Subgame perfect equilibrium in mean-payoff games

Jean-François Raskin (Université libre de Bruxelles)

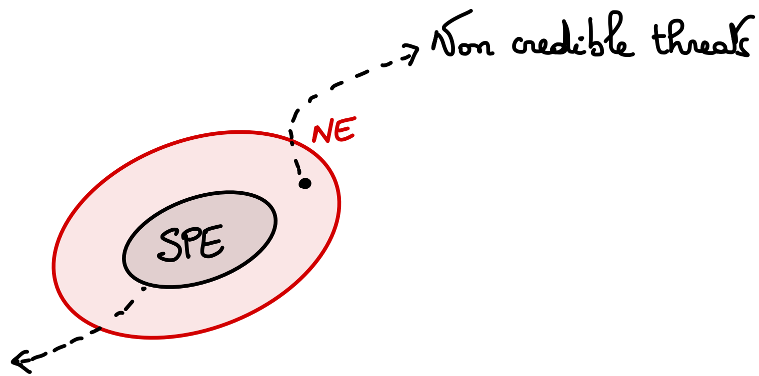
joint work with Marie van den Bogaard (Université Gustave Eiffel)

Léonard Brice (ENS Paris-Saclay)

Objectives of the talk

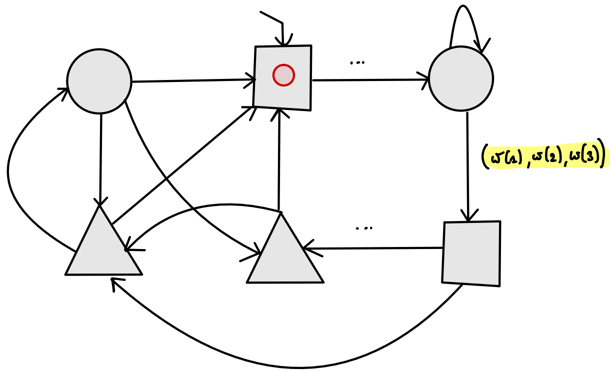
1. Recall SPE for sequential games

2. Expose 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 $i \in [1, N]$

$E \subseteq V \times V$, $w : E \to \mathbb{Z}^N$

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

↳ mean-payoff $\mu_i(\rho)$

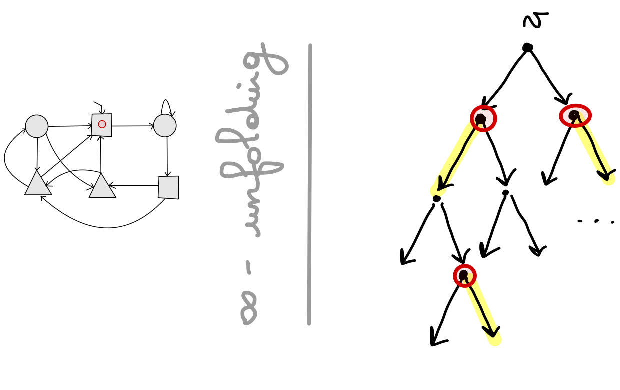
$$MP_i(\rho) = \liminf_{n \to \infty} \frac{1}{n} \cdot \sum_{j=0}^{n-1} w(v_j, v_{j+1})(i)$$

Rationality : each player wants to maximize his own payoff.

