Mean-payoff games with incomplete information

Paul Hunter, Guillermo Pérez, Jean-François Raskin

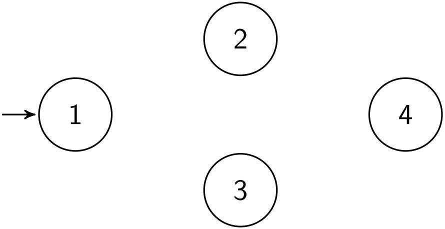
Université Libre de Bruxelles
COST Meeting @ Madrid

October, 2013

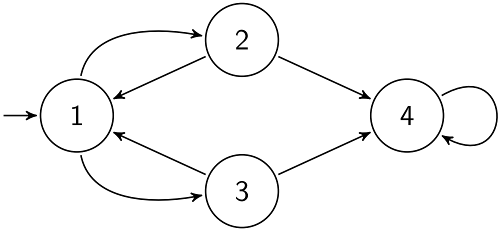
Outline

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

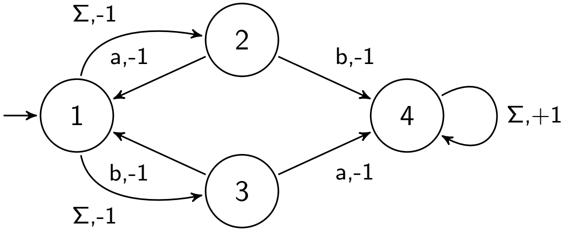


MPGs imperfect information: example



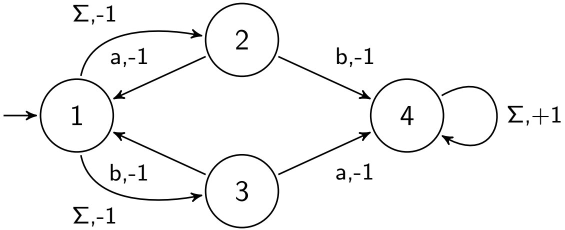
MPGs imperfect information: example

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



MPGs imperfect information: example

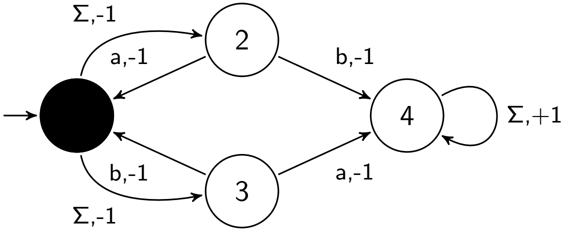
- $\Sigma = \{a, b\}$ and weights on the edges
- Game
  - to move token: $\exists$ve chooses σ and $\forall$dam chooses edge
  - to win ($\exists$ve): maximize average weight of edges traversed



UNIVERSITÉ LIBRE DE BRUXELLES

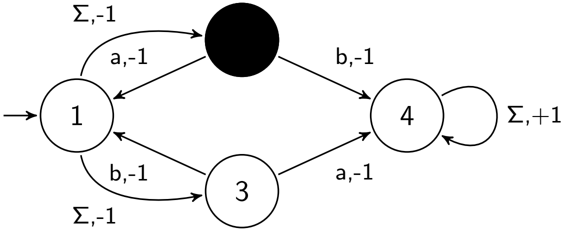
MPGs imperfect information: example

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



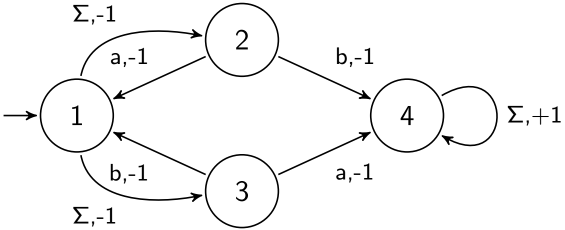
MPGs imperfect information: example

- $\Sigma = \{a, b\}$ and weights on the edges
- Game
  - to move token: $\exists$ve chooses σ 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



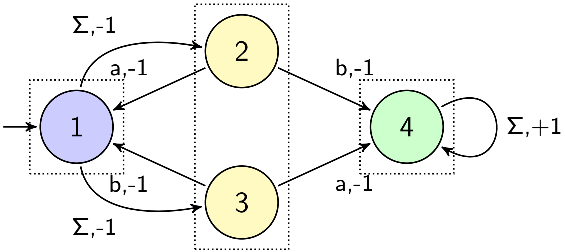
MPGs imperfect information: example

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



MPGs imperfect information: example

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



ULB UNIVERSITÉ LIBRE DE BRUXELLES

Mean-payoff game

Definition (MPGs)

- Mean-payoff games are 2-player games of infinite duration played on (directed) weighted graphs. $\exists$ve chooses an action, and $\forall$dam resolves non-determinism by choosing the next state.
- $\exists$ve wants to maximize the average weight of the edges traversed (the MP value).
- $\forall$dam wants to minimize the same value.

