

Subgame perfect equilibrium with an algorithmic perspective

Jean-François Rasbin (Université libre de Bruxelles)

joint work with Marie van den Bogaard (Université Gustav Eiffel)

Léonard Brice (ENS Paris-Saclay)

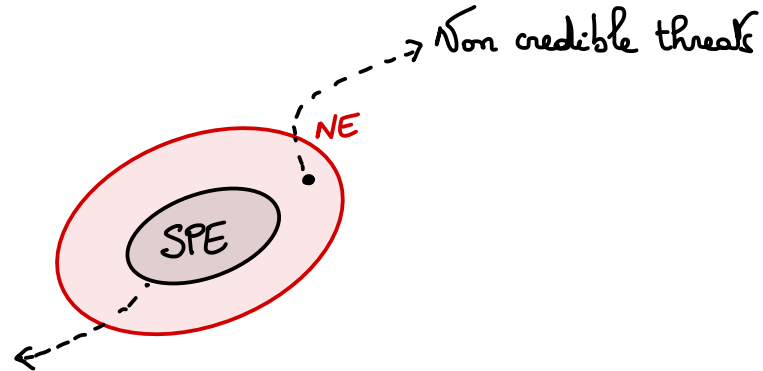
March 26, 2021.

Simons Institute for the Theory of Computing (Berkeley)

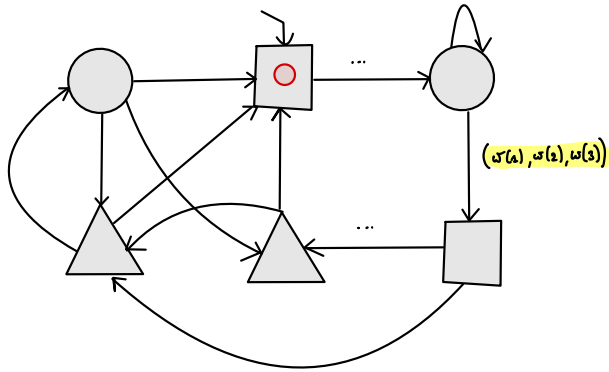
Objectives of the talk

- ① Recall SPE for sequential games
- ② Explore some recent progresses in algorithms to handle SPEs (for **mean-payoff** objectives) → left open in the literature

Q: how to obtain an effective representation of all possible outcomes produced by **rational** players?



N player turn-based graph games



Vertices are *partitioned*

$$V = V_1 \uplus V_2 \uplus \dots \uplus V_N$$

V_i = vertices of Player $i \in [1, N]$

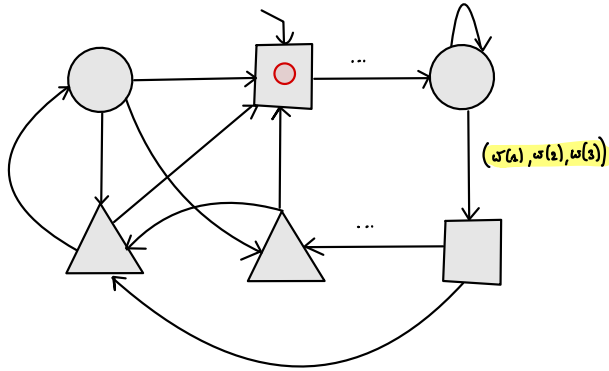
$$E \subseteq V \times V, \omega: E \rightarrow \mathbb{Z}^N$$

Play $p \in V^\omega$, payoffs given by $\mu_i: V^\omega \rightarrow \mathbb{R}, i \in [1, N]$

↳ *mean-payoff* $\mu_i(p)$

$$\text{MP}_i(p) = \liminf_{n \rightarrow \infty} \frac{1}{n} \cdot \sum_{j=0}^{n-1} \omega(v_j, v_{j+1})(i)$$

N player turn-based graph games



Vertices are *partitioned*

$$V = V_1 \uplus V_2 \uplus \dots \uplus V_N$$

V_i = vertices of Player $i \in [1, N]$

$$E \subseteq V \times V, \omega: E \rightarrow \mathbb{Z}^N$$

Play $p \in V^\omega$, payoffs given by $\mu_i: V^\omega \rightarrow \mathbb{R}, i \in [1, N]$

↳ *mean-payoff* $\mu_i(p)$

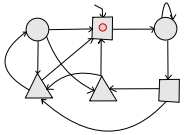
$$\text{MP}_i(p) = \liminf_{n \rightarrow \infty} \frac{1}{n} \cdot \sum_{j=0}^{n-1} \omega(v_j, v_{j+1})(i)$$

Rationality: each player wants to *maximize* his own payoff.

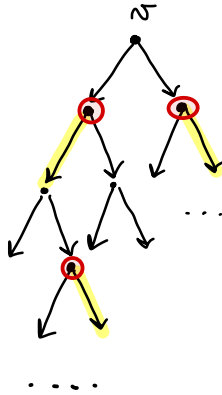
Strategies, profiles, outcomes

$$\sigma_i : V^* \cdot V_i \rightarrow E$$

$$\text{s.t. } \forall \pi, \omega \in V^* \cdot V_i : (\pi, \sigma_i(\pi, \omega)) \in E$$



∞-unfolding

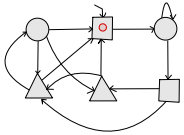


$\Sigma_i = \text{set of strategy of } P, i \in [1, N]$

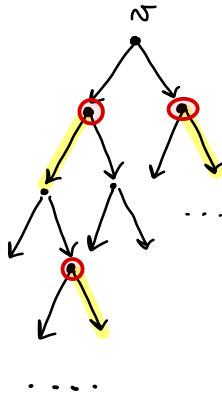
Strategies, profiles, outcomes

$$\sigma_i : V^* \cdot V_i \rightarrow E$$

$$\text{s.t. } \forall \pi \cdot v \in V^* \cdot V_i : (\pi, \sigma_i(\pi \cdot v)) \in E$$



∞ -unfolding



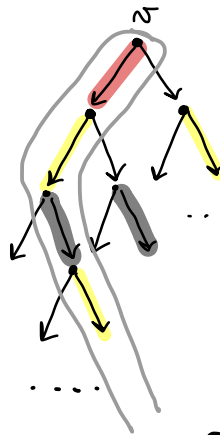
$\Sigma_i = \text{set of strategy of } P, i \in [1, N]$