Strategies, profiles, outcomes

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

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

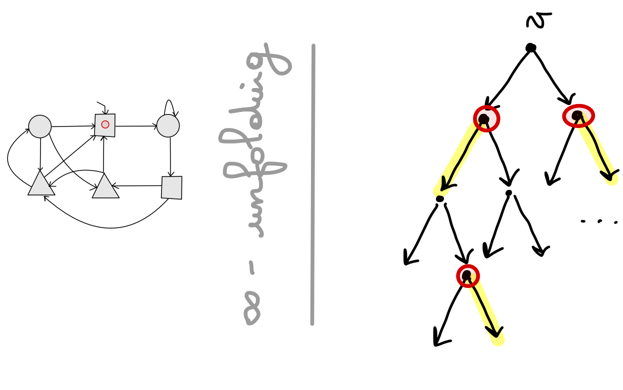


Σ_i = set of strategy of Pl. $i \in [1, N]$

Strategies, profiles, outcomes

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

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



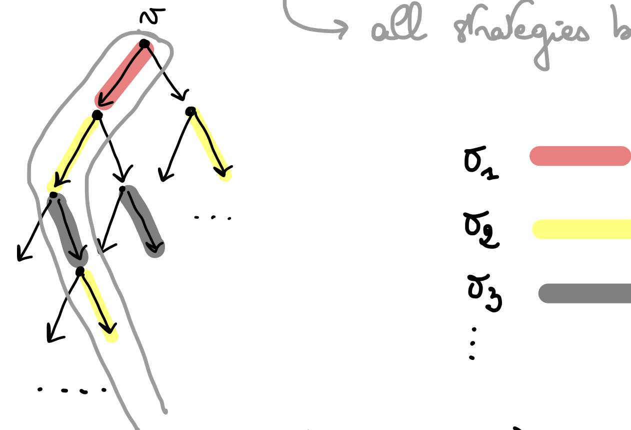
Σ_i = set of strategy of Pl. $i \in [1, N]$

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



$= Out_v(\sigma_1, \sigma_2, \dots, \sigma_N) = \rho \in V^\omega$

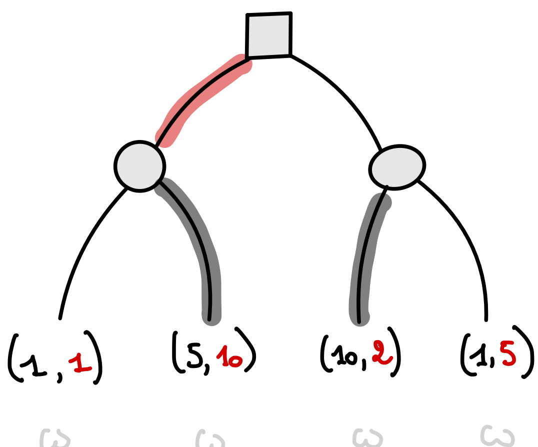
$Out_{v_0}(\bar{\sigma}) = v_0 v_1 v_2 \dots v_n \dots = \rho$

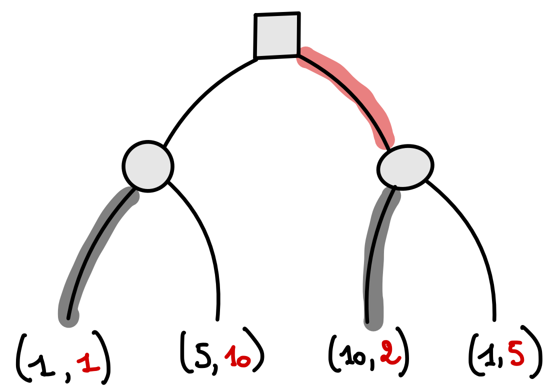
s.t. $v_0 = v$, $\forall j \geq 0$: if $\rho(j) \in V_i$: $v_{j+1} = \sigma_i(\rho(0..j))$

Nash equilibrium and subgame perfect equilibrium

A profile of strategies $(\sigma_1, \sigma_2, \dots, \sigma_N)$ is a Nash equilibrium (NE) in v_0, if for all $i \in [1, N]$, for all $\sigma_i' \in \Sigma_i$: $\mu_i(\mathrm{Out}_{v_0}(\bar{\sigma}_{-i}, \sigma_i')) \leq \mu_i(\mathrm{Out}_{v_0}(\bar{\sigma}_{-i}, \sigma_i))$

= No player has an incentive to deviate unilaterally.

