Energy and Mean-payoff Games

Algorithms and Imperfect Information

JF Raskin - Université Libre de Bruxelles

Joint work with
Laurent Doyen (LSV-ENS Cachan)
Raphaëlla Gentilini (ULB-Univ. Perugia)

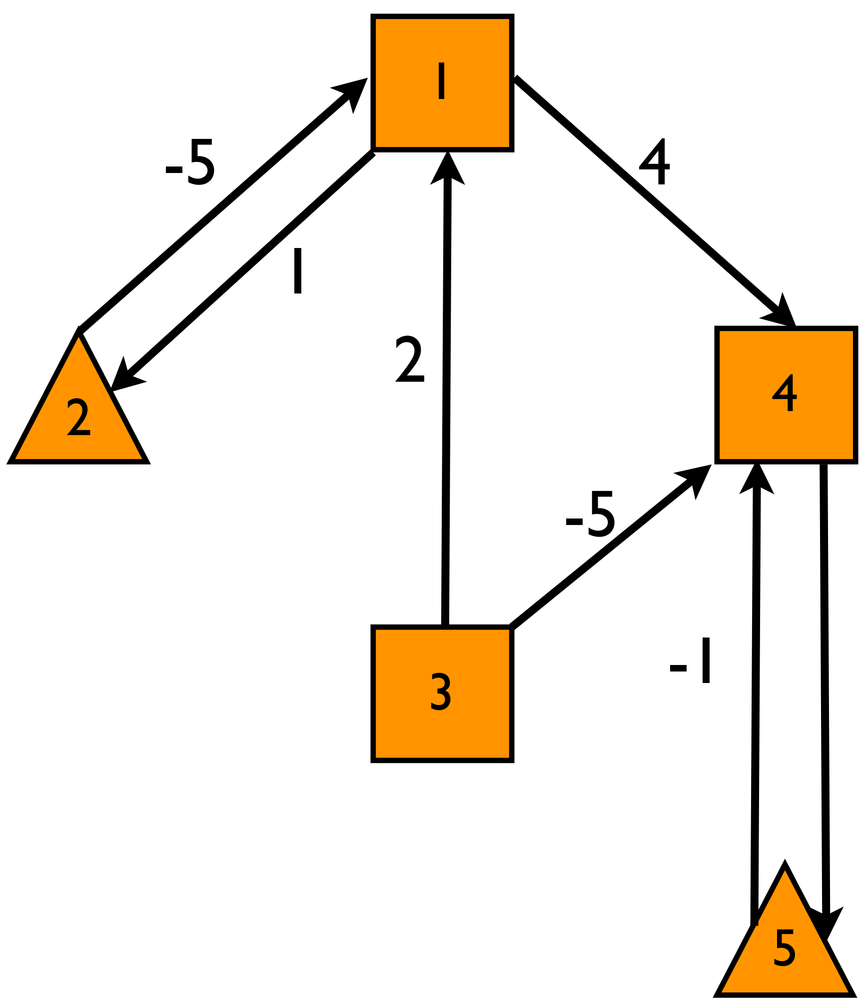
Gasics Workshop
RWTH Aachen
October 22nd and 23rd, 2009

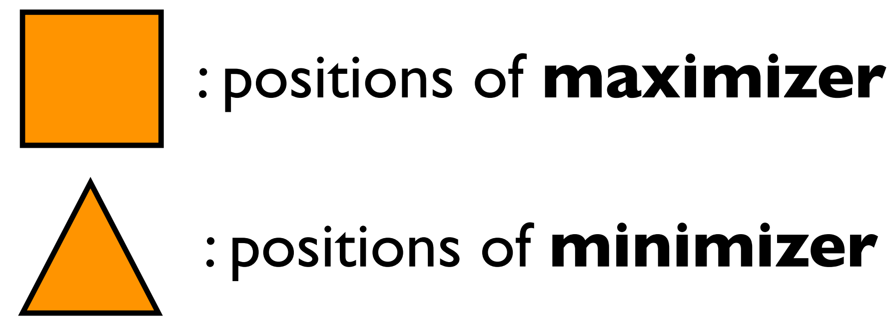
Plan of the talk

- Mean-payoff games and energy games
- An algorithm for solving energy games
- Energy games with imperfect information

Mean-payoff games and energy games

Mean-payoff Games





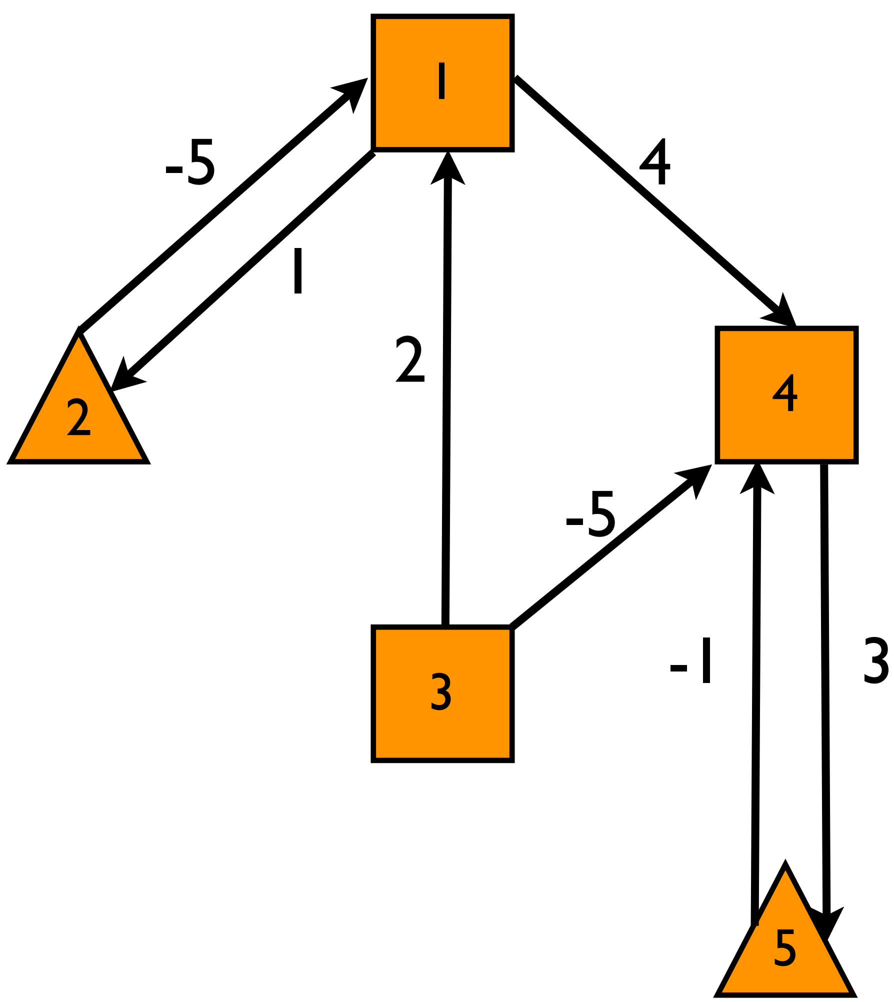
Edges are labelled with rewards

(1,4) (4,5) (5,4) ... (4,5) (5,4) ....
4 3 -1 3 -1 ....

$$= \mathbf{Lim\ Sup}_{n\to+\infty} \Sigma_{i=0,i=n}\, r_i / n = 1.$$

= **MP**((1,4) (4,5) (5,4) ... (4,5) (5,4) ....)

Mean-payoff Games



Strategies for Maximizer

π: Pos*. PosMax→edge.

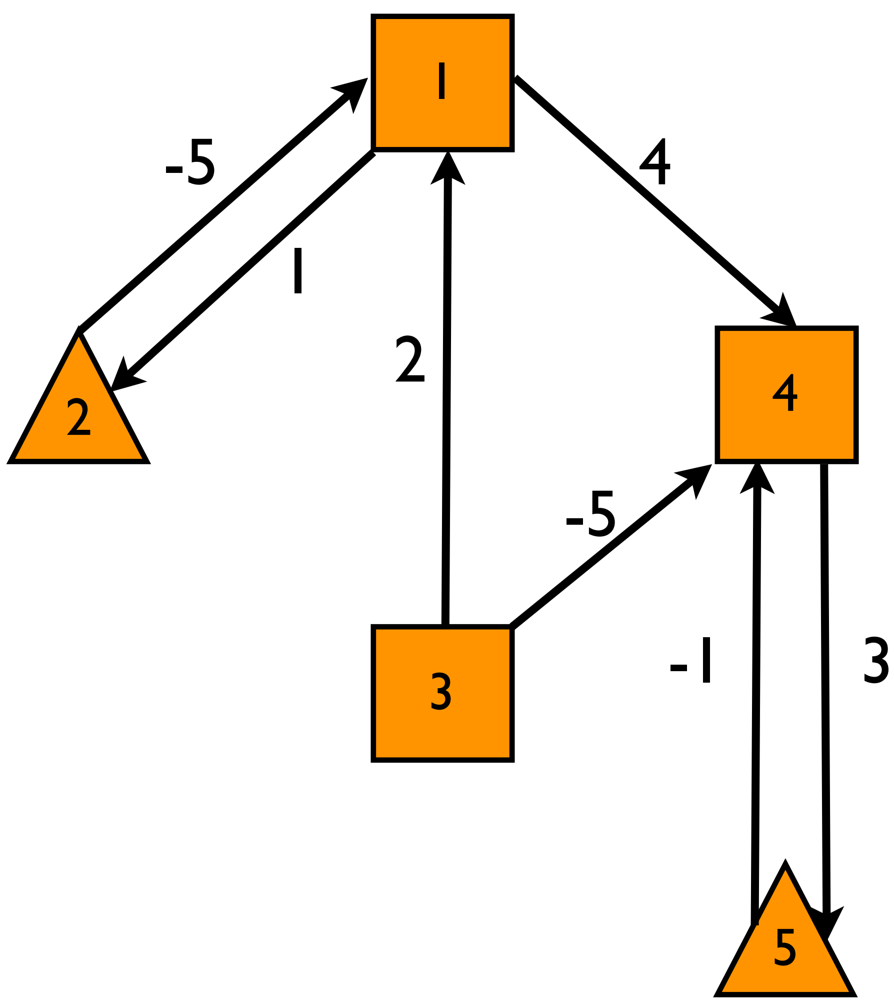
Memoryless: **π:PosMax→edge**.
Π=strategies of Maximizer,
Π_m=memoryless strategies.

Strategies for Minimizer

σ: Pos*. PosMin→edge.

Memoryless: **σ:PosMin→edge**.
Σ=strategies of Minimizer,
Σ_m=memoryless strategies.

Mean-payoff Games



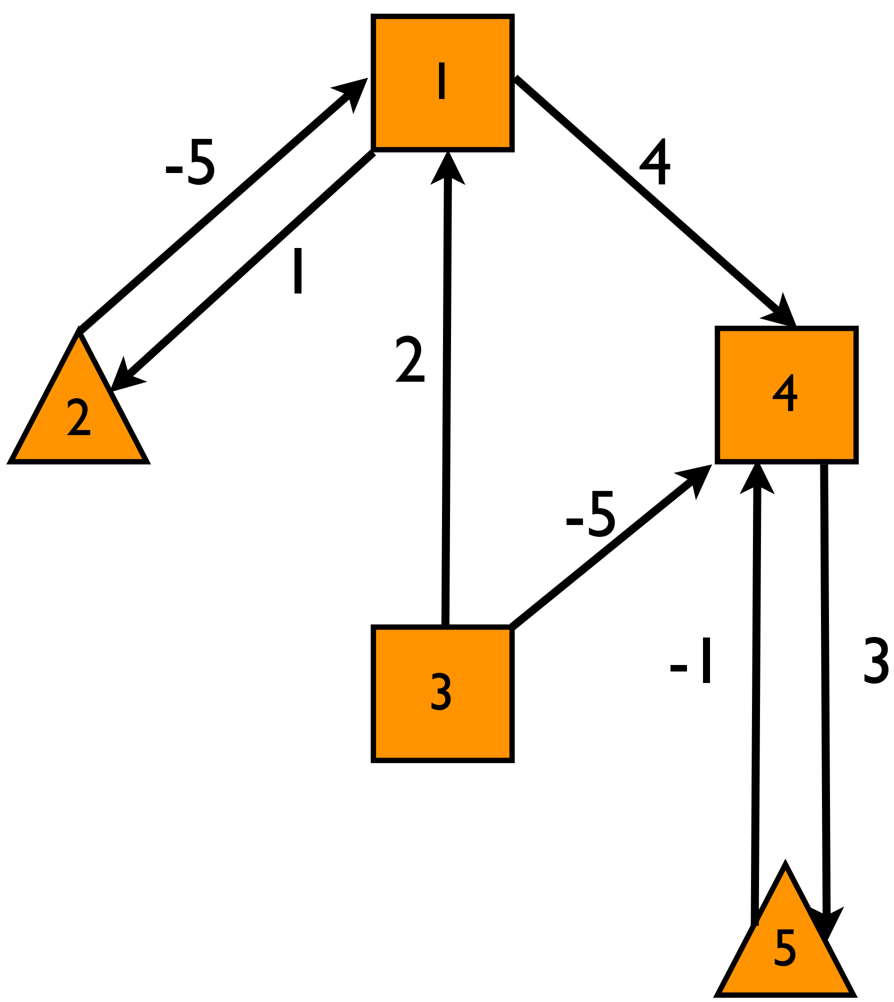
Outcome(**s**,π,σ)=infinite sequence of edges when players play π and σ from **s**.

The value of the game in **s** under π and σ is

MP(**Outcome**(**s**,π,σ)).

The decision problem for mean-payoff games asks given an initial state **s** if Maximizer has **a strategy** to ensure a mean-payoff greater or equal to a integer value v from **s** no matter what Minimizer plays.