Profiles:

$$(\sigma_1, \sigma_2, \dots, \sigma_N) \in \Sigma_1 \times \Sigma_2 \times \dots \times \Sigma_N$$

$$= (\sigma_i, \bar{\sigma}_{-i})$$

all strategies but σ_i



- σ_1
- σ_2
- σ_3
- ...

$$= \text{Out}_{\sigma}(\sigma_1, \sigma_2, \dots, \sigma_N) = p \in V^{\omega}$$

$$\text{Out}_{\sigma_0}(\bar{\sigma}) = \sigma_0 \sigma_1 \sigma_2 \dots \sigma_n \dots = p$$

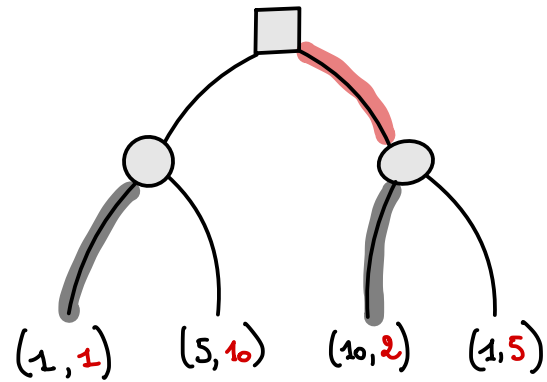
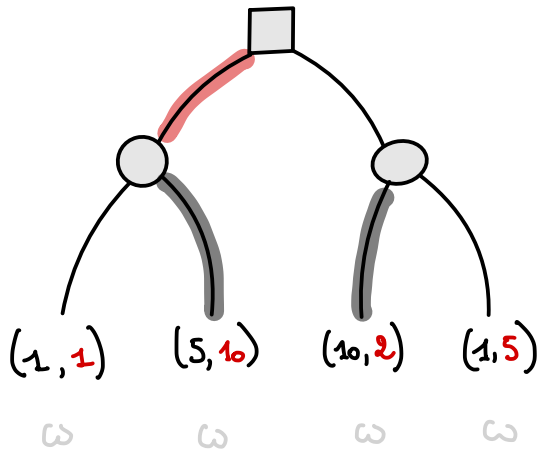
$$\text{s.t. } \sigma_0 = \sigma, \forall j \geq 0 : \text{if } p(j) \in V_i : \sigma_{j+1} = \sigma_i(p(0..j))$$

Nash equilibrium and subgame perfect equilibrium

A profile of strategies $(\sigma_1, \sigma_2, \dots, \sigma_N)$ is a **Nash equilibrium (NE)** in σ_0 , if for all $i \in [1, N]$, for all $\sigma_i' \in \Sigma_i$: $\mu_i(Q_{\sigma_0}(\bar{\sigma}_{-i}, \sigma_i')) \leq \mu_i(Q_{\sigma_0}(\bar{\sigma}_{-i}, \sigma_i))$
= No player has an incentive to deviate unilaterally.

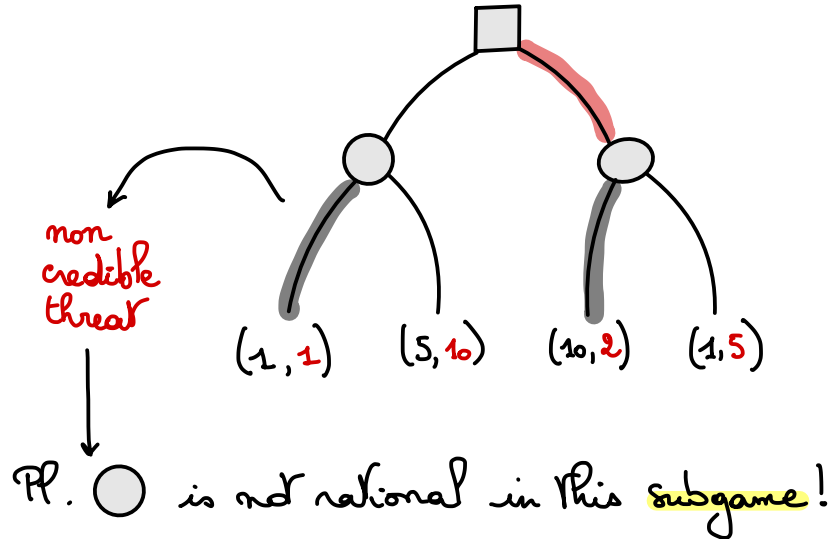
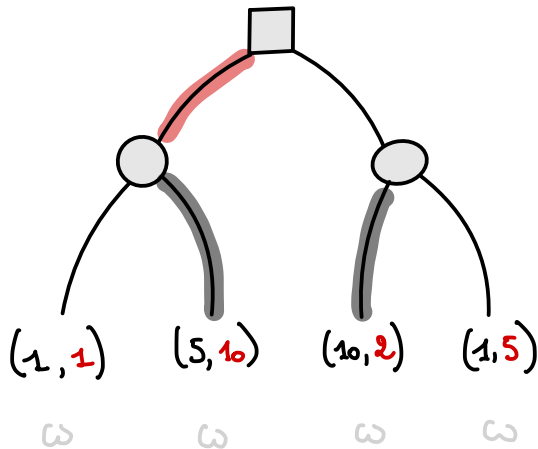
Nash equilibrium and subgame perfect equilibrium

A profile of strategies $(\sigma_1, \sigma_2, \dots, \sigma_N)$ is a **Nash equilibrium (NE)** in σ_0 , if for all $i \in [1, N]$, for all $\sigma_i' \in \Sigma_i$: $\mu_i(Q_{\sigma_0}(\bar{\sigma}_{-i}, \sigma_i')) \leq \mu_i(Q_{\sigma_0}(\bar{\sigma}_{-i}, \sigma_i))$
 = No player has an incentive to deviate unilaterally.