Strategies, Mean-payoff value

Definition (Strategies for ∃ve)

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

Strategies, Mean-payoff value

Definition (Strategies for ∃ve)

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

Definition (MP value)

Given the transition relation Δ and the weight function $w : \Delta \mapsto \mathbb{Z}$ 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 ∃ve)

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

Definition (MP value)

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

Problem (Winner of a MPG)

Given a threshold $\nu \in \mathbb{N}$, the MPG is won by ∃ve iff $MP \geq \nu$. W.l.o.g assume $\nu = 0$.

UNIVERSITÉ LIBRE DE BRUXELLES

MPGs

Theorem (Ehrenfeucht and Mycielski [1979])

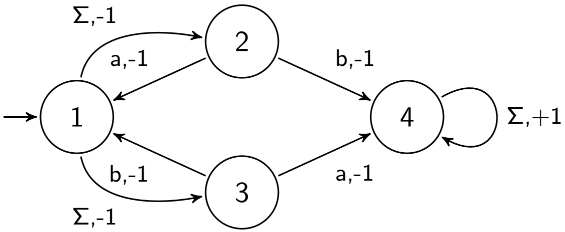
- *MPGs are determined, i.e. if $\exists$ve doesn't have a winning strategy then $\forall$dam does (and viceversa).*
- *Positional strategies suffice for either $\forall$dam or $\exists$ve to win a MPG.*

MPGs

Theorem (Ehrenfeucht and Mycielski [1979])

- *MPGs are determined, i.e. if ∃ve doesn't have a winning strategy then ∀dam does (and viceversa).*
- *Positional strategies suffice for either ∀dam or ∃ve to win a MPG.*

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

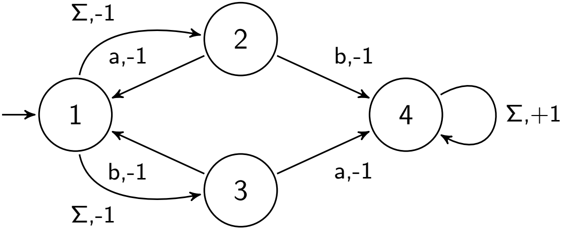


MPGs

Theorem (Ehrenfeucht and Mycielski [1979])

- *MPGs are determined, i.e. if ∃ve doesn't have a winning strategy then ∀dam does (and viceversa).*
- *Positional strategies suffice for either ∀dam or ∃ve to win a MPG.*

$\Sigma = \{a, b\}$ ∃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

MPG with imperfect information

Definition (MPGs with imperfect info.)

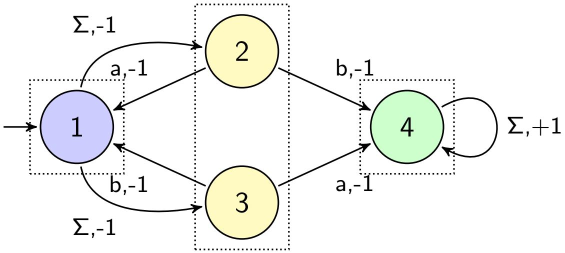
A MPG with imperfect information is played on a weighted graph given with a coloring of the state space that defines equivalence classes of indistinguishable states (observations).

MPG with imperfect information

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).

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

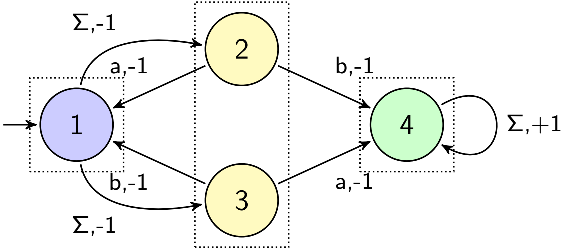


MPG with imperfect information

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 ∃ve nor ∀dam have a winning strategy anymore



UNIVERSITÉ LIBRE DE BRUXELLES

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".*
- *∃ve learns about the game by using memory.*
- *Determining who wins is undecidable.*
- *May require infinite memory to be won by ∃ve .*

UNIVERSITÉ LIBRE DE BRUXELLES

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

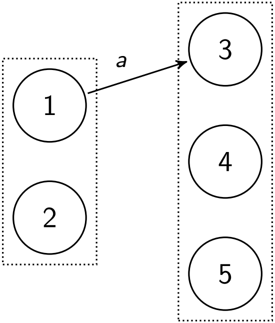
3 Conclusions

ULB UNIVERSITÉ LIBRE DE BRUXELLES

Don't lie to ∃ve

Definition

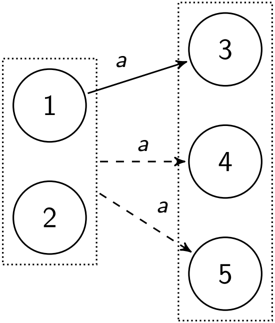
A game of imperfect information is of incomplete information if for every $(q, \sigma, q') \in \Delta$, then for every s' in the same observation as q' there is a transition $(s, \sigma, s') \in \Delta$ where s is in the same observation as q.