Mean-payoff Games



Theorem (Memoryless determinacy)[EM79]
All following expressions are equal:

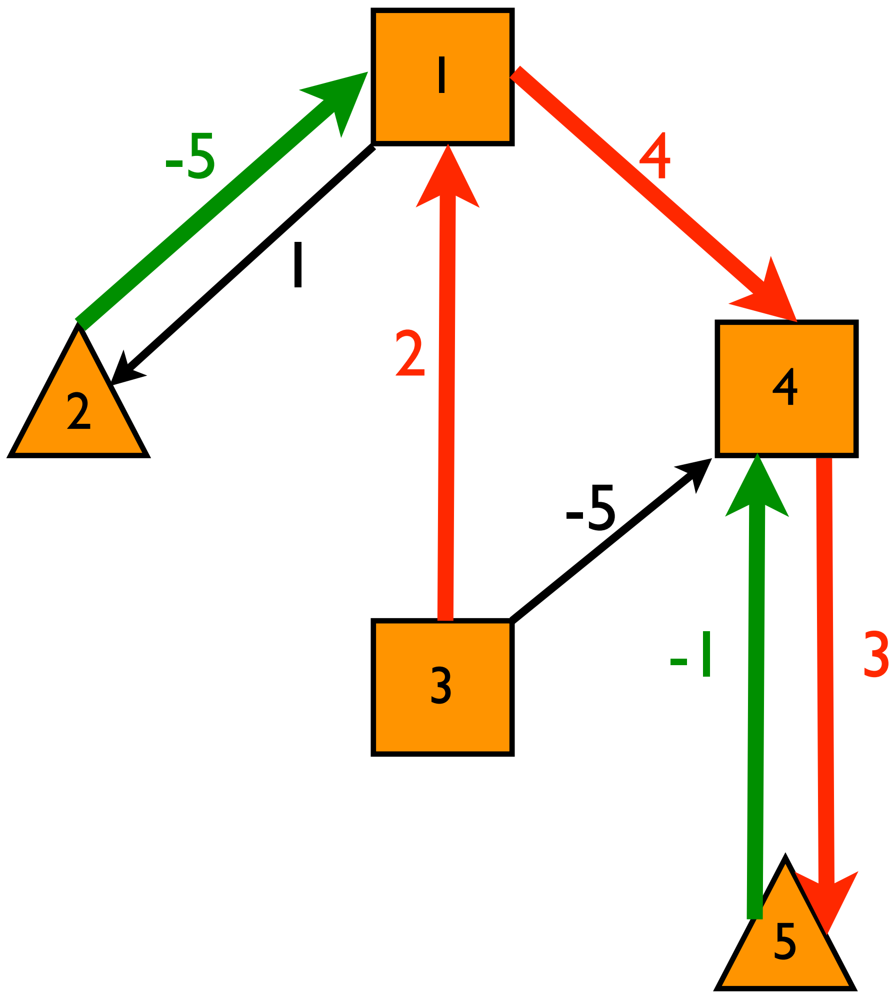
$$\mathbf{Sup}_{\pi\in\Pi}\ \mathbf{Inf}_{\sigma\in\Sigma}\ \mathbf{MP}(\mathbf{Outcome}(\mathbf{s},\pi,\sigma))$$

$$= \mathbf{Inf}_{\sigma\in\Sigma}\ \mathbf{Sup}_{\pi\in\Pi}\ \mathbf{MP}(\mathbf{Outcome}(\mathbf{s},\pi,\sigma))$$

$$= \mathbf{Sup}_{\pi\in\Pi_m}\ \mathbf{Inf}_{\sigma\in\Sigma_m}\ \mathbf{MP}(\mathbf{Outcome}(\mathbf{s},\pi,\sigma))$$

$$= \mathbf{Inf}_{\sigma\in\Sigma_m}\ \mathbf{Sup}_{\pi\in\Pi_m}\ \mathbf{MP}(\mathbf{Outcome}(\mathbf{s},\pi,\sigma))$$

Mean-payoff Games



Theorem (Memoryless determinacy)[EM79]
All following expressions are equal:

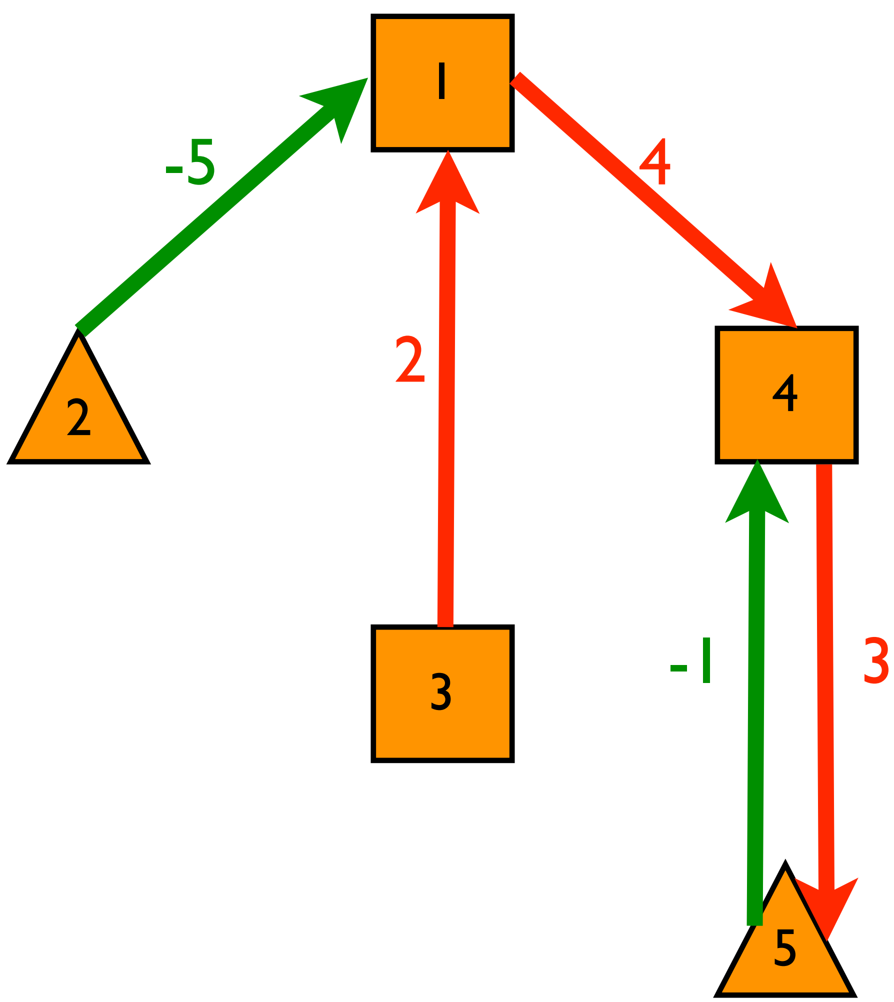
$$\mathbf{Sup}_{\pi\in\Pi}\ \mathbf{Inf}_{\sigma\in\Sigma}\ \mathbf{MP}(\mathbf{Outcome}(\mathbf{s},\pi,\sigma))$$

$$= \mathbf{Inf}_{\sigma\in\Sigma}\ \mathbf{Sup}_{\pi\in\Pi}\ \mathbf{MP}(\mathbf{Outcome}(\mathbf{s},\pi,\sigma))$$

$$= \mathbf{Sup}_{\pi\in\Pi_m}\ \mathbf{Inf}_{\sigma\in\Sigma_m}\ \mathbf{MP}(\mathbf{Outcome}(\mathbf{s},\pi,\sigma))$$

$$= \mathbf{Inf}_{\sigma\in\Sigma_m}\ \mathbf{Sup}_{\pi\in\Pi_m}\ \mathbf{MP}(\mathbf{Outcome}(\mathbf{s},\pi,\sigma))$$

Mean-payoff Games



Theorem (Memoryless determinacy)[EM79]
All following expressions are equal:

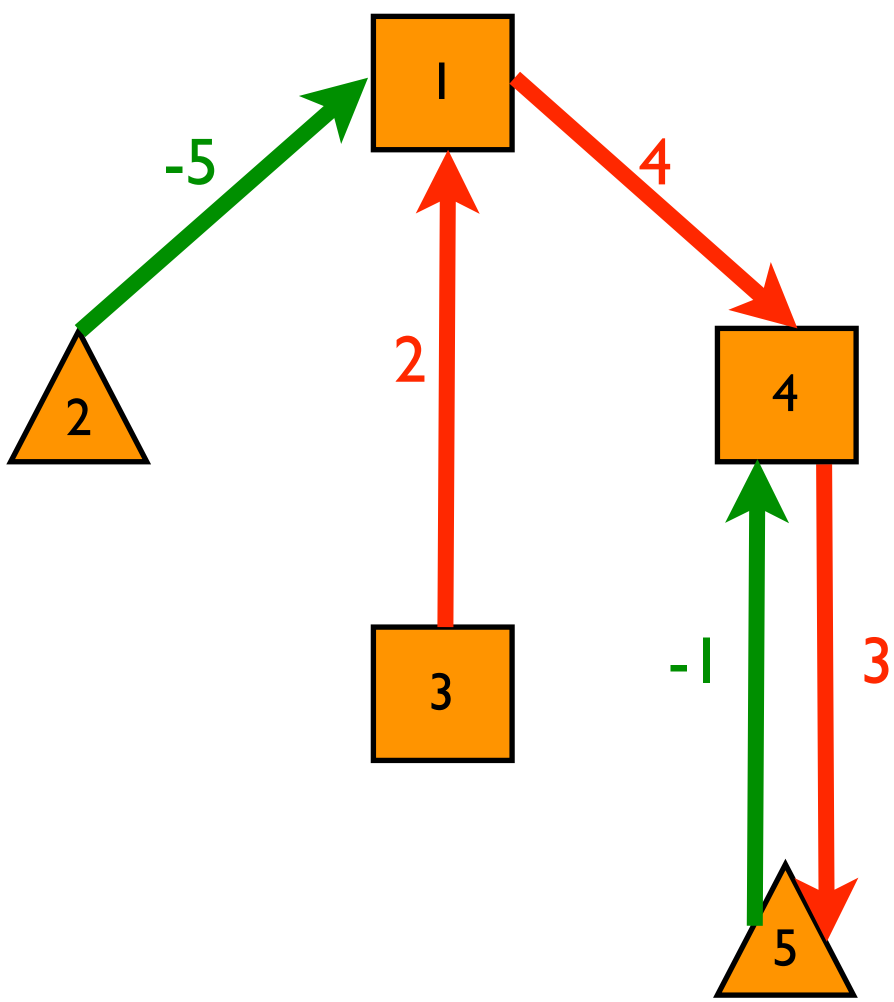
$\mathbf{Sup}_{\pi\in\Pi}\ \mathbf{Inf}_{\sigma\in\Sigma}\ \mathbf{MP}(\mathbf{Outcome}(\mathbf{s},\pi,\sigma))$

$= \mathbf{Inf}_{\sigma\in\Sigma}\ \mathbf{Sup}_{\pi\in\Pi}\ \mathbf{MP}(\mathbf{Outcome}(\mathbf{s},\pi,\sigma))$

$= \mathbf{Sup}_{\pi\in\Pi_m}\ \mathbf{Inf}_{\sigma\in\Sigma_m}\ \mathbf{MP}(\mathbf{Outcome}(\mathbf{s},\pi,\sigma))$

$= \mathbf{Inf}_{\sigma\in\Sigma_m}\ \mathbf{Sup}_{\pi\in\Pi_m}\ \mathbf{MP}(\mathbf{Outcome}(\mathbf{s},\pi,\sigma))$

Mean-payoff Games



Theorem (Memoryless determinacy)[EM79]
All following expressions are equal:

$$\mathbf{Sup}_{\pi\in\Pi}\ \mathbf{Inf}_{\sigma\in\Sigma}\ \mathbf{MP}(\mathbf{Outcome}(s,\pi,\sigma))$$

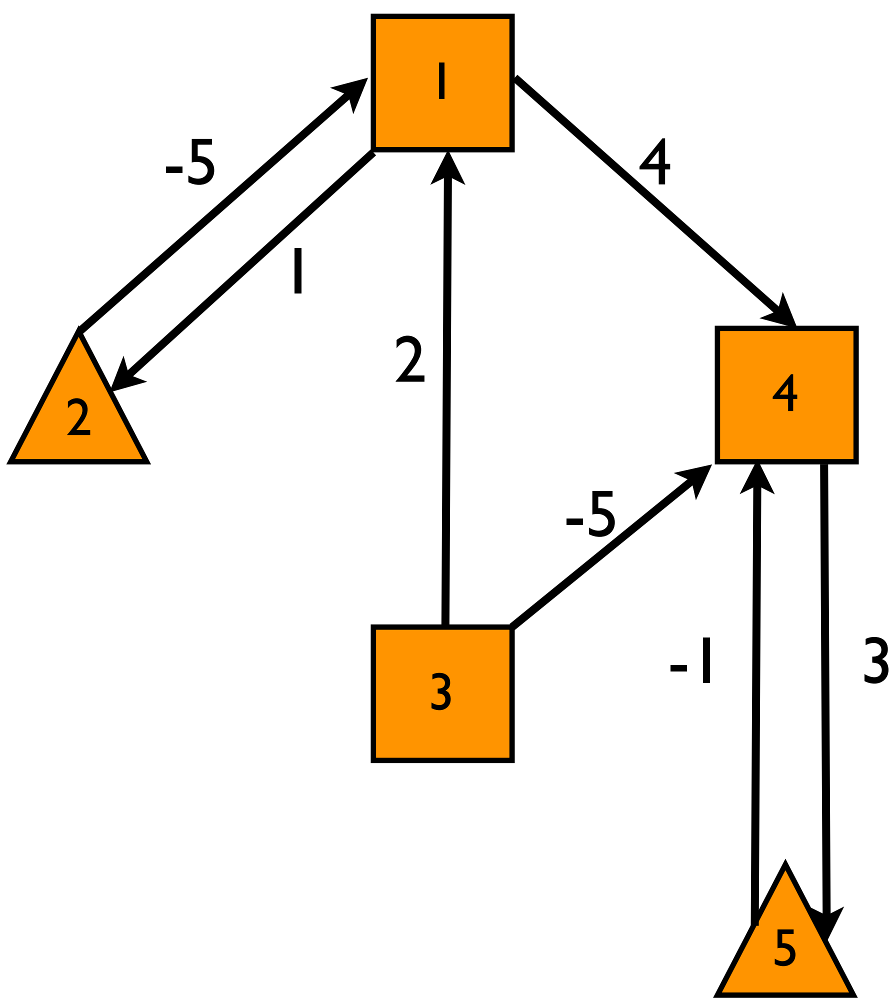
$$= \mathbf{Inf}_{\sigma\in\Sigma}\ \mathbf{Sup}_{\pi\in\Pi}\ \mathbf{MP}(\mathbf{Outcome}(s,\pi,\sigma))$$

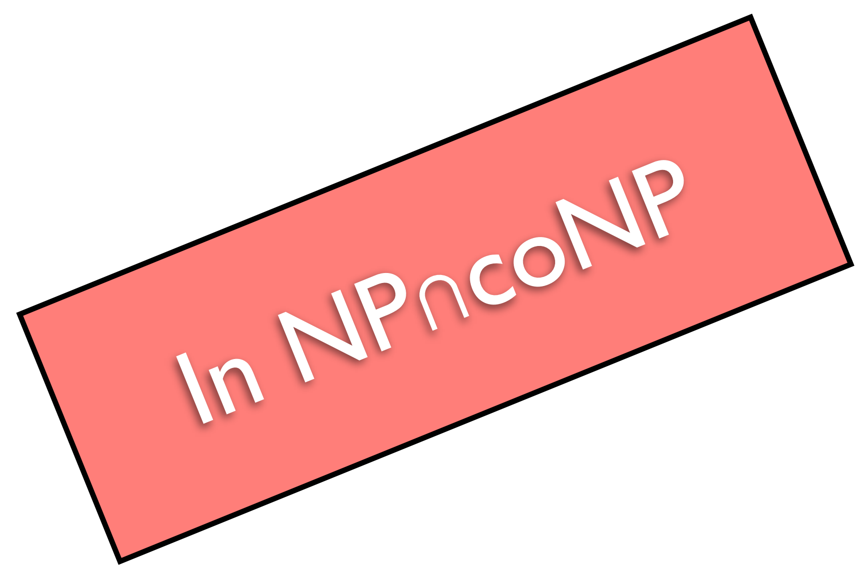
$$= \mathbf{Sup}_{\pi\in\Pi_m}\ \mathbf{Inf}_{\sigma\in\Sigma_m}\ \mathbf{MP}(\mathbf{Outcome}(s,\pi,\sigma))$$

$$= \mathbf{Inf}_{\sigma\in\Sigma_m}\ \mathbf{Sup}_{\pi\in\Pi_m}\ \mathbf{MP}(\mathbf{Outcome}(s,\pi,\sigma))$$

All states have value 1 !