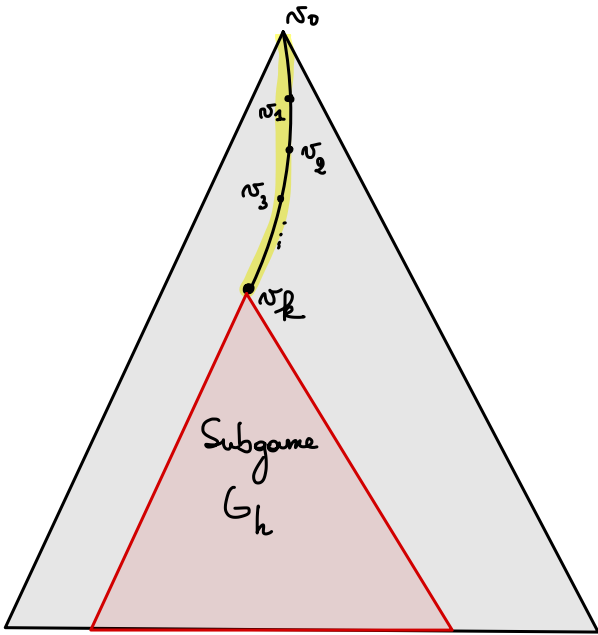
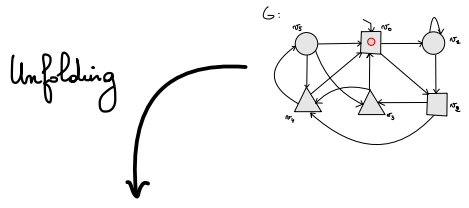


Nash equilibrium and subgame perfect equilibrium

A profile of strategies $(\sigma_1, \sigma_2, \dots, \sigma_N)$ is a **Nash equilibrium (NE)** in σ_0 , if for all $i \in [1, N]$, for all $\sigma_i' \in \Sigma_i$: $\mu_i(Q_{\sigma_0}(\bar{\sigma}_{-i}, \sigma_i')) \leq \mu_i(Q_{\sigma_0}(\bar{\sigma}_{-i}, \sigma_i))$
 = No player has an incentive to deviate unilaterally.

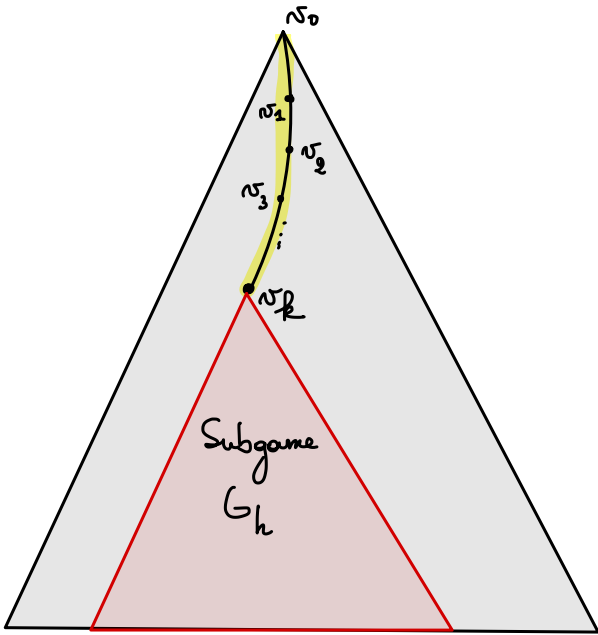
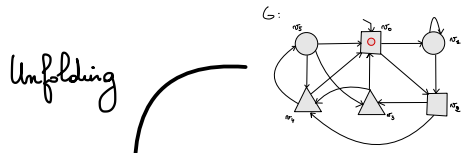


Subgame

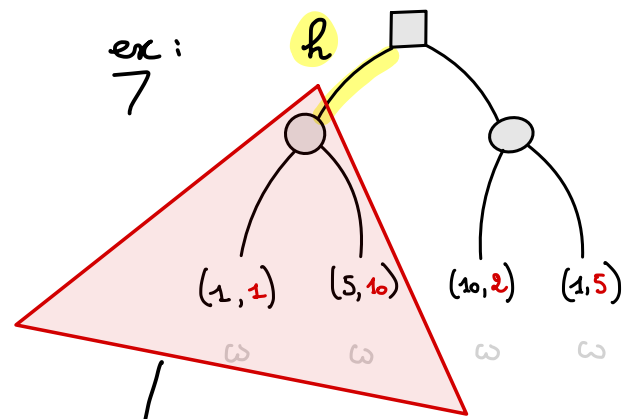


Every history $h = n_0 n_1 \dots n_R$ defines a subgame G_h

Subgame



Every history $h = v_0 v_1 \dots v_k$ defines a subgame G_h



Subgame induced by h
for $h = h'.v$,

$$Out_h(\bar{\sigma}) = h'. Out_v(\bar{\sigma}^l)$$

Nash equilibrium and subgame perfect equilibrium

A profile of strategies $(\sigma_1, \sigma_2, \dots, \sigma_N)$ is a **subgame perfect equilibrium (SPE)**

if for all subgames G_h of G , for all P. $i \in [1, N]$, for all $\sigma_i^j \in \Sigma_i$:

$$\mu_i(\text{Pay}_h(\bar{\sigma}_{-i}, \sigma_i^j)) \leq \mu_i(\text{Pay}_h(\bar{\sigma}_{-i}, \sigma_i^h)).$$

→ Players must be rational in all subgames (→ no non-credible threats)

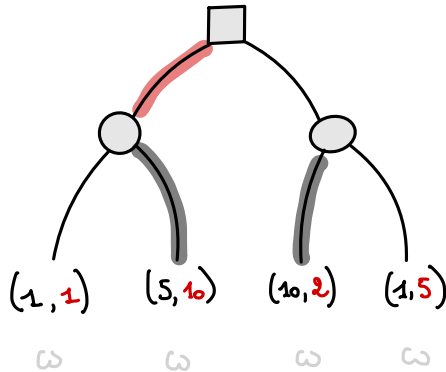
Nash equilibrium and subgame perfect equilibrium

A profile of strategies $(\sigma_1, \sigma_2, \dots, \sigma_N)$ is a **subgame perfect equilibrium (SPE)**

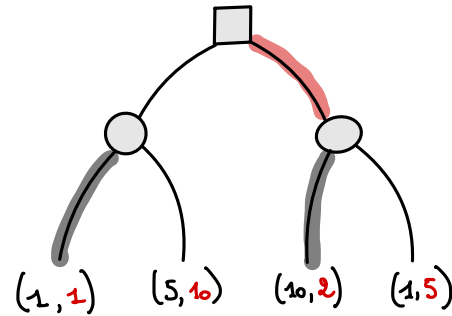
if for all subgames G_h of G , for all P. $i \in [1, N]$, for all $\sigma_i^h \in \Sigma_i$:

$$u_i(\sigma_{-i}^h, \sigma_i^h) \leq u_i(\sigma_{-i}^h, \sigma_i^h).$$

→ Players must be rational in all subgames (→ no non-credible threats)