Don't lie to ∃ve

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.

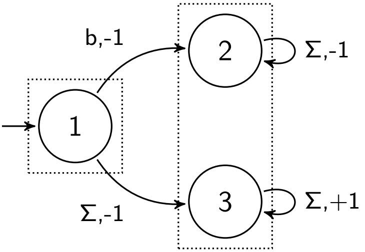


Don't lie to ∃ve

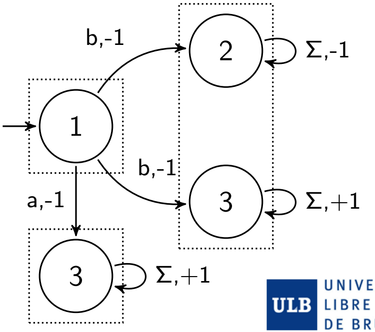
Lemma (imperfect to incomplete info.)

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

G



G'



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

UNIVERSITÉ LIBRE DE BRUXELLES

Observe that in an MPG of incomplete information:

1. the view ∃ve has of a play in the game is $o_0\sigma_0 o_1\sigma_1 \ldots$,

Incomplete information peculiarities

Observe that in an MPG of incomplete information:

1. the view ∃ve 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),

UNIVERSITÉ LIBRE DE BRUXELLES

Observe that in an MPG of incomplete information:

1. the view $\exists$ve 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. $\forall$dam can have a two step strategy: choose observations first,

UNIVERSITÉ LIBRE DE BRUXELLES

Incomplete information peculiarities

Observe that in an MPG of incomplete information:

1. the view $\exists$ve 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. $\forall$dam can have a two step strategy: choose observations first,
4. "delay" the specific choice of states for later!

UNIVERSITÉ LIBRE DE BRUXELLES

$\forall$dam and determinacy

Definition

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

UNIVERSITÉ LIBRE DE BRUXELLES

$\forall$dam and determinacy

Definition

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

Definition (New winning condition)

Let ψ be a play in the game. $\exists$ve wins if all paths in $\gamma(\psi)$ are winning for her. $\forall$dam wins if there is some path which is winning for him.

UNIVERSITÉ LIBRE DE BRUXELLES

$\forall$dam and determinacy

Definition

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

Definition (New winning condition)

Let ψ be a play in the game. $\exists$ve wins if all paths in $\gamma(\psi)$ are winning for her. $\forall$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 γ). The new winning condition is coSuslin and hence determined.*

Outline

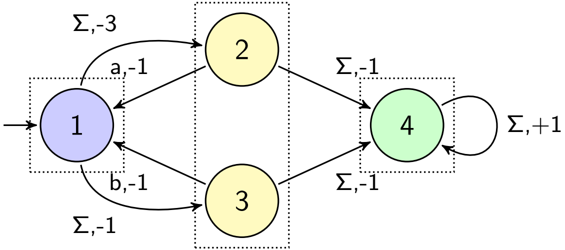
1. MPG variations
   - Mean-payoff games
   - Imperfect information
2. Tackling MPGs with imperfect information
   - Incomplete information
   - Observable determinacy
   - Decidable subclasses
   - Pure games with incomplete information
3. Conclusions

UNIVERSITÉ LIBRE DE BRUXELLES

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.

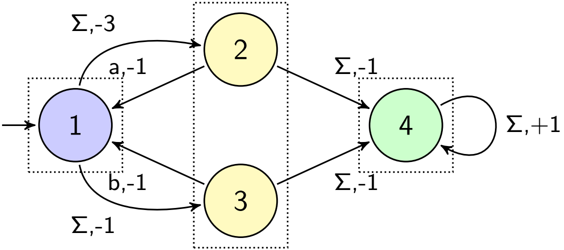


obs: blue
play: f_I
cur. f: $f_I(1) = 0$

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.