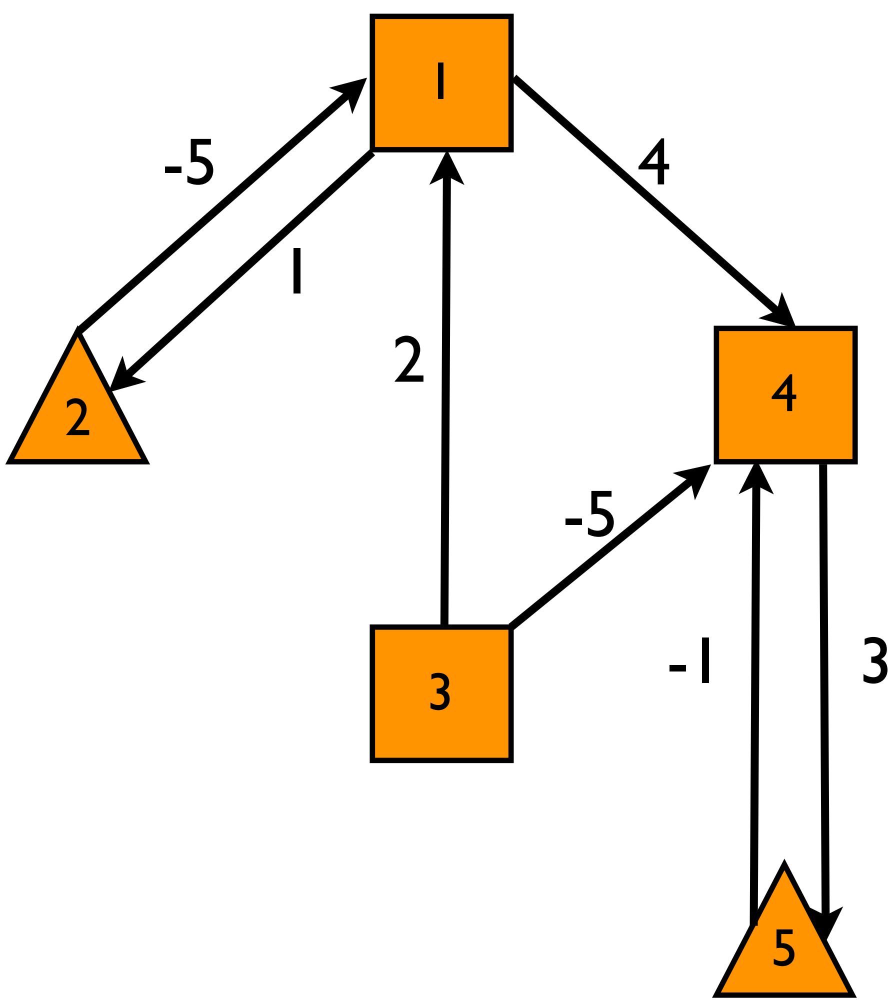
Mean-payoff Games

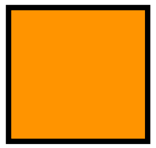


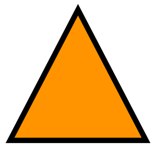


The decision problem for mean-payoff games asks given an initial state **s** if Maximizer has **a strategy** to ensure a mean-payoff greater or equal to a integer value v from **s** no matter what Minimizer plays.

Energy Games [CdAHS03,BFLM08]




: positions of **maximizer**


: positions of **minimizer**

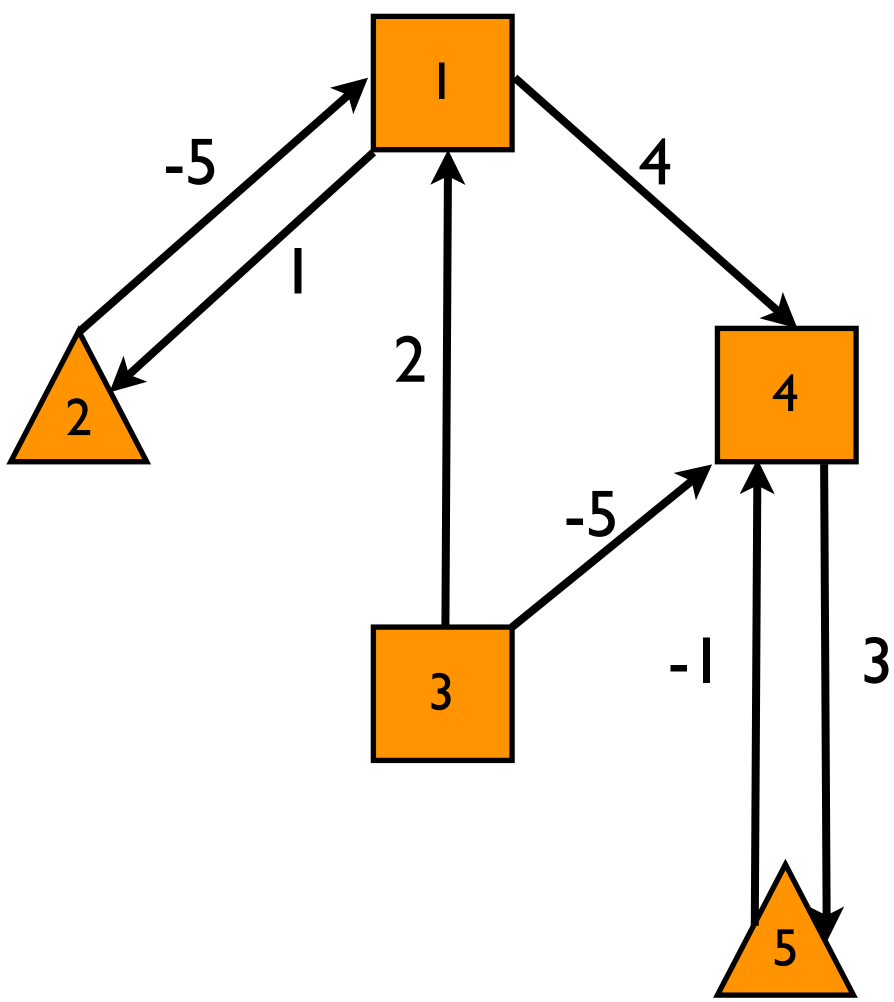
Edges are labelled with energy consumptions or energy gains

Initial energy level : 7

Play : (1,2) (2,1) (1,4) (4,5) (5,4) (4,5) (5,4) ...

EL : 7 8 3 7 10 9 10 9 ...

Energy Games

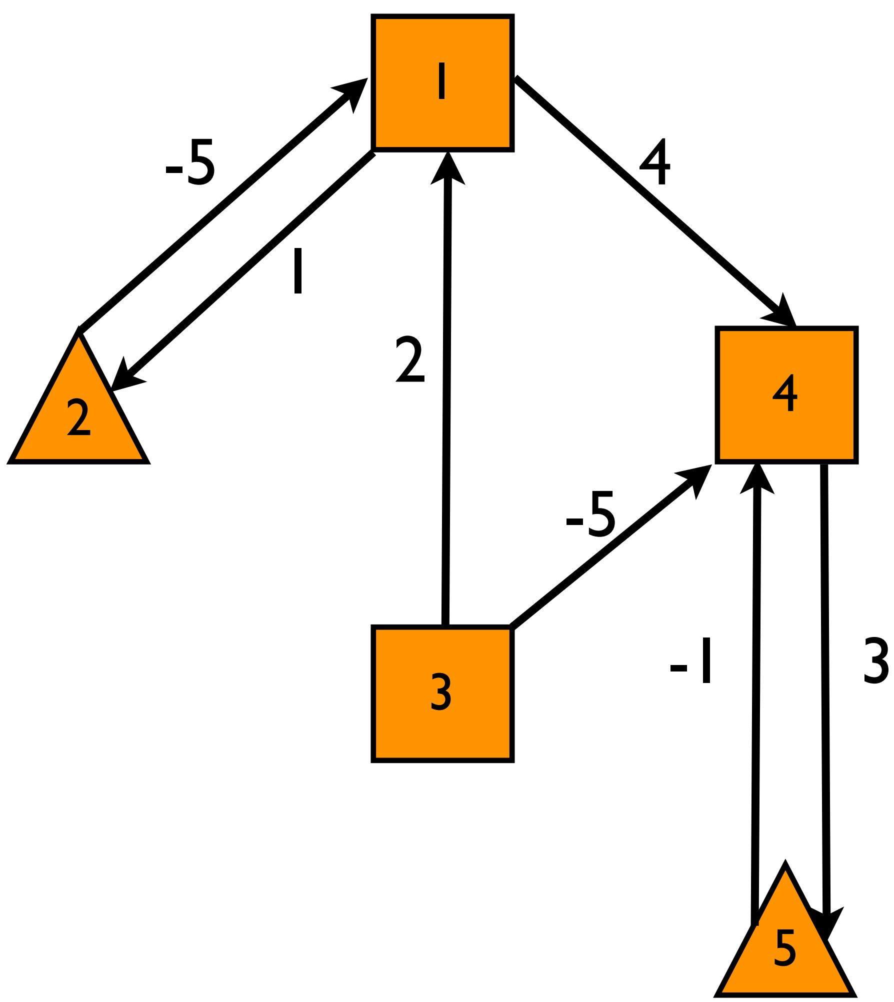


The **decision problem for energy games** asks given an initial state **s** if there exist

1) an **initial energy level**

2) and π **for Maximizer** to maintain a positive energy level at all time no matter what Minimizer plays.

Energy Games



Thm [BFLM08] **(Memoryless determinacy)**
All following expressions are equivalent:

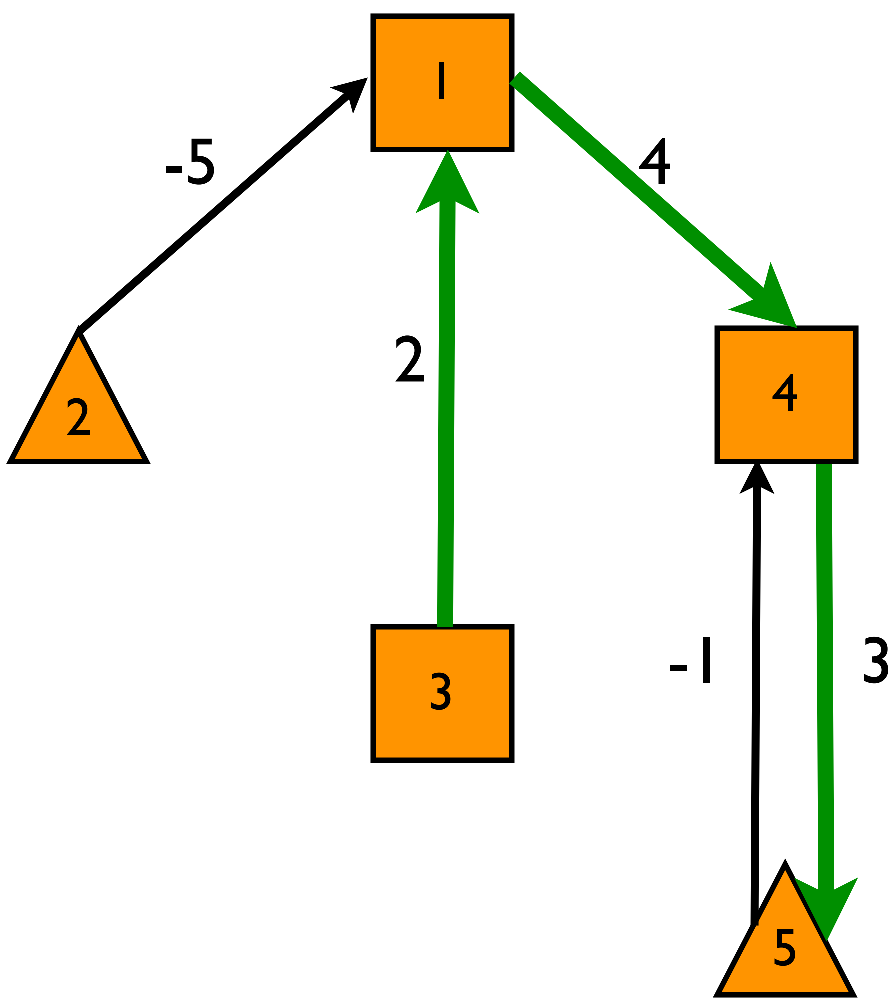
$\exists \mathbf{I}\ \exists \pi \in \Pi\ \forall \sigma \in \Sigma\ \forall\ i \geq 0\ \mathbf{EL}(\mathbf{I}, \mathrm{Outcome}(s,\pi,\sigma)(0..i)) \geq 0$

iff $\exists \mathbf{I}\ \forall \sigma \in \Sigma\ \exists \pi \in \Pi\ \forall\ i \geq 0\ \mathbf{EL}(\mathbf{I}, \mathrm{Outcome}(s,\pi,\sigma)(0..i)) \geq 0$

iff $\exists \mathbf{I}\ \exists \pi \in \Pi_m\ \forall \sigma \in \Sigma_m\ \forall\ i \geq 0\ \mathbf{EL}(\mathbf{I}, \mathrm{Outcome}(s,\pi,\sigma)(0..i)) \geq 0$

iff $\exists \mathbf{I}\ \forall \sigma \in \Sigma_m\ \exists \pi \in \Pi_m\ \forall\ i \geq 0\ \mathbf{EL}(\mathbf{I}, \mathrm{Outcome}(s,\pi,\sigma)(0..i)) \geq 0$

Mean-payoff Games and Energy Games



Thm. The decision problem for **MPG** and the decision problem for **EG** are **log-space equivalent**.