SPE



~~SPE~~

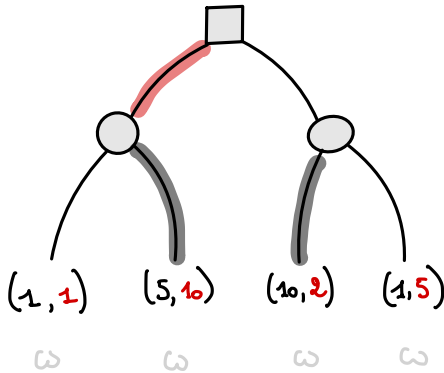
Nash equilibrium and subgame perfect equilibrium

A profile of strategies $(\sigma_1, \sigma_2, \dots, \sigma_N)$ is a **subgame perfect equilibrium (SPE)**

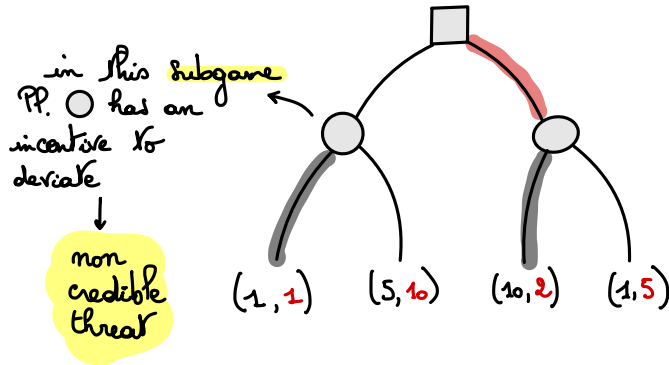
if for all subgames G_h of G , for all P. $i \in [1, N]$, for all $\sigma_i^h \in \Sigma_i$:

$$u_i(\text{Out}_h(\bar{\sigma}_{-i}^h, \sigma_i^h)) \leq u_i(\text{Out}_h(\bar{\sigma}_{-i}^h, \sigma_i^h)).$$

→ Players must be rational in all subgames (→ no non-credible threats)



SPE

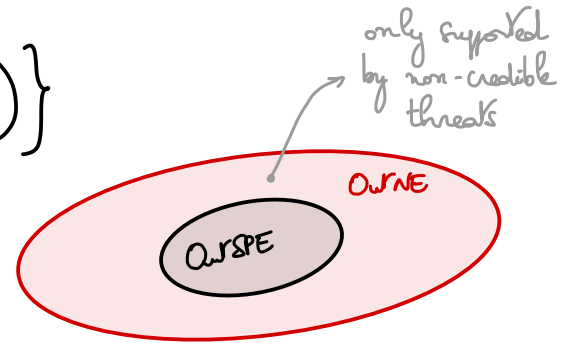


~~SPE~~

Outcomes supported by equilibria

$$Q_{\text{NE}}(G) = \bigcup_{\bar{\sigma} \in \text{NE}} \{ \text{Outcome}_{\bar{\sigma}}(\bar{\sigma}) \}$$

$$Q_{\text{SPE}}(G) = \bigcup_{\bar{\sigma} \in \text{SPE}} \{ \text{Outcome}_{\bar{\sigma}}(\bar{\sigma}) \}$$



↳ Sets of behaviors induced by NE/SPE -rational players in G

SPE - Known results

Existence

- guaranteed when $(\mu_i)_{i \in [1, N]}$ are **continuous** (ex: discounted sum, quantitative reach.)

↳ can be extended to lower semi-continuous [Fleisch et al.]

or $(\mu_i)_{i \in [1, N]}$ are **B-omega-regular objectives** (ex: parity)

↳ Established by Ummels (2006) using transfinite induction

- **not** guaranteed for **mean-payoff objectives** (see example)

Effective Representation

$\text{QwSPE}(G)$

- for **quantitative reachability** [CONCUR'19]

- ... also for **B-omega regular obj.** (ex: parity) [Ummels 06]

- **TODAY**: **mean-payoff objectives** [arXiv: 2101.10685]

alternating tree automata

Effective representation of the set of outcomes of NE/SPE

why?

Effective representation of the set of outcomes of NE/SPE

Why?

① Existence problem for SPE : $\text{Out-SPE}(G) \neq \emptyset$

Effective representation of the set of outcomes of NE/SPE

Why?

① Existence problem for SPE : $Out-SPE(G) \neq \emptyset$

② Quantitative rational verification ^{→ B, NE} [Wooldridge et al. 17] for NE.

→ \forall NE/SPE : do **all** behaviors by rational agents satisfy some spec. ψ ?

all: $Out-NE(G) \not\models [\psi]$ $Out-SPE(G) \not\models [\psi]$

Some: $Out-NE(G) \cap [\psi] \neq \emptyset$ $Out-SPE(G) \cap [\psi] \neq \emptyset$

→ idem for $\Gamma_{\sigma_0} \otimes G_{0 \cup [1, N]}$
 ↪ System, $[1, N] = Env.$

Effective representation of the set of outcomes of NE/SPE

Why?

① Existence problem for SPE : $\text{Out-SPE}(G) \neq \emptyset$

② Quantitative rational verification [Wooldridge et al. 17] for NE.

→ \forall NE/SPE : do **all** behaviors by rational agents satisfy some spec. ψ ?

all: $\text{Out-NE}(G) \not\models [\psi]$ $\text{Out-SPE}(G) \not\models [\psi]$

Some: $\text{Out-NE}(G) \cap [\psi] \neq \emptyset$ $\text{Out-SPE}(G) \cap [\psi] \neq \emptyset$

→ idem for $\Gamma_{\sigma_0} \otimes G_{0 \cup [1, N]}$
↳ System, $[1, N] = \text{Env.}$

③ Cooperative rational synthesis: $\exists p \in \text{Out-}\left\{ \begin{matrix} \text{SPE} \\ \text{NE} \end{matrix} \right\}(G) : p \models \psi_0$

[Kupferman et al. 12]

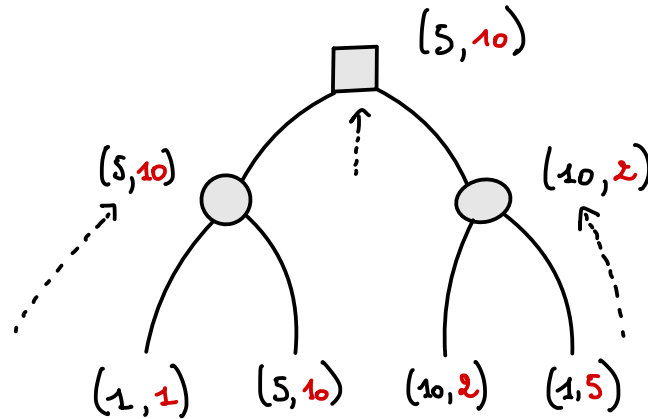
(first seminar)

↳ constrained- \exists

↳ Spec. Syst.