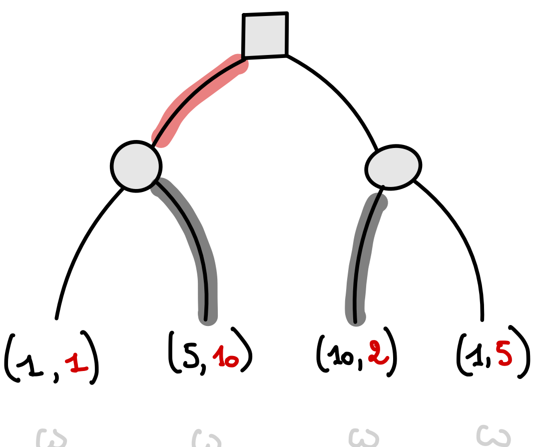


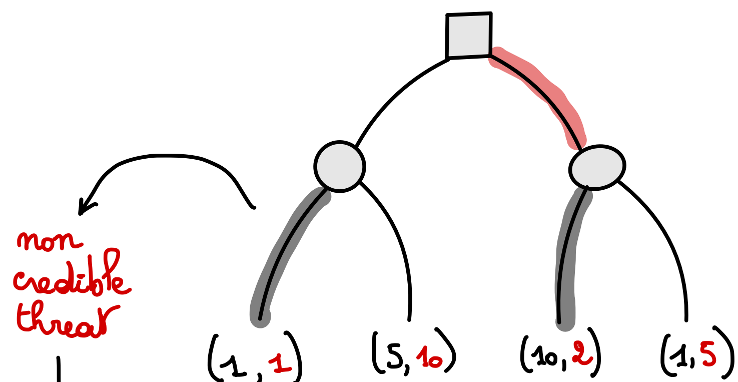


Nash equilibrium and subgame perfect equilibrium

A profile of strategies $(\sigma_1, \sigma_2, \dots, \sigma_N)$ is a Nash equilibrium (NE) in v_0, if for all $i \in [1, N]$, for all $\sigma_i' \in \Sigma_i$: $\mu_i(\mathrm{Out}_{v_0}(\bar{\sigma}_{-i}, \sigma_i')) \leq \mu_i(\mathrm{Out}_{v_0}(\bar{\sigma}_{-i}, \sigma_i))$

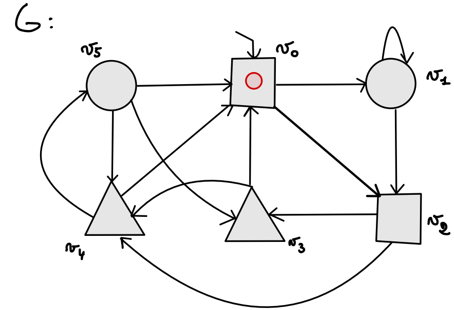
= No player has an incentive to deviate unilaterally.



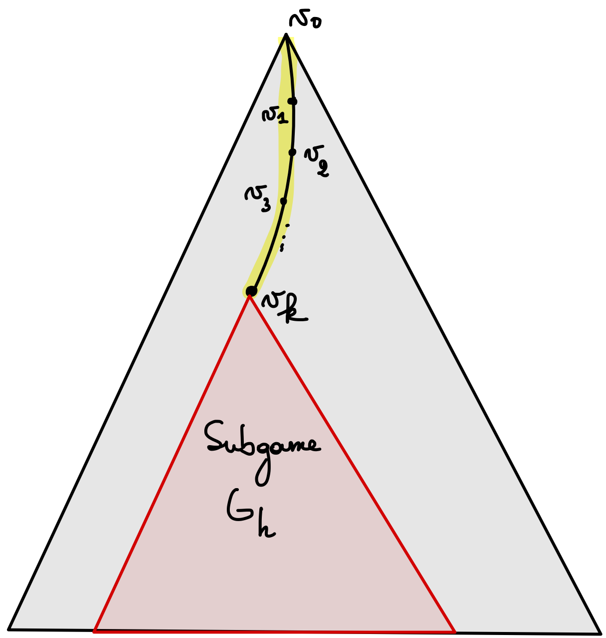


Pl. ◯ is not rational in this subgame!