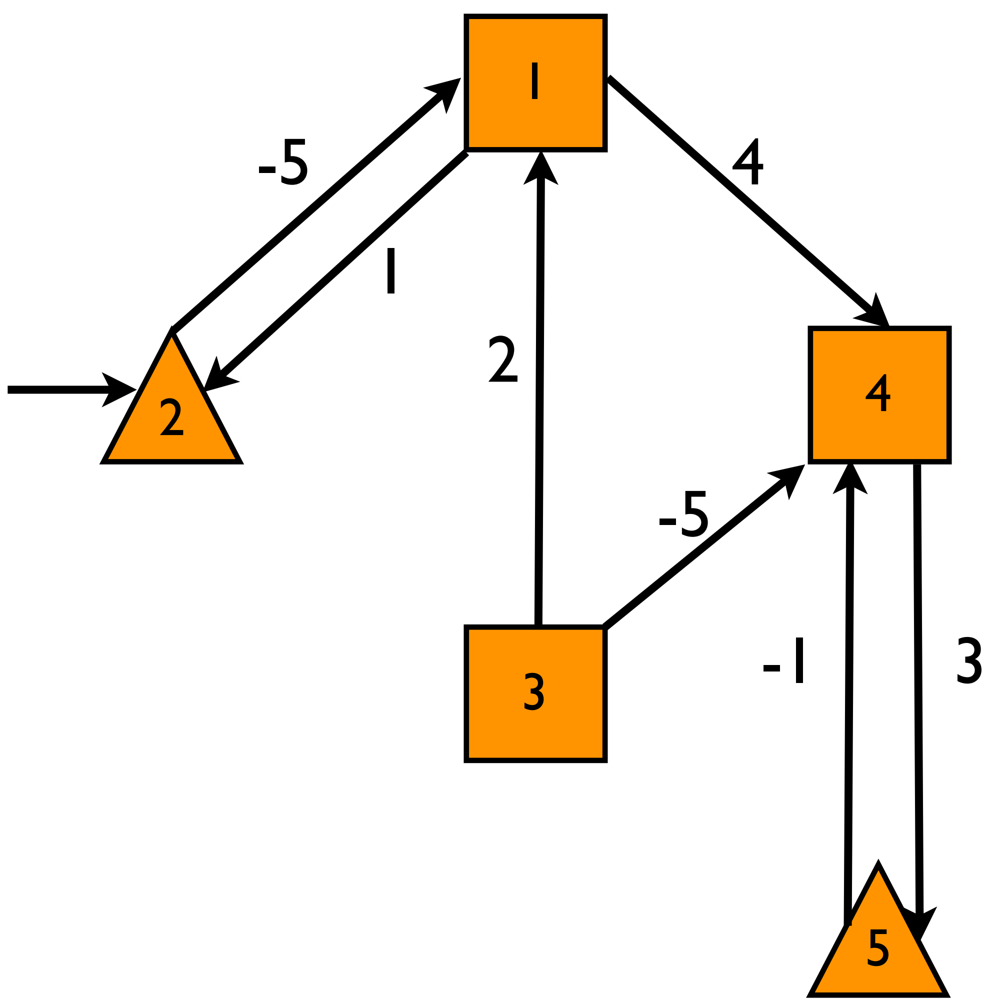
Why ?

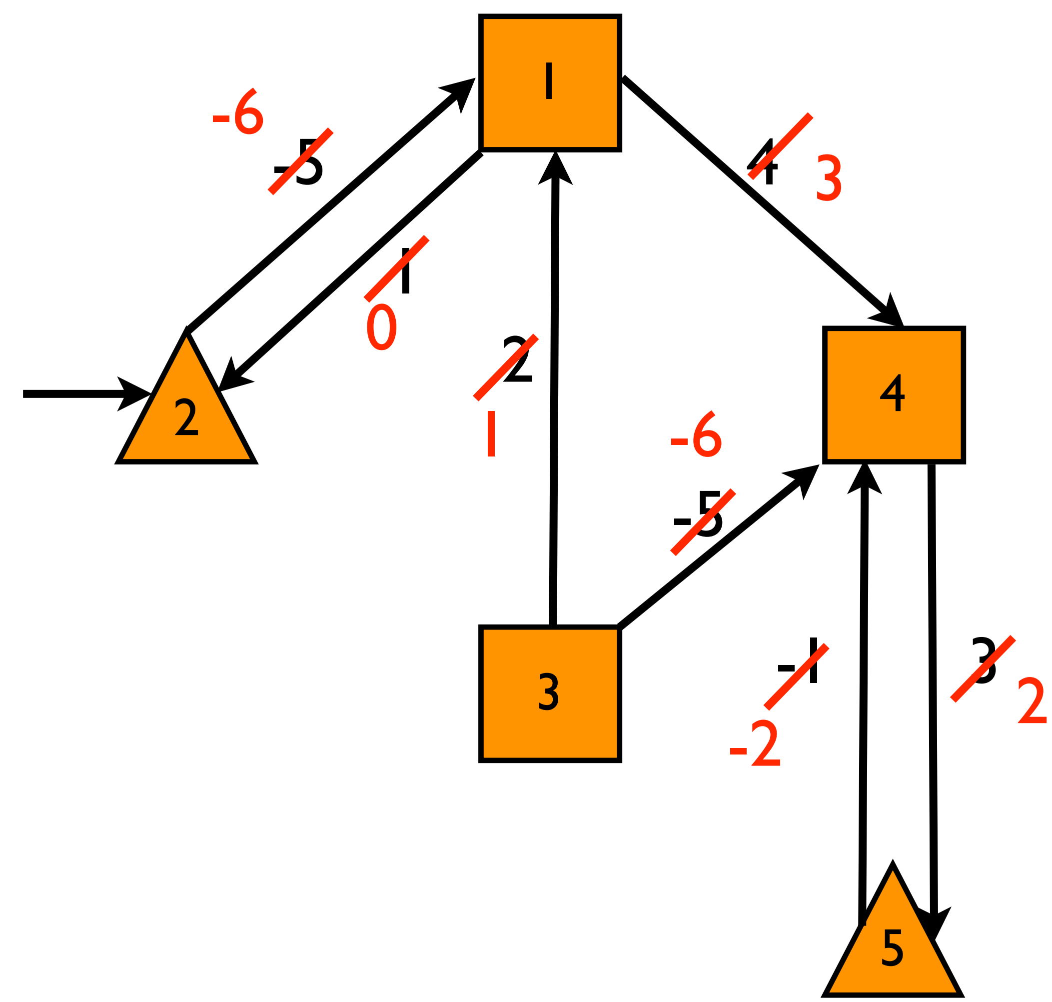
Lemma. For all MPG G, for all value v, for all memoryless π of Maximizer, π secures at least v iff all the reachable cycles in G(π) have mean value at least v.

Lemma. For all ENG G, for all memoryless π of Maximizer, π allows to maintain a positive energy level from **s** iff all the reachable cycles in G(π) from **s** have positive value.

Thm. Maximizer has π in G for value at least v in MPG **iff** Maximizer has a winning strategy in the ENG G-v.

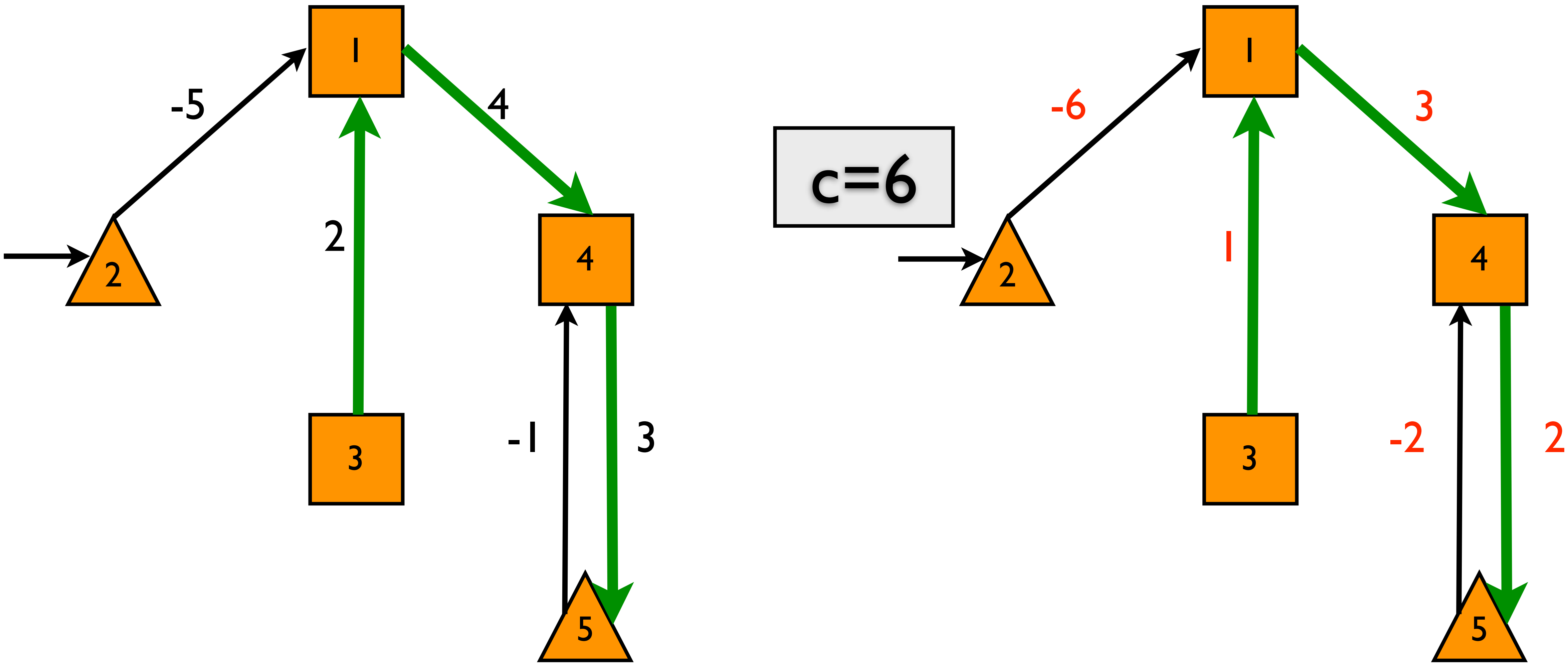
Mean-payoff Games and Energy Games





Maximizer ensure value 1 in MPG **iff** Maximizer win EG G-1.

Mean-payoff Games and Energy Games



Maximizer ensure value 1 in MPG **iff** Maximizer win EG G-1.

An algorithm for solving energy games

Complexities

W is the maximal absolute weight in G	Mean-payoff games	Energy games
Decision problem	$O(E \cdot V^2 \cdot W)$ [ZP96]	
Strategy synthesis	$O(E \cdot V^3 \cdot W \cdot \log(E/V))$ [ZP96]	

More results in the paper...

Algorithmic complexities

W is the maximal absolute weight in G	Mean-payoff games	Energy games
Decision problem	O(E•V•W) (this paper) O(E•V^2•W) [ZP96]	O(E•V•W) (this paper)
Strategy synthesis	O(E•V•W) (this paper) O(E•V^3•W•log(E/V)) [ZP96]	O(E•V•W) (this paper)

More results in the paper...

Algorithmic complexities

W is the maximal absolute weight in G	Mean-payoff games	Energy ga[illegible]s
Decision problem	$O(E \cdot V \cdot W)$ (this paper) $O(E \cdot V^2 \cdot W$[illegible]	[illegible]G) [illegible]is paper)
Strategy synthesis[illegible]	[illegible]aper) [illegible]V•log(E/V)) [ZP96]	$O(E \cdot V \cdot W)$ (this paper)

All are pseudo-polynomial algorithms

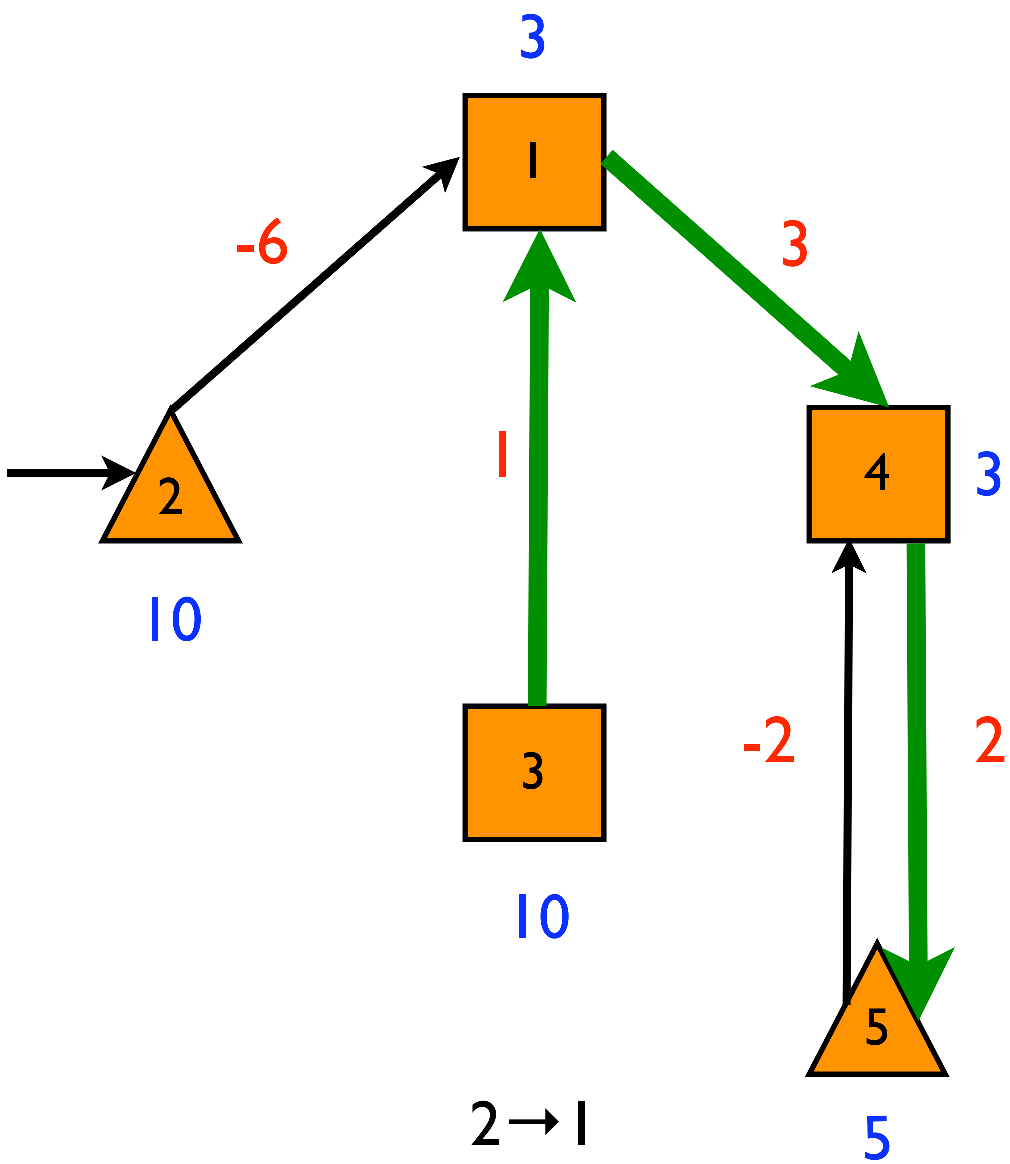
More results in the paper...