How to reason on SPE ?

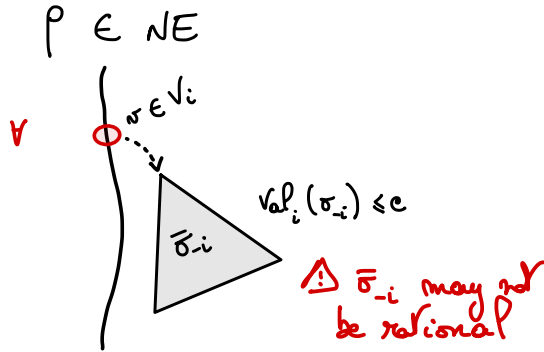
→ finite trees : conceptually easy using **backward induction**



→ infinite trees : **backward induction** does **not** generalize well...

Starting point:
NE in infinite duration games

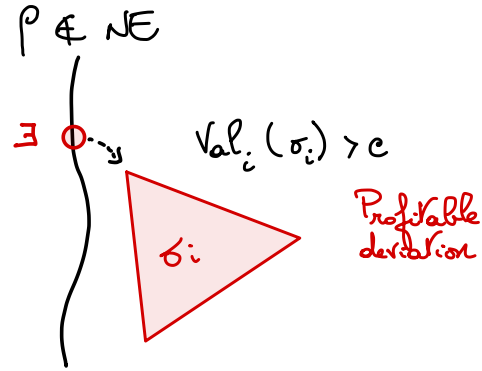
NE - Deviation - Punishment



$Val_i(p) = c$ and

$c \Rightarrow \inf_{\sigma_i} \sup_{\sigma_{-i}} Val(O_{w,r}(\sigma_i, \sigma_{-i}))$

worst-case value for P_i



$Val_i(p) = c$ and

$\sup_{\sigma_i} \inf_{\sigma_{-i}} Val(O_{w,r}(\sigma_i, \sigma_{-i})) > c$

→ This works for any prefix independent payoff function for which worst-case value exists and can be realized

→ OK for HP

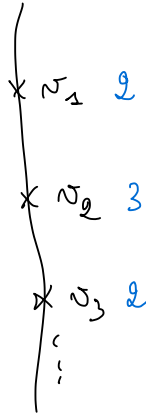
→ [Brhaye et al. 13]

Set of outcomes supported by NE - MP

→ Requirement: $\lambda: V \rightarrow \mathbb{R} \cup \{-\infty, +\infty\}$

→ A path $p = v_0 v_1 \dots v_n \dots$ is λ -consistent if

$$\forall i \in [1, N]: \underline{MP}_i(p) \geq \max_{v \in \text{Visit}(p) \cap V_i} \lambda(v_i)$$



Set of outcomes supported by NE - MP

→ Requirement: $\lambda: V \rightarrow \mathbb{R} \cup \{-\infty, +\infty\}$

→ A path $p = \sigma_0 \sigma_1 \dots \sigma_n \dots$ is λ -consistent if

$$\forall i \in [1, N]: \underline{MP}_i(p) \geq \max_{\sigma \in \text{Visit}(p) \cap V_i} \lambda(\sigma_i)$$

→ **Key**: The **worst-case value** that Pl. i can force against all the other players.

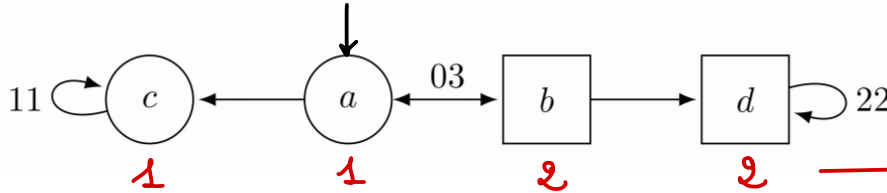
→ $\forall \sigma \in V_i$: let $\lambda_1(\sigma) = \sup_{\sigma_i \in \Sigma_i} \inf_{\sigma_{-i} \in \Sigma_{-i}} \underline{MP}_i(\text{Out}_\sigma(\sigma_i, \sigma_{-i}))$ ↳ requirement induced by WCV

Theorem: $p = \sigma_0 \sigma_1 \dots \sigma_n \dots$, $p \in \text{OutNE}(G)$ iff p is λ_1 -consistent.

↳ p gives Pl. i at least the wcv that he can force along p → if not → profitable deviation.

Set of outcomes supported by NE - MP

- an example



↳ The set of λ_1 -consistent paths in G are :

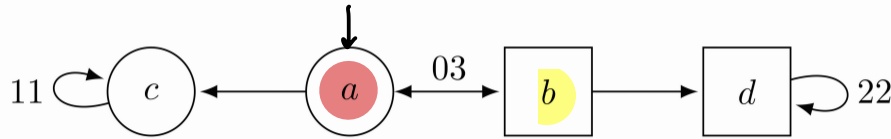
$$\{a \rightarrow c^\omega\} \cup \bigcup_{k \in \mathbb{N}} \{a \rightarrow (b \rightarrow a)^k \rightarrow b \rightarrow d^\omega\}$$

λ_1
 worst-case values

Corollary: The set of λ -consistent paths is recognized by a multi-MP automaton. This language is not necessarily ω -regular.