Subgame



Unfolding



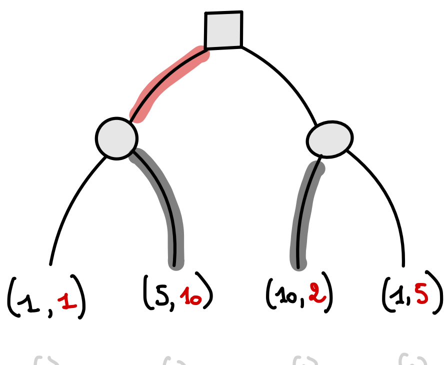
Every history $h = v_0 v_1 \dots v_k$ defines a <u>subgame</u> G_h

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.P. $i \in [1, N]$, for all $\sigma_i^{h\prime} \in \Sigma_i$:

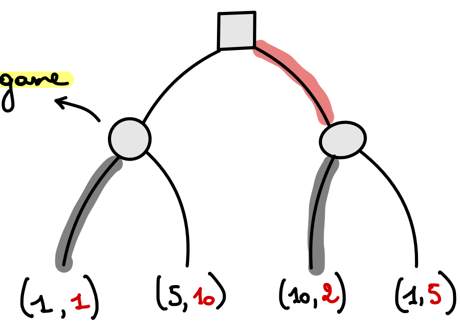
$$\mu_i\left(\mathrm{Out}_h\left(\bar{\sigma}_{-i}^h, \sigma_i^{h\prime}\right)\right) \leq \mu_i\left(\mathrm{Out}_h\left(\bar{\sigma}_{-i}^h, \sigma_i^h\right)\right).$$

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



SPE

in this subgame P.P. ○ has an incentive to deviate
↓
non credible threat



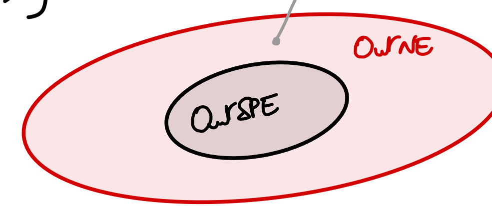
~~SPE~~

Outcomes supported by equilibria

$$\mathrm{OutNE}(G) = \bigcup_{\bar{\sigma} \in NE} \{\mathrm{Outcome}_{v_0}(\bar{\sigma})\}$$

$$\mathrm{OutSPE}(G) = \bigcup_{\bar{\sigma} \in SPE} \{\mathrm{Outcome}_{v_0}(\bar{\sigma})\}$$

only supported by non-credible threats



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

Effective representation of the set of outcomes of NE/SPE

Why ?

1. Existence problem for SPE : $\mathrm{Out.SPE}(G) \overset{?}{\neq} \emptyset$

→ left-open in the literature for MP games

Effective representation of the set of outcomes of NE/SPE

why ?

① Existence problem for SPE : $\mathrm{Out.SPE}(G) \overset{?}{=} \emptyset$

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

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

all: $\mathrm{Out.NE}(G) \overset{?}{\subseteq} [\![\psi]\!]$ $\qquad$ $\mathrm{Out.SPE}(G) \overset{?}{\subseteq} [\![\psi]\!]$

Some: $\mathrm{Out.NE}(G) \cap [\![\psi]\!] \overset{?}{\neq} \emptyset$ $\qquad$ $\mathrm{Out.SPE}(G) \cap [\![\psi]\!] \overset{?}{\neq} \emptyset$

→ idem for $M_{\sigma_0} \otimes G_{0 \cup [1,N]}$

↳ System , $[1,N]$ = Env.

Effective representation of the set of outcomes of NE/SPE

why ?