Solving Energy Games

Lemma. For all memoryless π of Maximizer, π allows to maintain a positive energy level

iff

all the cycles in G(π) have positive value.



2→1
10≥3-(-6)=9, OK

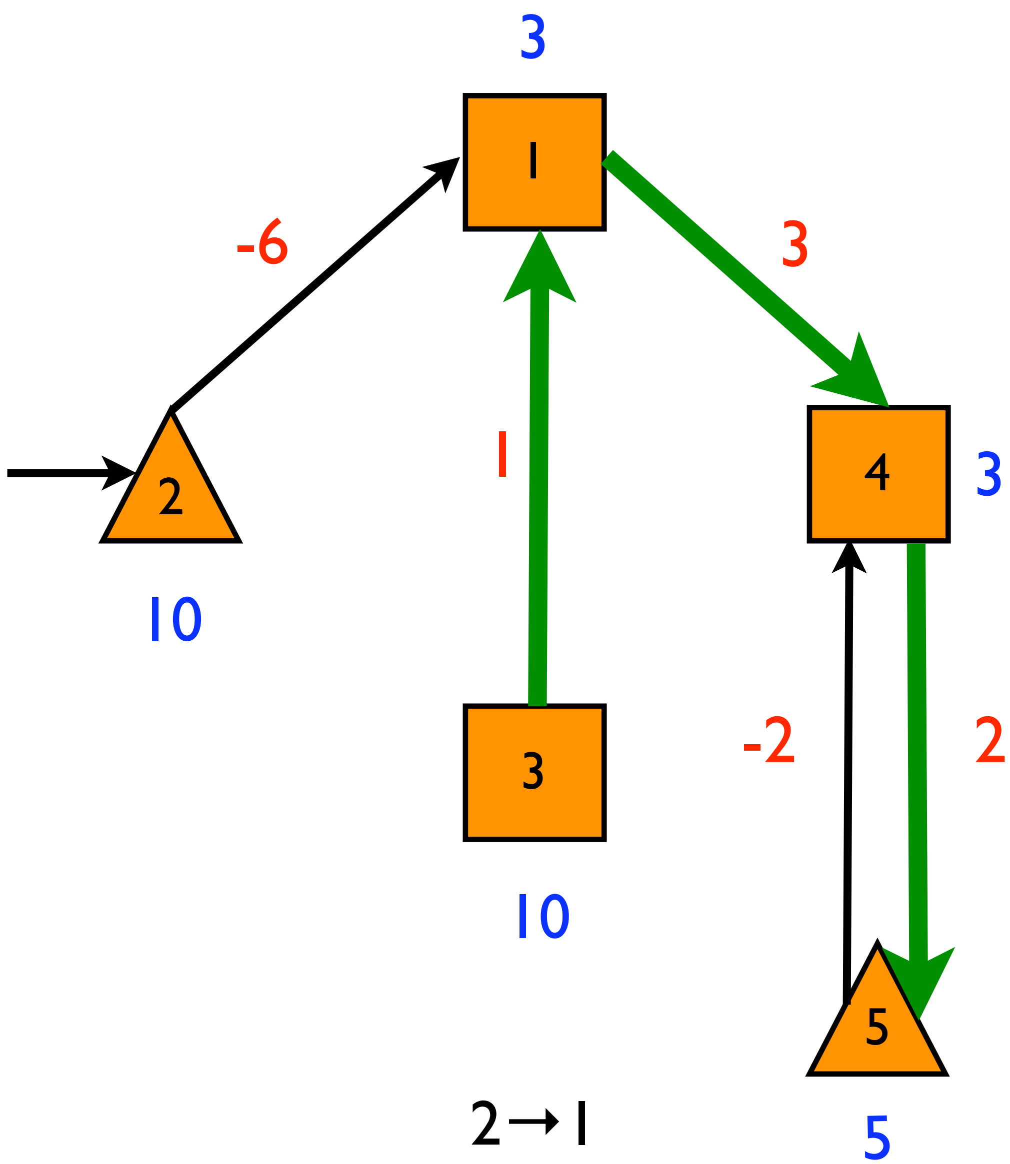
5→4
5≥3-(-2)=5, OK

Solving Energy Games

Lemma. For all memoryless π of Maximizer, π allows to maintain a positive energy level
iff
all the cycles in G(π) have positive value.

When does G(π) have only positive cycles ?

$f: V \to \mathbb{N}$ **is energy progress measure**
iff **for all $(v,v') \in E$: $f(v) \geq f(v') - w(v,v')$**



2→1
$10 \geq 3-(-6)=9$, OK

5→4
$5 \geq 3-(-2)=5$, OK

Solving Energy Games

Lemma. For all memoryless π of Maximizer, π allows to maintain a positive energy level

iff

all the cycles in G(π) have positive value.

When does G(π) have only positive cycles ?

$f: V \to \mathbb{N}$ **is energy progress measure**
iff **for all $(v,v') \in E$: $f(v) \geq f(v') - w(v,v')$**

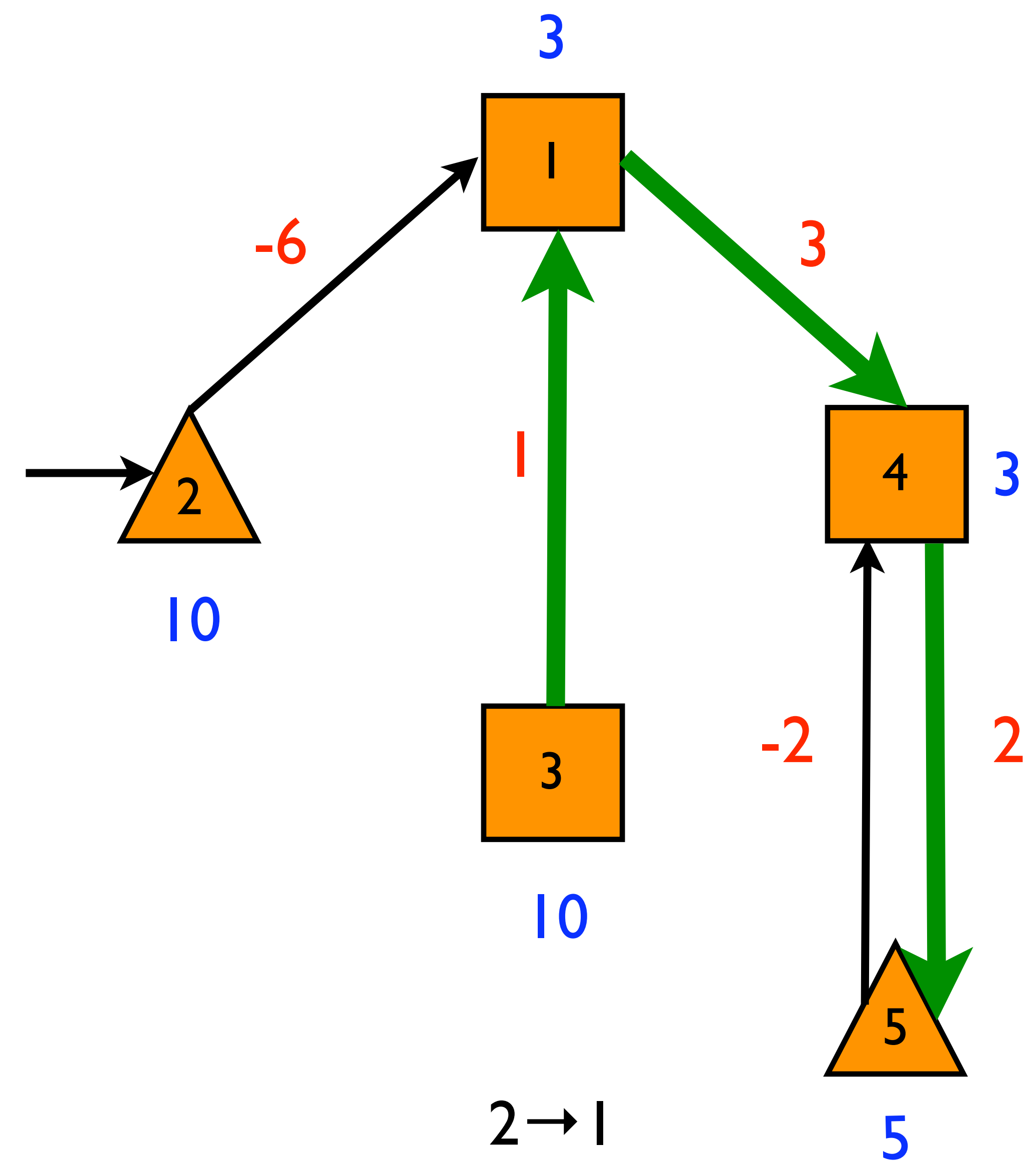
Theorem. G(π) has only positive cycles

iff

there exists an energy progress measure for G(π).

The codomain of f can be bounded by

$M_G = \Sigma_{v \in V}$ **max** $(\{0\} \cup \{ -w(v,v') \mid (v,v') \in E \})$



2→1
10≥3-(-6)=9, OK

5→4
5≥3-(-2)=5, OK

How to synthesize energy progress measures ?

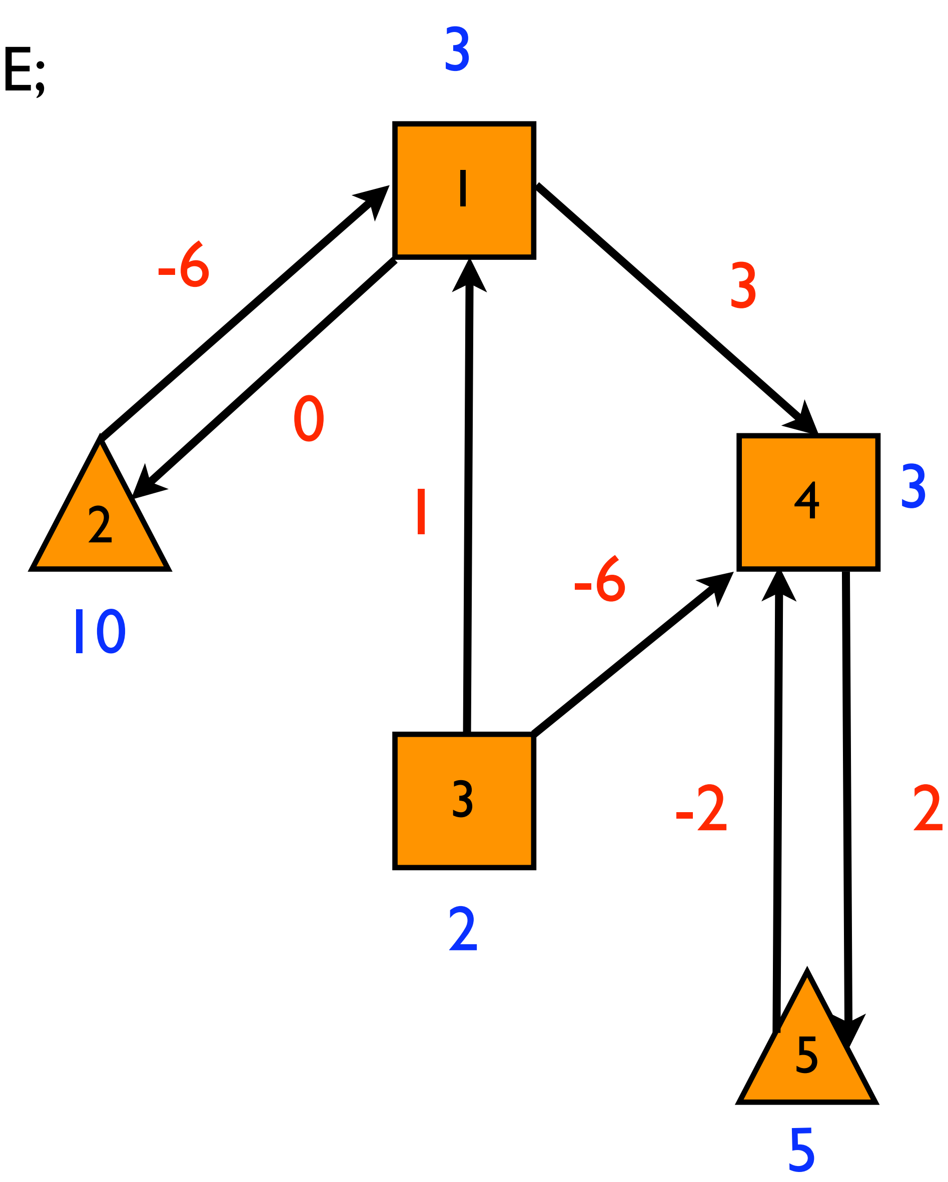
From graphs to games.