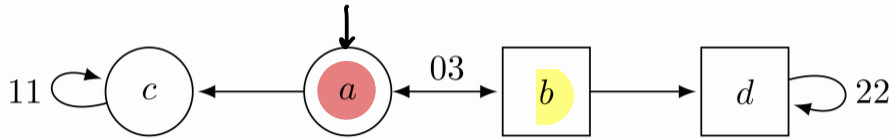
[Alur et al. 09, Chatterjee et al. 10]

A MP game without SPE



- PP. \circ can secure 1 from a ($a \rightarrow c$)
- PP. \square can secure 2 from b ($b \rightarrow d$)
- So there is no NE in which $a \leftrightarrow b$ is taken for ever as PP. \circ would have an incentive to leave ($a \rightarrow c$) but then PP. \square would prefer to leave before PP. \circ

A MP game without SPE



→ PP. \circ can secure 1 from a ($a \rightarrow c$)

→ PP. \square can secure 2 from b ($b \rightarrow d$)

→ So there is no NE in which $a \leftrightarrow b$ is taken for ever as PP. \circ would have an incentive to leave ($a \rightarrow c$) but then PP. \square would prefer to leave before PP. \circ

∴ → From a PP. \circ knows that PP. \square will leave, PP. \circ has then no incentive to do it before (as he will then get '2 instead of 1')

But → then PP. \square has no interest to leave as he receives 3 on the cycle.

→ Need to iterate the reasoning on worst-case value.

Generalization :
The negotiation function

? : Given λ_1 and v , can the player that controls v improve the value that she can obtain against the other players if the other players are not willing to give away their worst-case value (λ_1)?

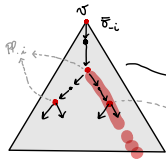
The negotiation function

? : Given λ_2 and v , can the player that controls v improve the value that she can obtain against the other players if the other players are not willing to give away their worst-case value (λ_1)?

$$\text{Nego} : [\lambda \rightarrow \mathbb{R} \cup \{-\infty, +\infty\}] \rightarrow [\lambda \rightarrow \mathbb{R} \cup \{-\infty, +\infty\}]$$

$$\text{let } v \in V_i, \quad \text{Nego}(\lambda)(v) = \inf_{\bar{v}_{-i} \in \lambda \text{Rat}(v)} \sup_{\sigma_i \in \Sigma_i} \underline{MP}_i(\text{Out}_v(\sigma_i, \bar{v}_{-i}))$$

↗ $\inf \phi = +\infty$



\bar{v}_{-i} is $\lambda \text{Rat}(v)$: in all \bar{v}_{-i} consistent history h , there is a \bar{v}_{-i} continuation that is λ -consistent.

↳ worst case value against λ -rational players.

How to compute $\text{Nego}(\cdot)$?

$$\text{Nego}(\lambda)(v) = \inf_{\bar{\sigma}_{-i} \in \lambda \text{Rat}(v)} \sup_{\sigma_i \in \Sigma_i} \underline{MP}_i(\text{Out}_v(\sigma_i, \bar{\sigma}_{-i}))$$

Handwritten notes: $v \in V_i$ (pointing to v), $\lambda \text{Rat}(v)$ (highlighted)

if the inf is always realizable \rightarrow **STEADY NEGOTIATION**
 \hookrightarrow OK for \underline{MP}

What is the relative worst-case value that P_i can force against λ -rational adversaries?

Prover want to prove $\text{Nego}(\lambda)(v) \leq \alpha$ to Challenger

P **C**

P **C** game to determine if $\text{Nego}(\lambda)(v) \leq \alpha$?

P proposes outcomes $p = \sigma_0 \sigma_1 \dots \sigma_n \dots \in \lambda$ -consistent if this is possible
 (if not possible FAIL)

C either accepts and the game ends (ACCEPT)
 or proposes a deviation after a history h that belongs to P.i (DEV)

$p = \sigma_0 \sigma_1 \dots \sigma_k \overset{\sigma'_k}{\neq} \sigma_{k+1} \dots \rightarrow$ current prefix $\pi = \sigma_0 \sigma_1 \dots \sigma_k \sigma'$

P proposes an outcome $p = \pi \sigma_0 \sigma_1 \dots \sigma_k \dots$ that extends π and such that $\sigma_0 \sigma_1 \dots \sigma_k \dots$ is λ -consistent
 ...
 (if not possible FAIL)

P **C** game - winning condition

either **C** puts an end to deviations $\rightarrow p = \pi_1 \cdot \pi_2 \cdot \dots \cdot \pi_n \cdot p'$ and P. i receives $\underline{MP}_i(p)$
(ACCEPT)

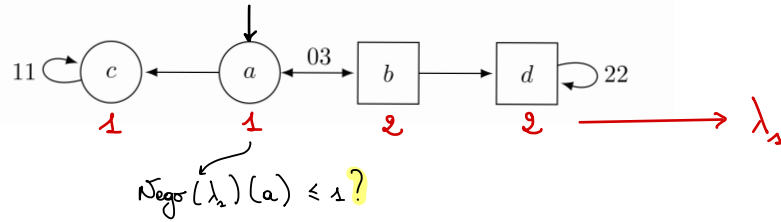
or **C** deviates ∞ -many times and the resulting play is $p^\infty = \pi_1 \cdot \pi_2 \cdot \dots \cdot \pi_n \dots$
and P. i receives $\underline{MP}_i(p^\infty)$. (DEV ∞)

or **P** fails at some round to propose a λ -confident continuation (FAIL)
and P. i receives $+\infty$.

P wins iff P. i receives a payoff $\leq \alpha$.
(This zero-sum game is determined)

How to compute Nego(.) ?

- an example - $Nego(\lambda_2)$



\mathbb{P} : $a \cdot c^\omega$, $a \cdot c^\omega$ is λ_1 -consistent and $\underline{MP}_0(a \cdot c^\omega) = 1$

\mathbb{C} : deviation : $a \rightarrow b$

\mathbb{P} : from b, the only λ_2 -consistent paths are $(ba)^* \cdot d^\omega$
 even if $(ab)^\omega$ is tempting

and $\underline{MP}_1((ba)^* \cdot d^\omega) = 2 \Rightarrow \mathbb{C}$ wins

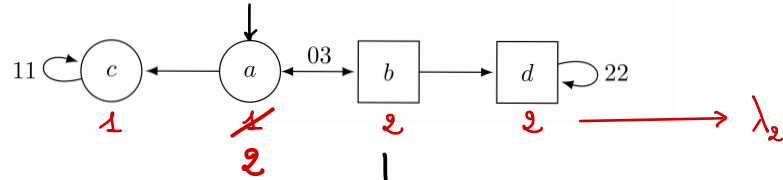
~~$Nego(\lambda_2)(a) < 1$~~

\rightarrow generalisation : $Nego(\lambda_2)(a) = 2$

it does not give 1 to \mathbb{C}

How to compute $\text{Nego}(\cdot)$?