① Existence problem for SPE : $Out_{SPE}(G) \overset{?}{=} \emptyset$

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

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

all: $Out_{NE}(G) \overset{?}{\subseteq} [\![\psi]\!]$ $\quad$ $Out_{SPE}(G) \overset{?}{\subseteq} [\![\psi]\!]$

Some: $Out_{NE}(G) \cap [\![\psi]\!] \overset{?}{\neq} \emptyset$ $\quad$ $Out_{SPE}(G) \cap [\![\psi]\!] \overset{?}{\neq} \emptyset$

→ idem for $M_{\sigma_0} \otimes G_{0 \cup [1,N]}$

↳ System, $[1,N]$ = Env.

③ Cooperative rational synthesis : $\exists \rho \in Out_{\left\{\substack{SPE \\ NE}\right\}}(G) : \rho \models \varphi_0$

[Kupferman et al. 16]

↳ constrained-$\exists$ $\quad$ ↳ Spec. syst.

Starting point:
NE in infinite duration games

Set of outcomes supported by NE - MP

→ Requirement : $\lambda : V \to \mathbb{R} \cup \{-\infty, +\infty\}$ % lower bound on reward

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

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

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

→ $\forall v \in V_i$: let $\lambda_1(v) = \sup_{\sigma_i \in \Sigma_i} \inf_{\sigma_{-i} \in \Sigma_{-i}} \underline{MP}_i(Out_v(\sigma_i, \sigma_{-i}))$ % requirement induced by WCV

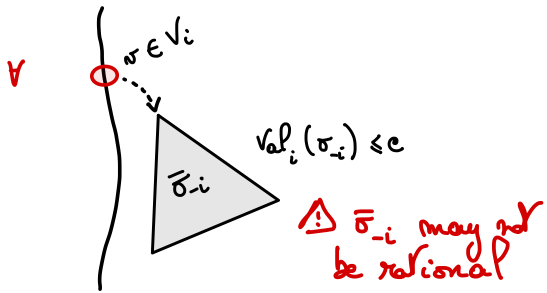
Theorem: $\rho = v_0 v_1 \dots v_n \dots$, $\rho \in OutNE(G)$ iff ρ is λ_1-consistent.

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

NE - Deviation - Punishment

→ using $Val_i(s)$ to characterize all NE
↳ worst-case value for Pl. i

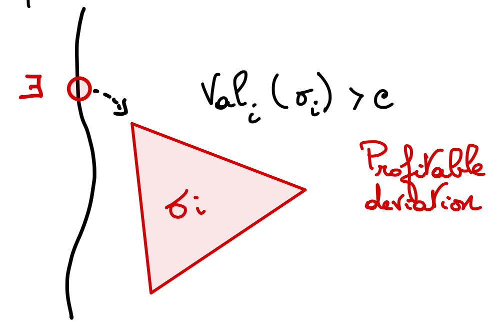
$\rho \in NE$



$Val_i(\rho) = c$ and

$c \geq \inf_{\bar{\sigma}_{-i}} \sup_{\sigma_i} Val(Out_v(\sigma_i, \bar{\sigma}_{-i}))$ .... = ... worst-case value for Pl_i

$\rho \notin NE$



$Val_i(\rho) = c$ and

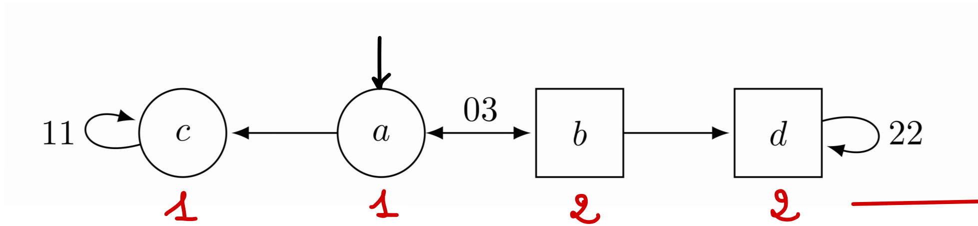
$\sup_{\sigma_i} \inf_{\bar{\sigma}_{-i}} Val(Out_v(\sigma_i, \bar{\sigma}_{-i})) > c$

→ This works for any prefix independent payoff function for which worst-case value exists and can be realized → [Brihaye et al. 13]

→ OK for MP

Set of outcomes supported by NE - MP

- an example



$\to \lambda_1 \to$ worst-case values

The set of λ_1-consistent paths in G are :

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

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

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

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)?

