text{Obs}| + 1$) and in polynomial time check that it is neither good nor bad.
- For hardness we reduce from the HAMILTONIAN-CYCLE problem.
HAM-CYCLE as an MPG

\[ q_1 \rightarrow \Sigma, 0 \rightarrow q_-, q_+ \rightarrow v_0, +1 \rightarrow v_1, +1 \rightarrow v_2, +1 \rightarrow \ldots \rightarrow v_{n-2}, +1 \rightarrow v_{n-1}, +1 \rightarrow v_n, +1 \rightarrow \tau, 0 \rightarrow \tau, -1 \rightarrow \Sigma, -n \rightarrow \Sigma, 0 \rightarrow q_1 \]

P. Hunter, G. Pérez, J.F. Raskin (ULB)  
MPGs with incomplete info.  
October, 2013  
26 / 28
### Summary

1. **Done:** incomplete info., observable determinacy, subclasses
2. **Cooking:** other asymmetric information types, other quantitative games, mixed strategies

<table>
<thead>
<tr>
<th></th>
<th>Finite memory</th>
<th>Adequately pure</th>
<th>Pure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information in</td>
<td>incomplete</td>
<td>imperfect</td>
<td>incomplete</td>
</tr>
<tr>
<td>Finite memory</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class-membership</td>
<td>Undec⁷</td>
<td>PSPACE-complete</td>
<td>NEXP-hard, in EXPSPACE</td>
</tr>
<tr>
<td>Winner-det.</td>
<td>R-c</td>
<td>PSPACE-complete</td>
<td>EXP-complete</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

¹ gray = Degorre et al. [2010], other colors are new results
References I