- an example - $\text{Nego}(\lambda_2)$



$\text{Nego}(\lambda_2)(b) \leq 2$?

\mathbb{P} : $b \cdot d^\omega$

\mathbb{C} : deviation $b \rightarrow a$

ω \mathbb{P} : $(a \cdot b)^* \cdot d^\omega$

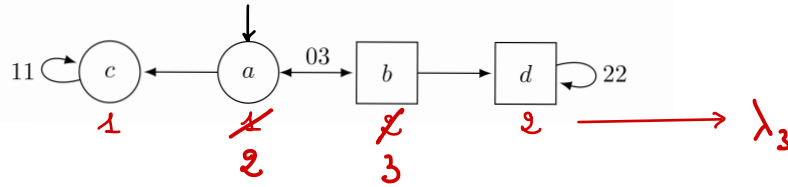
\mathbb{C} : deviation $b \rightarrow a$

→ outcome: $(ba)^\omega$ and $\text{MP}_\square((ba)^\omega) = 3$

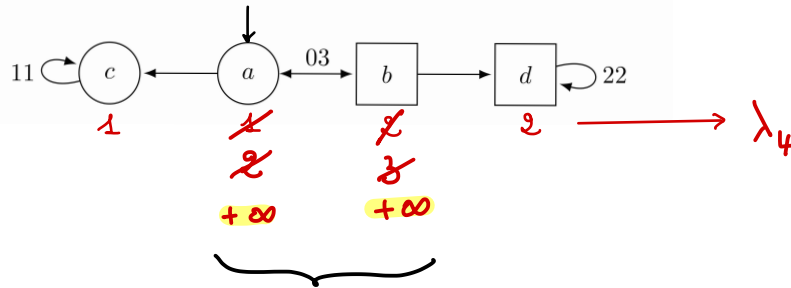
→ ~~$\text{Nego}(\lambda_2)(b) \leq 2$~~

→ generalisation: $\text{Nego}(\lambda_2)(b) = 3$

How to compute Nego(.) ?



There is **no** λ_3 -consistent path from a nor from b !
(FAIL)



no SPE starting from a or b !

Properties of the negotiation function

Let λ_0 be s.t. $\lambda_0(\sigma) = -\infty, \forall \sigma \in V$. no constraint!

Then $\text{Nego}(\lambda_0)(\sigma) = \lambda_1(\sigma) = \inf_{\sigma_{-i} \in \hat{Z}_{-i}} \sup_{\sigma_i \in \hat{Z}_i} \underline{MP}_i(\text{Out}_\sigma(\sigma_i, \sigma_{-i}))$

= $\sup_{\sigma_i \in \hat{Z}_i} \inf_{\sigma_{-i} \in \hat{Z}_{-i}} \underline{MP}_i(\text{Out}_\sigma(\sigma_i, \sigma_{-i}))$

"worst-case value"

all profiles are λ_0 -rational

Theorem. $\text{Nego}(\lambda_0)$ characterizes NEs!

Properties of the negotiation function

Let λ^* be s.t. $\text{Nego}(\lambda^*) = \lambda^*$, i.e. λ^* is a **fixed point** of Nego .

Lemma 1. $\forall \lambda^*$ -consistent paths ρ , $\exists \bar{\sigma} \in \text{SPE} : \rho = \text{Out}(\bar{\sigma})$.

Lemma 2. $\forall \bar{\sigma} \in \text{SPE} : \exists \lambda^*$ s.t. $\text{Nego}(\lambda^*) = \lambda^*$ and $\text{Out}(\bar{\sigma})$ is λ^* -consistent.

Theorem. The set of fixed points of the function Nego is a characterization of outcomes of SPEs.

Because Nego is **monotone** and the set of requirements form a **complete lattice** and in addition the set of λ -consistent paths is upward-closed then we have the following stronger result:

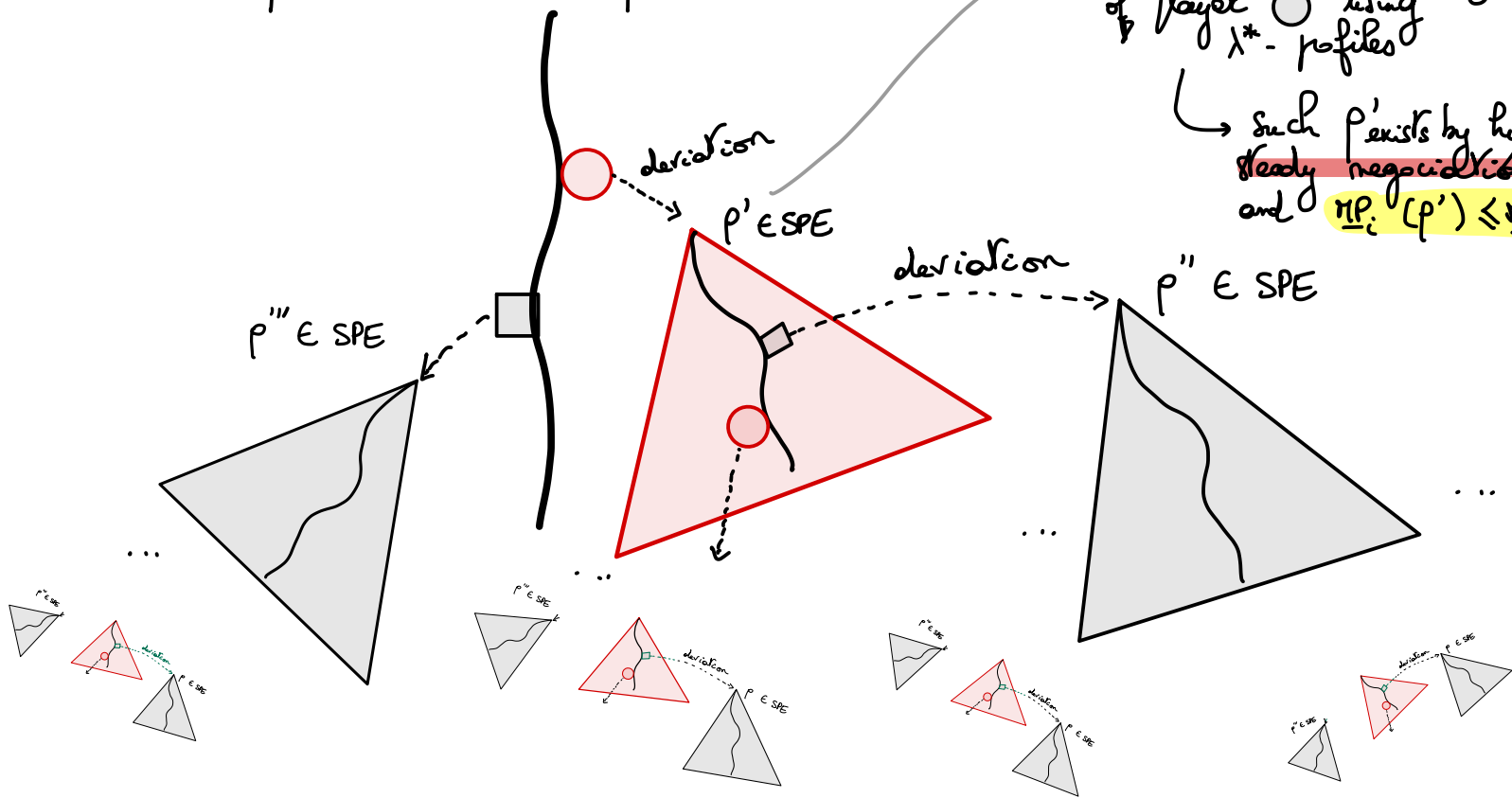
Corollary. The set of outcomes of SPEs is characterized by the **LFP** of Nego .