Let G be a weighted game graph. Let $\mathscr{F}$ the set of functions f :V → $M_G \cup \{\top\}$.

f∈$\mathscr{F}$ is a **game energy progress measure** iff

-if v belongs to Maximizer: f(v) ≥ f(v') - w(v,v') for **some** (v,v')∈E;

-if v belongs to Minimizer: f(v) ≥ f(v') - w(v,v') for **all** (v,v')∈E.



Let us consider state 3 which belongs to Maximizer.

3→4

2≥3-(-6)=9, KO

3→1

2≥3-1=2, OK

⟹ OK for state 3.

How to synthesize energy progress measures ?

$f_1 \sqsubseteq f_2$ when $f_1(v) \leq f_2(v)$ for all $v \in V$.

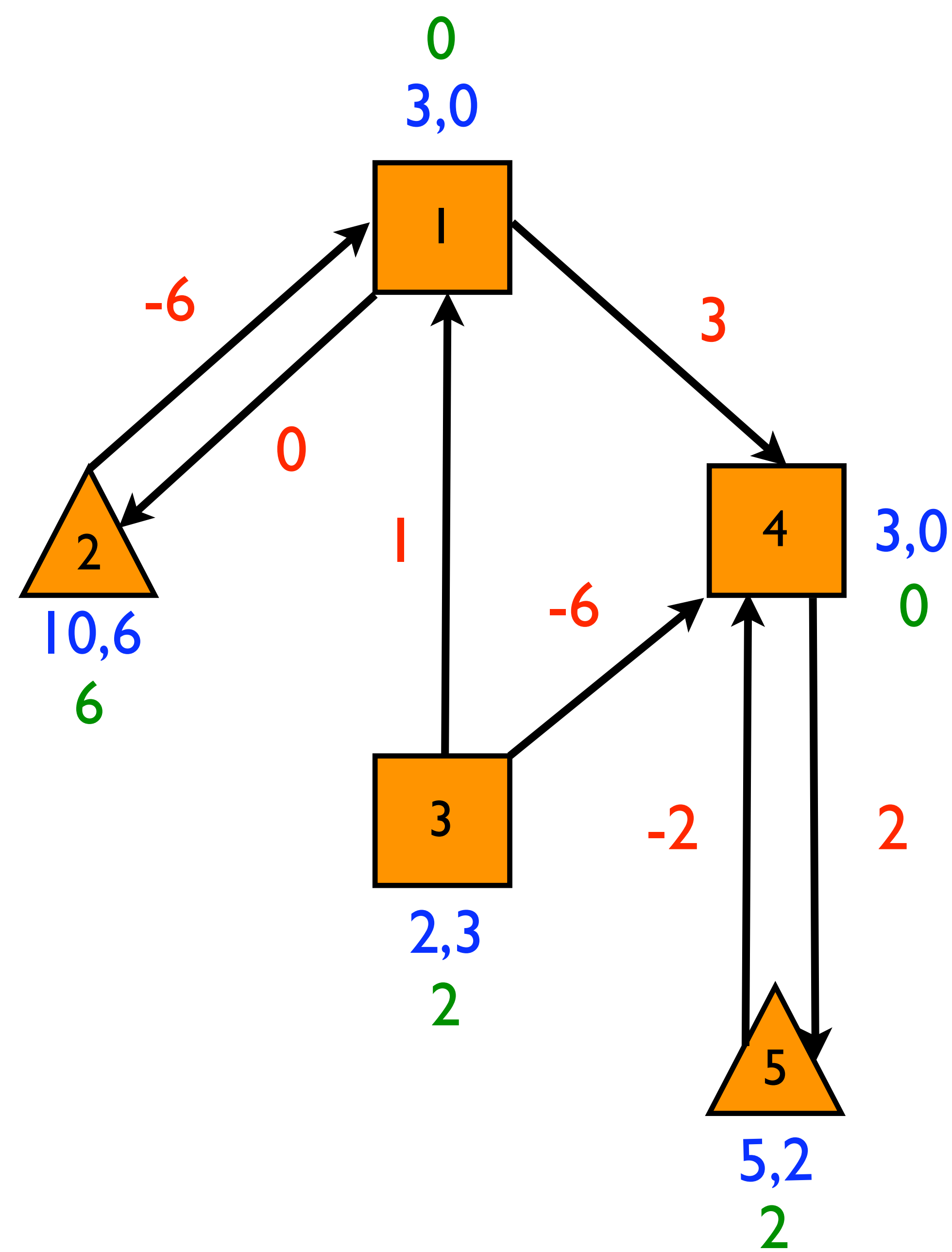
$(\mathscr{F}, \sqsubseteq)$ is a cpo.

Let $f_1, f_2 : V \rightarrow M_G \cup \{\top\}$,

$\mathbf{min}\{f_1,f_2\} : v \in V \rightarrow \mathbf{min}(f_1(v), f_2(v))$.

$\mathbf{min}\{f_1,f_2\} = \sqcap\{f_1,f_2\}$ (glb for $\sqsubseteq$).

Lemma. If f_1 and f_2 are game progress measures, then $\mathbf{min}\{f_1,f_2\}$ is also a game progress measure.

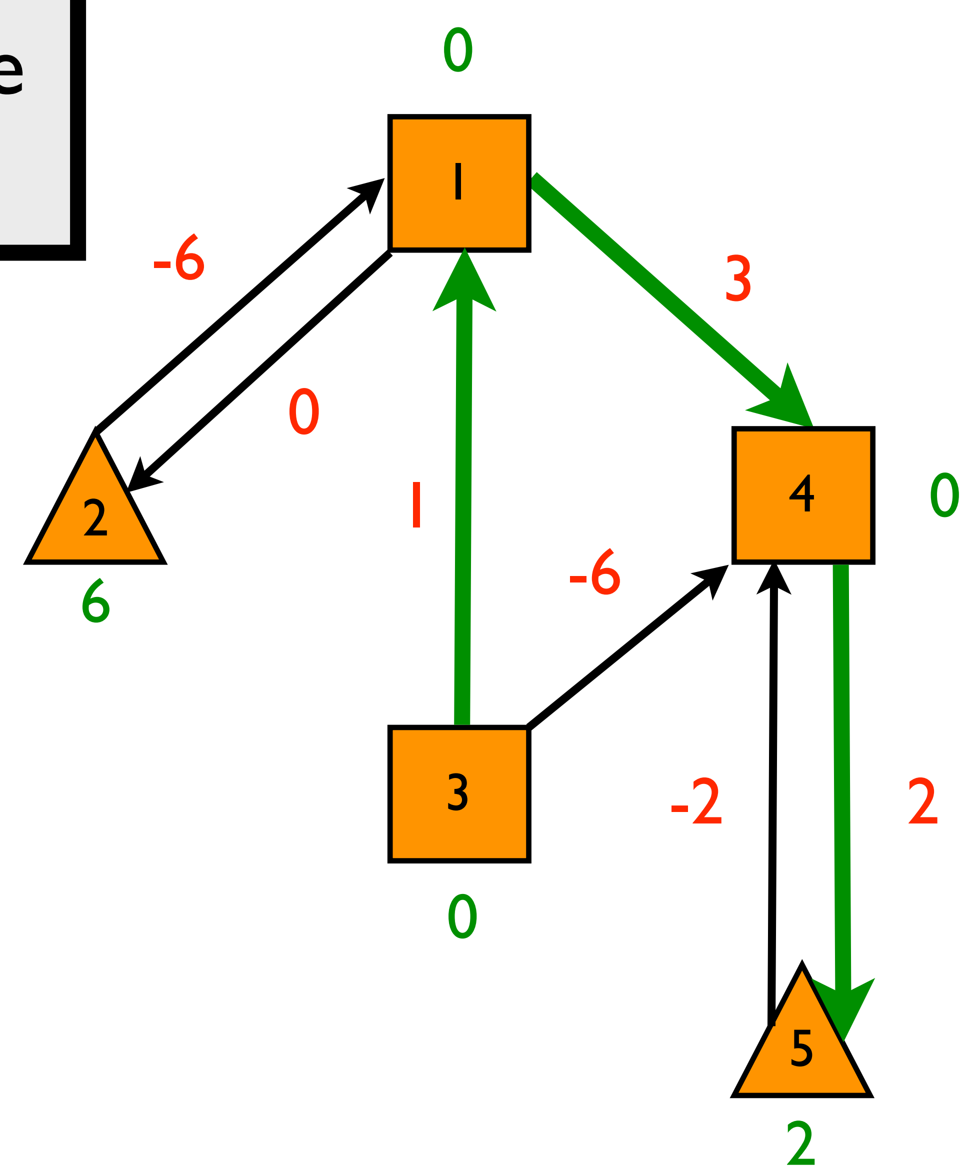


How to synthesize energy progress measures ?

Thm. Let $F \subseteq \mathscr{F}$ be the progress measures of G.

1) $\sqcap F$ is smallest game progress measure.

2) The set $\{ v \mid \sqcap F(v) < \top \}$ is the set of vertices from which Maximizer has a strategy to ensure positive energy level at all time.



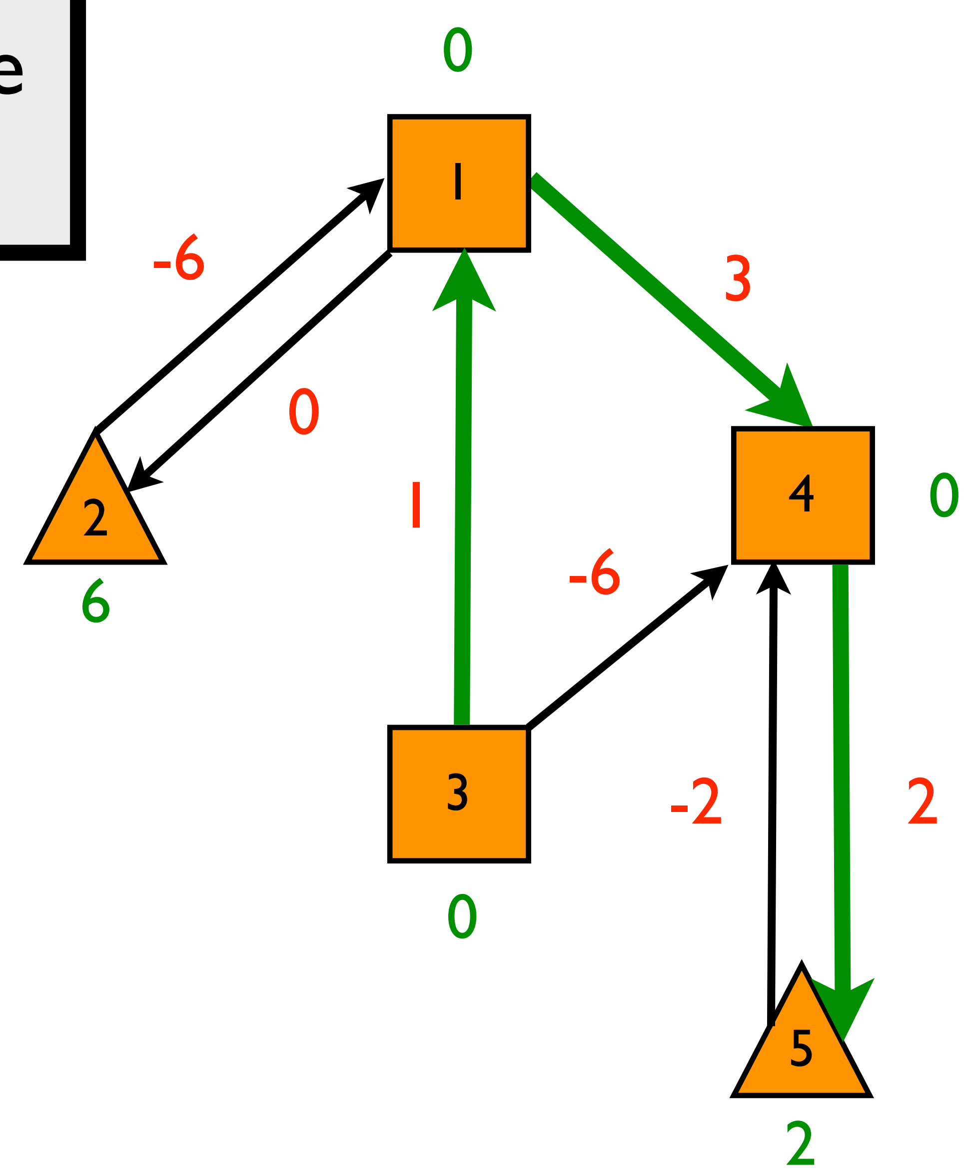
How to synthesize energy progress measures ?

Thm. Let $F \subseteq \mathscr{F}$ be the progress measures of G.

1) $\sqcap F$ is smallest game progress measure.

2) The set $\{ v \mid \sqcap F(v) < \top \}$ is the set of vertices from which Maximizer has a strategy to ensure positive energy level at all time.

$\sqcap F(v)$ is the minimal credit needed by Maximizer to win from v.



How to synthesize energy progress measures ?