→ Clearly, a rational player will never give away his/her worst-case value !

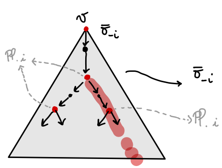
The negotiation function

? : 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)?

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

$\inf \emptyset = +\infty$

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

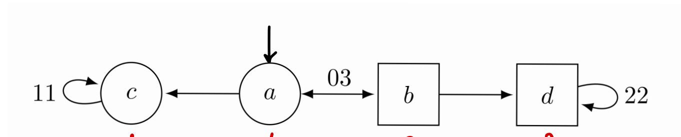


$\bar{\sigma}_{-i}$ is $\lambda\mathrm{Rat}(v)$: in all $\bar{\sigma}_{-i}$ consistent history h, there is a $\bar{\sigma}_{-i}$ continuation that is λ-consistent.

≈ worst case value against λ-rational players.

How to compute Nego(.) ?

– an example – $Nego(\lambda_1)$



1 1 2 2 → λ_1

$Nego(\lambda_1)(a) \leq 1$?

$\mathbb{P}$ want to minimize reward of $\mathbb{P}$.i

$\mathbb{C}$ want to max. reward of $\mathbb{P}$.i

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

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

$\mathbb{P}$: from b, the only λ_1-consistent paths are $(b\,a)^* \cdot d^\omega$

even if $(a\,b)^\omega$ is tempting

it does not give 1 to ○

Prover Challenger game

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

~~$Nego(\lambda_1)(a) \leq 1$~~

→ generalisation : $Nego(\lambda_1)(a) = 2$

+ Iterate

Properties of the negociation function

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

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

Lemma 2. $\forall \bar{\sigma} \in$ SPE : $\exists\, \lambda^*$ s.t. Nego$(\lambda^*) = \lambda^*$ and 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.

Additional properties

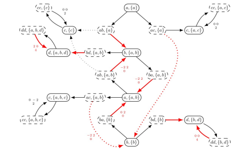


Fig. 10. A concrete negotiation game

① We can transform the P C 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

Additional properties

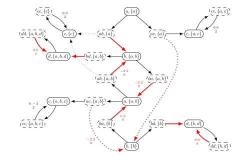


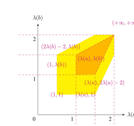
Fig. 10. A concrete negotiation game

① We can transform the $\boxed{P}\,\boxed{C}$ game into a finite state multi- mean payoff game (ic'15)

② This multi- mean payoff game allows us to effectively compute $\boxed{\text{Nego(.)}}$

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 $\boxed{\text{Nego(.)}}$ is effectively piecewise linear and λ^* can be obtained using linear algebraic techniques.



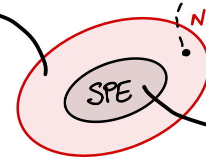
Conclusions and perspectives

→ 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 Nego(·) leads to an effective representation of OutSPE(G) for MP games (multi mean-payoff automata)

Non credible threats

characterized by worst-case value



characterized by fixed-points of Nego(·)

Conclusions and perspectives

→ Nego(·) 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. [Ummels '06] → We have recently proved NP-completeness

→ Our previous algorithm for quantitative reachability can be rephrased with Nego(·) [CONCUR '19]

→ Nego(·) provides a new algorithmic basis to do rational verification and synthesis based on SPEs.