LFP and SPE

p is λ^* -consistent $\Leftrightarrow p \in \text{SPE}$

p' is chosen such that p' minimize the payoff of player \bigcirc using λ^* -profiles

Such p' exists by hyp. of steady negotiations and $\underline{\pi}_i(p') \leq \underline{\pi}_i(p)$



Additional properties

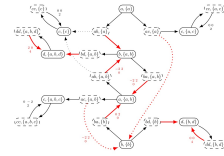
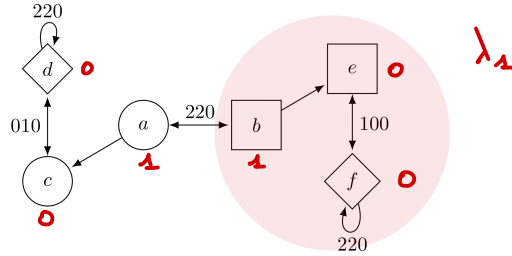


Fig. 10. A concrete negotiation game

- ① We can transform the **PC** game into a finite state **multi-mean payoff game** (ic'15)
- ② This multi-mean payoff game allows us to **effectively compute** **Nego(.)**
- ③ λ^* may not be reached from λ_0 by Kleene-Tarski iteration in finitely many steps

Non finite convergence



$\text{Nego}(\lambda_1)(a) = 1 \frac{1}{2}$

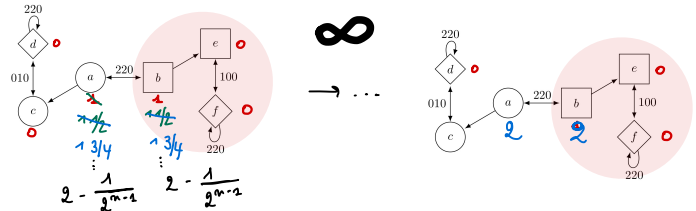
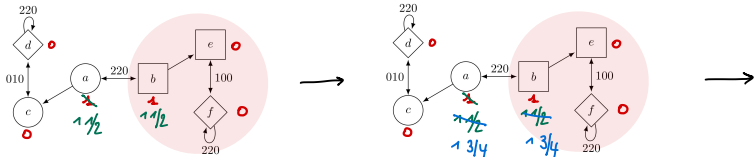
P cannot popote to go to the left with a value $\leq 1 \frac{1}{2}$

as C would deviate to "b". From "b"

P can popote on the right part $1 \frac{1}{2}$

to PP. \bigcirc (as he needs to give at least 1 to PP. \square)

Symmetrically: $\text{Nego}(\lambda_1)(b) = 1 \frac{1}{2}$



Additional properties

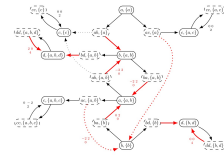
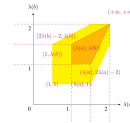


Fig. 10. A concrete negotiation game

- ① We can transform the **PC** game into a finite state **multi-mean payoff game** (ic'15)
- ② This multi-mean payoff game allows us to **effectively compute** **Nego(.)**
- ③ λ^* may not be reached from λ_0 by Kleene-Tarski iteration in finitely many steps
- ④ ... **BUT** thanks to good properties of multi-mean payoff games, we can show that **Nego(.)** is **effectively piecewise linear** and λ^* can be obtained using linear algebraic techniques.

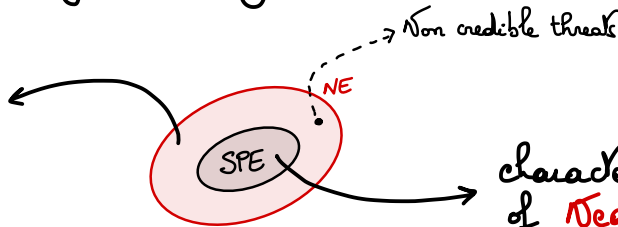


Conclusions and perspectives

→ SPE provides a natural notion of rational behaviors in infinite duration games played on graphs

→ Worst-case value relative to rational adversary formalized by the fixed points of $\text{Wego}(\cdot)$ leads to an effective representation of $\text{OutSPE}(G)$ for IP games (multi mean-payoff automata)

characterized
by worst-case value



characterized by fixed-points
of $\text{Wego}(\cdot)$

Conclusions and perspectives

→ $\text{Nego}(\cdot)$ is also applicable to parity games (omega regular obj.)

↳ useful to close complexity gaps

ex: Constrained existence for SPE
is in ExpTime (emptiness of alternating automata)
and NP-hard. [Ummer's '06]

→ Our previous algorithm for quantitative reachability can be reframed with $\text{Nego}(\cdot)$ [CONCUR'19]

→ $\text{Nego}(\cdot)$ provides a new algorithmic basis to do rational verification and synthesis based on SPEs.