Theorem. Let $F \subseteq \mathscr{F}$, $\sqcap F$ is lfp $\delta : \mathscr{F} \rightarrow \mathscr{F}$:

- -if v of **Maximizer**: $\delta(f)(v) =$ **min** $\{ f(v') - w(v,v') \mid (v,v') \in E \}$
- -if v of **Minimizer**: $\delta(f)(v) =$ **max** $\{ f(v') - w(v,v') \mid (v,v') \in E \}$

Remember:

$f \in \mathscr{F}$ is a **game energy progress measure** iff

-if v belongs to Maximizer: $f(v) \geq f(v') - w(v,v')$ for **some** $(v,v') \in E$;

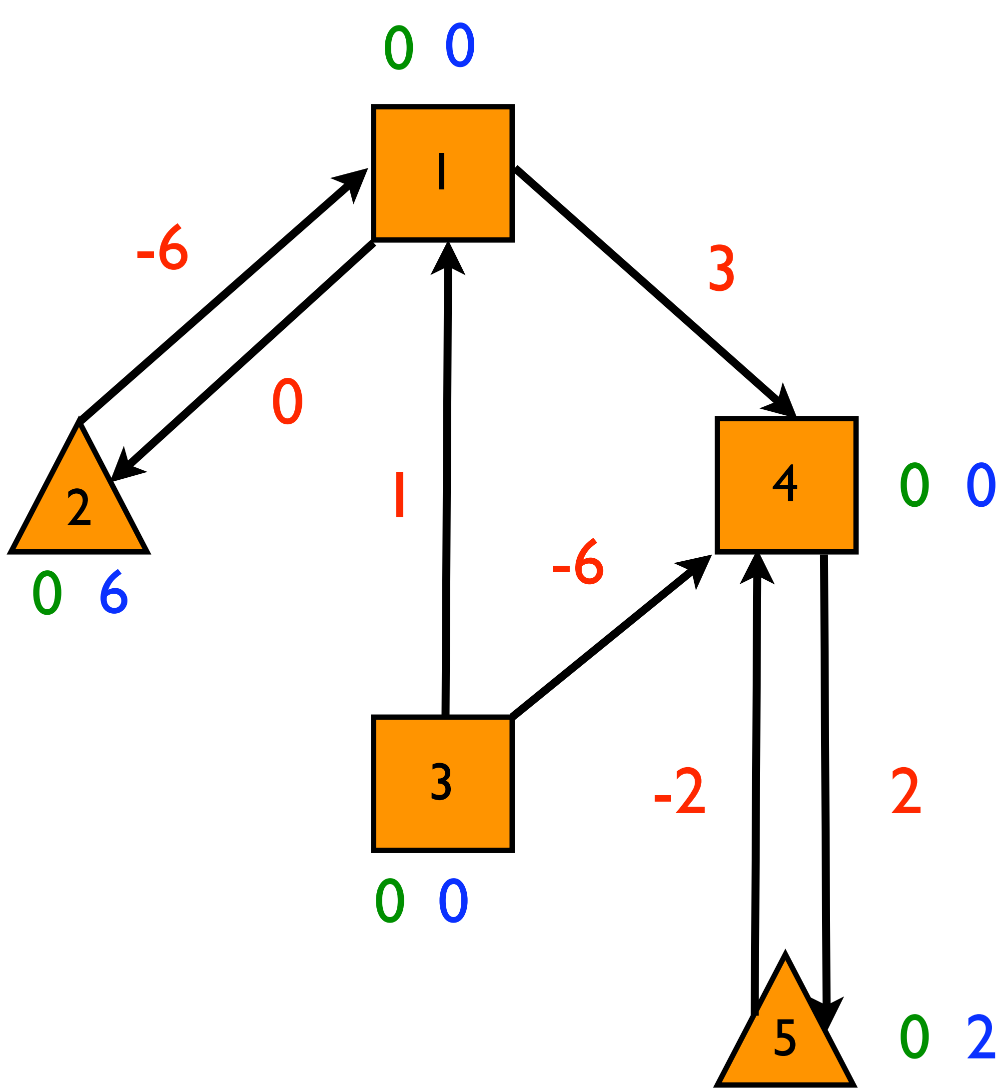
-if v belongs to Minimizer: $f(v) \geq f(v') - w(v,v')$ for **all** $(v,v') \in E$.

How to synthesize energy progress measures ?

Theorem. Let $F \subseteq \mathscr{F}$, $\sqcap F$ is lfp $\delta : \mathscr{F} \rightarrow \mathscr{F}$:

-if v of **Maximizer**: $\delta(f)(v) =$ **min** $\{ f(v') - w(v,v') \mid (v,v') \in E \}$

-if v of **Minimizer**: $\delta(f)(v) =$ **max** $\{ f(v') - w(v,v') \mid (v,v') \in E \}$



This least fixed point can be computed by successive approximation from $f_0=(0,0,0,0,0)$.

(0,0,0,0,0)

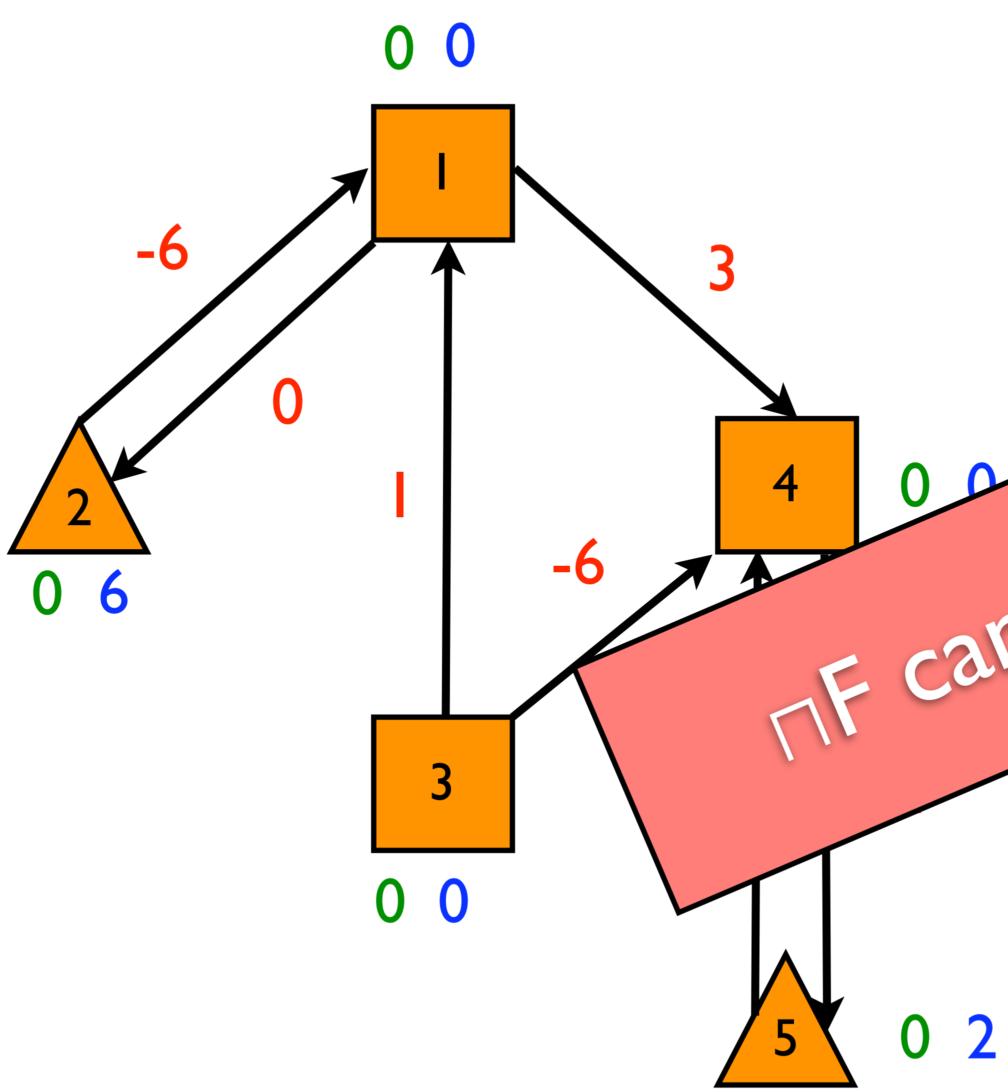
(0,6,0,0,2) $\Longrightarrow$ fixed point $= \sqcap F$

$\sqcap F(v)$ gives the necessary initial credit

How to synthesize energy progress measures ?

Theorem. Let F⊆$\mathscr{F}$, ⊓F is lfp δ : $\mathscr{F} \rightarrow \mathscr{F}$:

- -if v of **Maximizer**: δ(f)(v)= **min** { f(v')-w(v,v') | (v,v') ∈ E}
- -if v of **Minimizer**: δ(f)(v)= **max** { f(v')-w(v,v') | (v,v') ∈ E}



This least fix... d by
succe... $_0$=(0,0,0,0,0).

...,0,0,2) ⟹ fixed point
= ⊓F

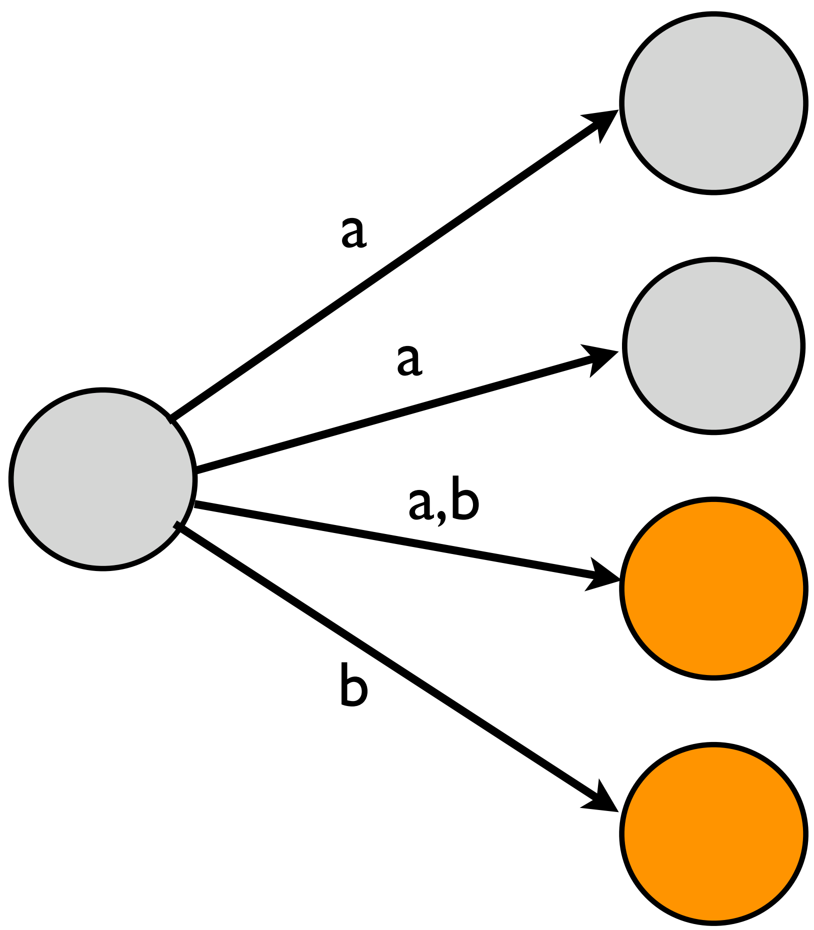
⊓F can be computed in O(E•M$_G$)

⊓F(v) gives the necessary initial credit

Energy games with imperfect information

Adding imperfect information to EG

- In game with **imperfect** information, Maximizer does not have a perfect **knowledge** of the current state of the game.
- Maximizer gets information while receiving **observations**.
- An observation is a class of **equivalent states**.
- A game is given as a tuple: $(Q,q_0,Act,Tr,Part,w)$, and we consider 3 variants:
  - $w : Act \to \mathbb{Z}$ (costs on actions, visible)
  - $w : Part \to \mathbb{Z}$ (cost on observations, visible)
  - $w : Tr \to \mathbb{Z}$ (costs on transitions, invisible)



Adding imperfect information to EG

- Now the game is played as follows:

  1. Maximizer chooses an **action**;
  2. Minimizer **resolves nondeterminism** by choosing the next position in the set of positions that are compatible with the action choosen by Maximizer;
  3. Maximizer receives the **observation** that is compatible with the position choosen by Minimizer.

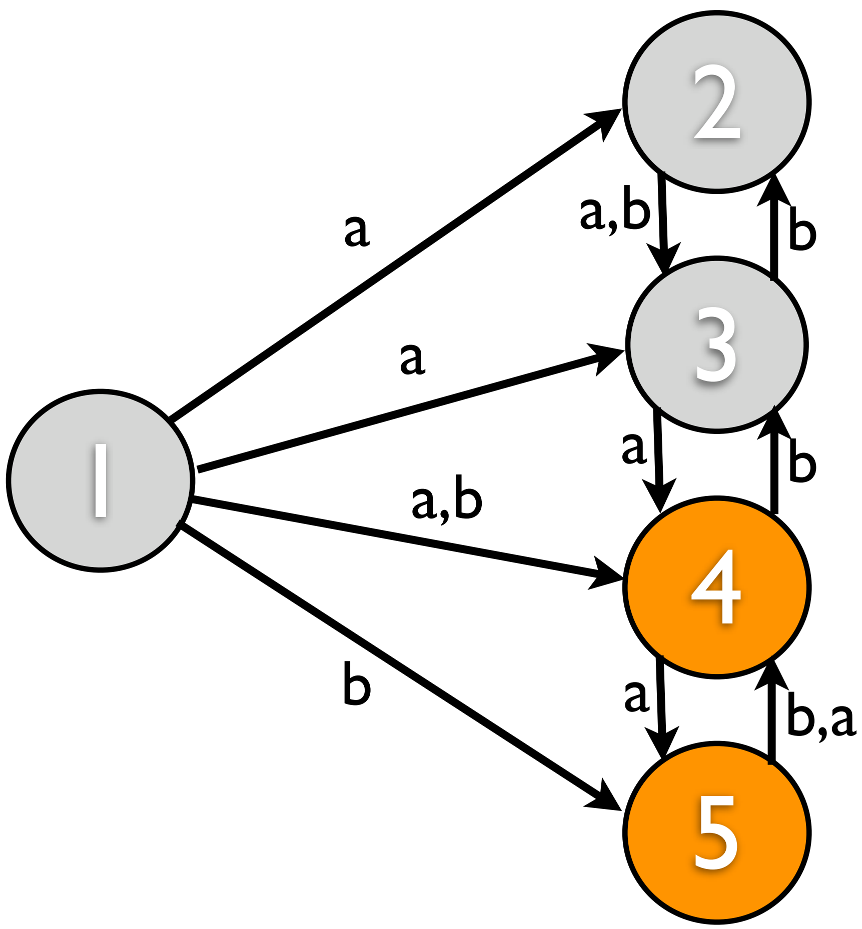
- The objective of Maximizer is to maintain energy level positive at all points in time.

- We will consider two variants :
  either the initial energy level is fixed or it is quantified existentially.

Knowledge

- In game with imperfect information, Maximizer does not see the current state of the game. She has only a **knowledge** of the current state of the game. We make the hypothesis that she knows about the initial energy level.
- We must distinguish the case where the energy consumption/energy gains are visible from the case where there are not:
  - in the visible case: the knowledge is a set of states in which the game can possibly be and an energy level, i.e. (S,v) where $S \subseteq Q$;
  - in the invisible case: the knowledge is a set of pairs of (q,v) where q is a state and v is an energy level. It can be argued that only the worst case are interesting, so in this case the knowledge if a function f : $Q \rightarrow \mathbb{N} \cup \{+\infty\} \cup \{\bot\}$.