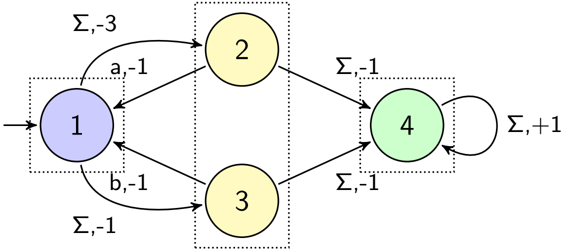
obs: blue-a-yellow
play: $f_I \; a \; f_1$
cur. f: $f_1(2) = -3, f_1(3) = -1$

UNIVERSITÉ LIBRE DE BRUXELLES

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

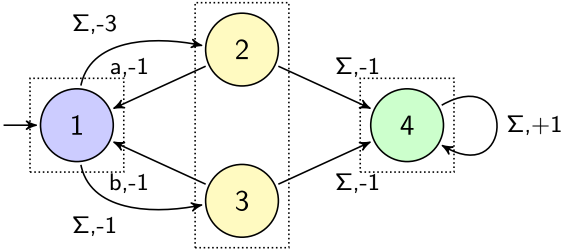
play: f_I a f_1 b f_2

cur. f: $f_2(4) = -4$

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.

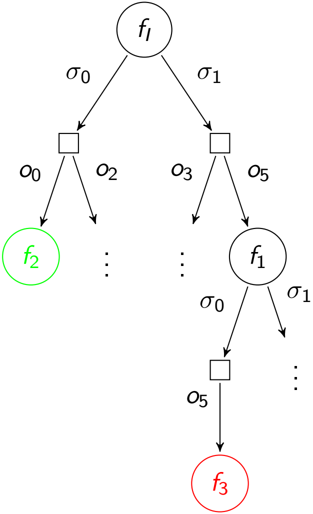


obs: blue-a-yellow-b-green-a-green

play: f_I 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 ∃ve)

∃ve has a finite memory winning strategy in G if and only if she has a winning strategy in H.

Theorem (Strat. transfer for ∀dam)

If ∀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: $\exists$ve can force good leaves or $\forall$dam can force bad leaves.

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.

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.

UNIVERSITÉ LIBRE DE BRUXELLES

Relevant problems

Let $\mathcal{A}$ 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 $\mathcal{A}$?

Problem (Winner determination)

Does ∃ve have a winning strategy in G?

Summary

	Finite memory	Adequately pure		Pure	
Information		incomplete	imperfect	incomplete	imperfect
Class-membership	Undec[1]	PSPACE-complete	NEXP-hard, in EXPSPACE	coNP-complete	coNEXP-complete
Winner-det.	R-c	PSPACE-complete	EXP-complete	NP ∩ coNP	EXP-complete

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

UNIVERSITÉ LIBRE DE BRUXELLES

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

UNIVERSITÉ LIBRE DE BRUXELLES

Does ∃ve win pure G?

Theorem

Deciding if ∃ve has a winning strategy in a given pure MPG with incomplete information is in NP ∩ coNP.

Based on Björklund et al. [2004].

Observe* that positional strategies suffice for ∃ve to win pure games with incomplete information. □

Is G pure?

Theorem

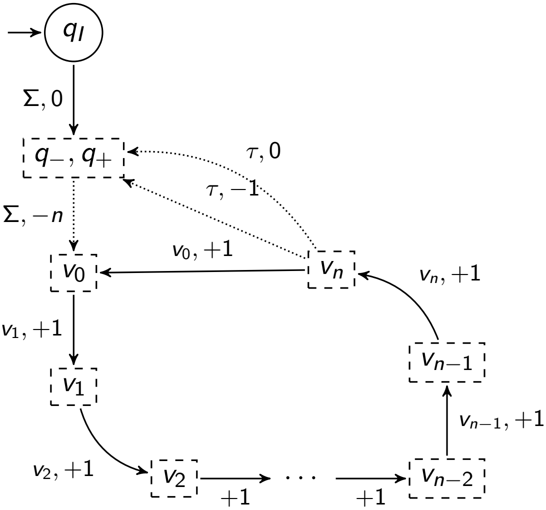
The class membership problem for pure games with incomplete information is coNP*-complete.*

Proof.

- One can "guess" a branch in H (of size at most $|Obs| + 1$) 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



UNIVERSITÉ LIBRE DE BRUXELLES

Summary

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

	Finite memory	Adequately pure		Pure	
Information		incomplete	imperfect	incomplete	imperfect
Class-membership	Undec[1]	PSPACE-complete	NEXP-hard, in EXPSPACE	coNP-complete	coNEXP-complete
Winner-det.	R-c	PSPACE-complete	EXP-complete	NP ∩ coNP	EXP-complete

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

References I

Björklund, H., Sandberg, S., and Vorobyov, S. (2004). Memoryless determinacy of parity and mean payoff games: a simple proof. Theoretical Computer Science, 310(1):365–378.

Degorre, A., Doyen, L., Gentilini, R., Raskin, J.-F., and Toruńczyk, S. (2010). Energy and mean-payoff games with imperfect information. In Computer Science Logic, pages 260–274. Springer.

Ehrenfeucht, A. and Mycielski, J. (1979). Positional strategies for mean payoff games. International Journal of Game Theory, 8:109–113.

Galperin, H. and Wigderson, A. (1983). Succinct representations of graphs. Information and Control, 56(3):183–198.