Illustration

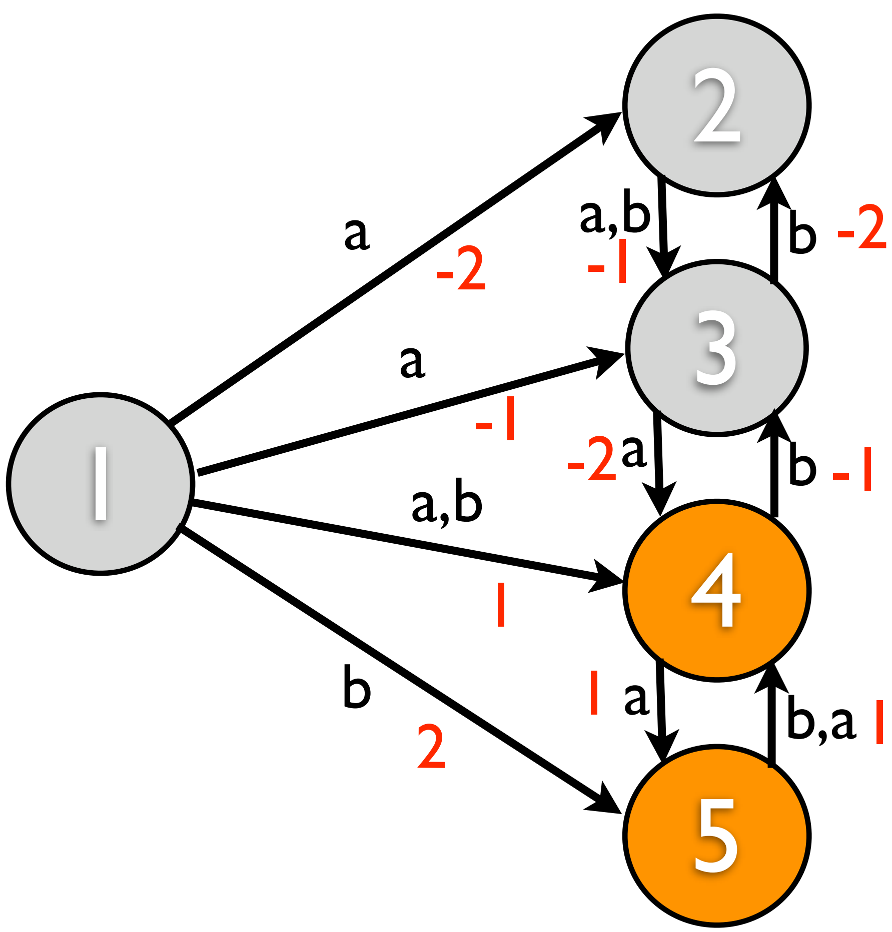


Assume, we start with energy level 10.

Assume that grey state consumes 1 and an orange state delivers 1.

	a	b	b	a	a	...
	2	3	2	3	4	...
	g	g	g	g	o	...
{1}	{2,3}	{2,3}	{2,3}	{3}	{4}	...
10	9	8	7	6	7	

Illustration



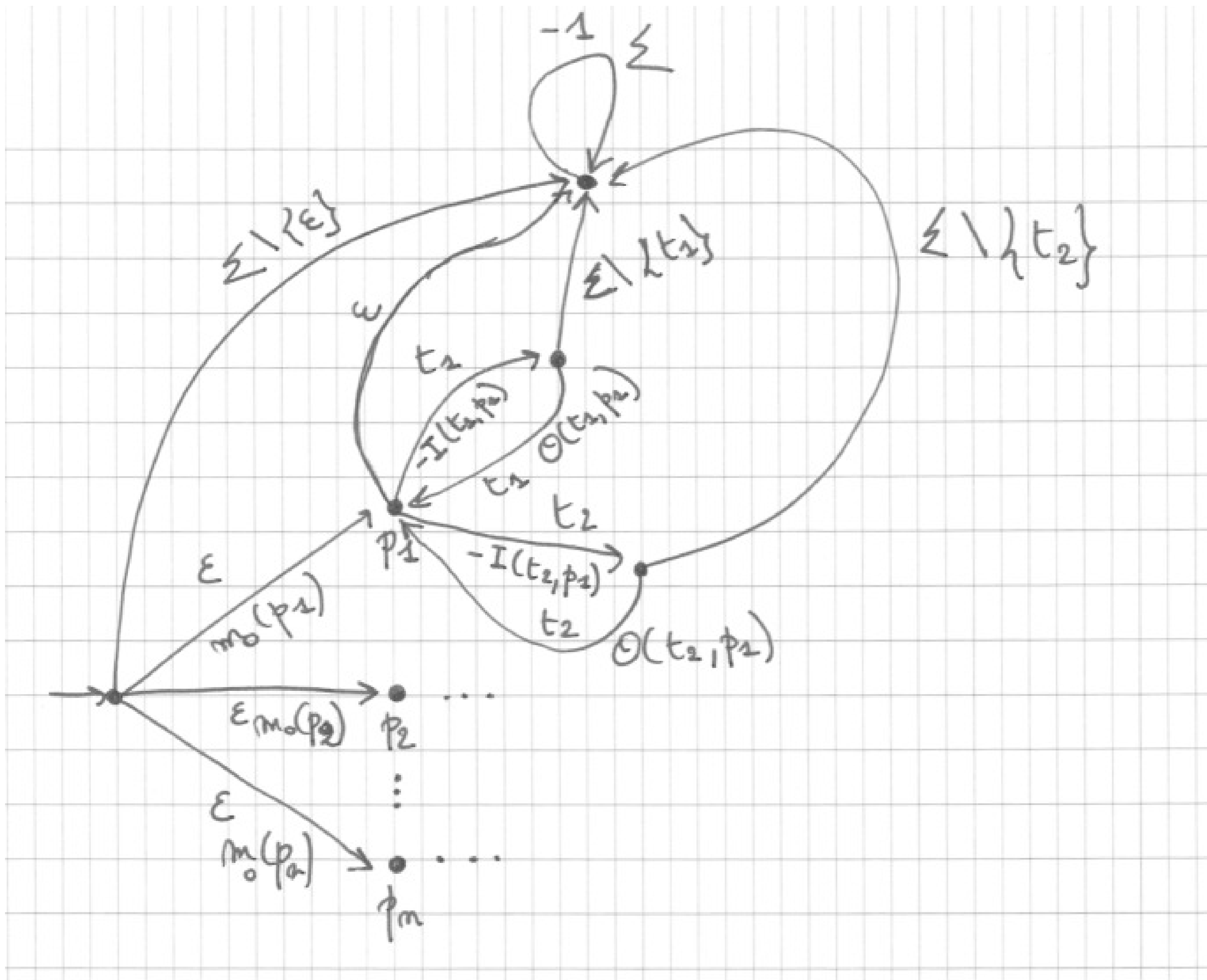
Assume, we start with energy level 10.

Assume costs are on transitions.

	a	b	b	...
	2	3	2	...
	g	g	g	...
{(1,10)}	{(2,8),(3,9)}	{(3,7),(2,7)}	{(2,5),(3,6)}	...

Invisible energy gains/consumptions and Petri Nets

For any Petri net **N** with initial marking $\mathbf{m_0}$, we can construct an EG with imperfect information and energy gains/consumptions on transitions, such that there exists an infinite execution starting from $\mathbf{m_0}$ in **N** iff Maximizer can win the EG with fixed initial energy level.

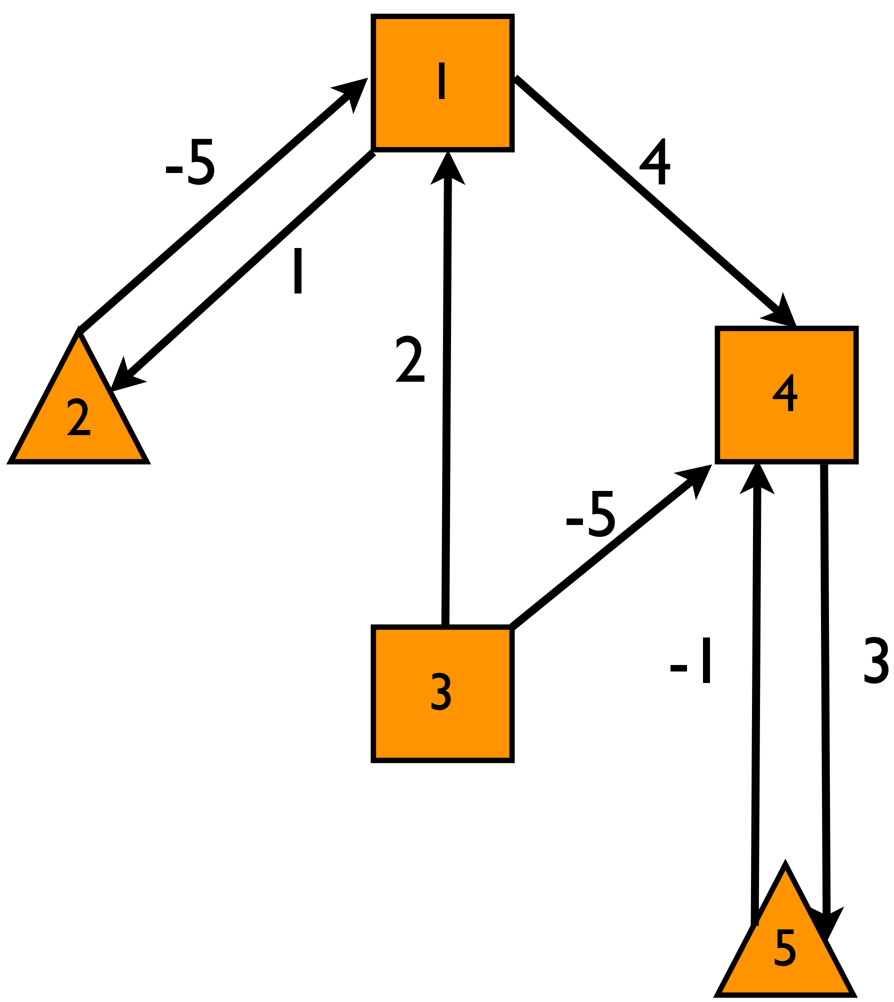


First Results

	Visible energy gains/comsumptions	Invisible energy gains/consumptions
Fixed initial energy level	ExpTime-C (subset constr.)	ExpSpace-H (PN infinite execution) Recursive (finite game tree - wqo)
Unfixed initial energy level	ExpTime-C (subset constr.)	?


iStockphoto

Mean-payoff Games



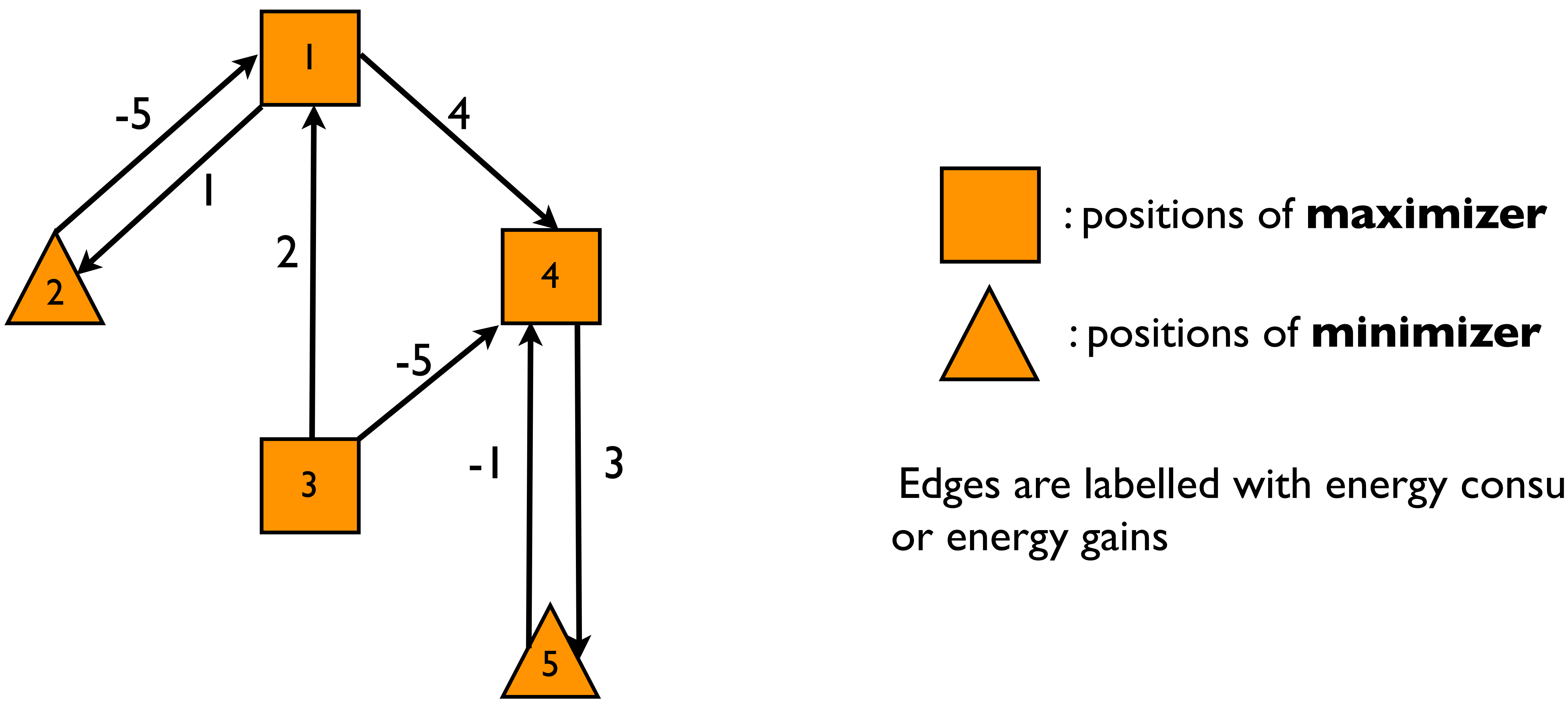
: positions of **maximizer**

: positions of **minimizer**

Edges are labelled with rewards

A **mean-payoff game** starts in a state. At each round, the owner of the state chooses an **edge** from that state. The game proceeds from the target state. The game is played for an **infinite** number of such rounds. The **limit average** of the rewards associated to edges that are traversed is the value of the game. Maximizer want to maximize this value, Minimizer want to minimize it.

Energy Games



Edges are labelled with energy consumptions or energy gains

A **energy game** starts in a state with an **initial energy level**. At each round, the owner of the state chooses an **edge** from that state. The game proceeds from the target state of this edge. The game is played for an **infinite** number of such rounds. At any point in time, the initial energy level plus the **sum of edge values** traversed so far defines the energy level of the game at that point. Maximizer want to maintain this value positive at all time, Minimizer want to reach a negative energy level.