Subgame Perfection

with an Algorithmic Perspective

Jean-François Raskin (Université libre de Bruxelles)
based on joint works with
Marie van den Bogaard (Université Gustave Eiffel)
and Léonard Brice (Université libre de Bruxelles)

Simons Institute - July - 2024

Objectives of the talk

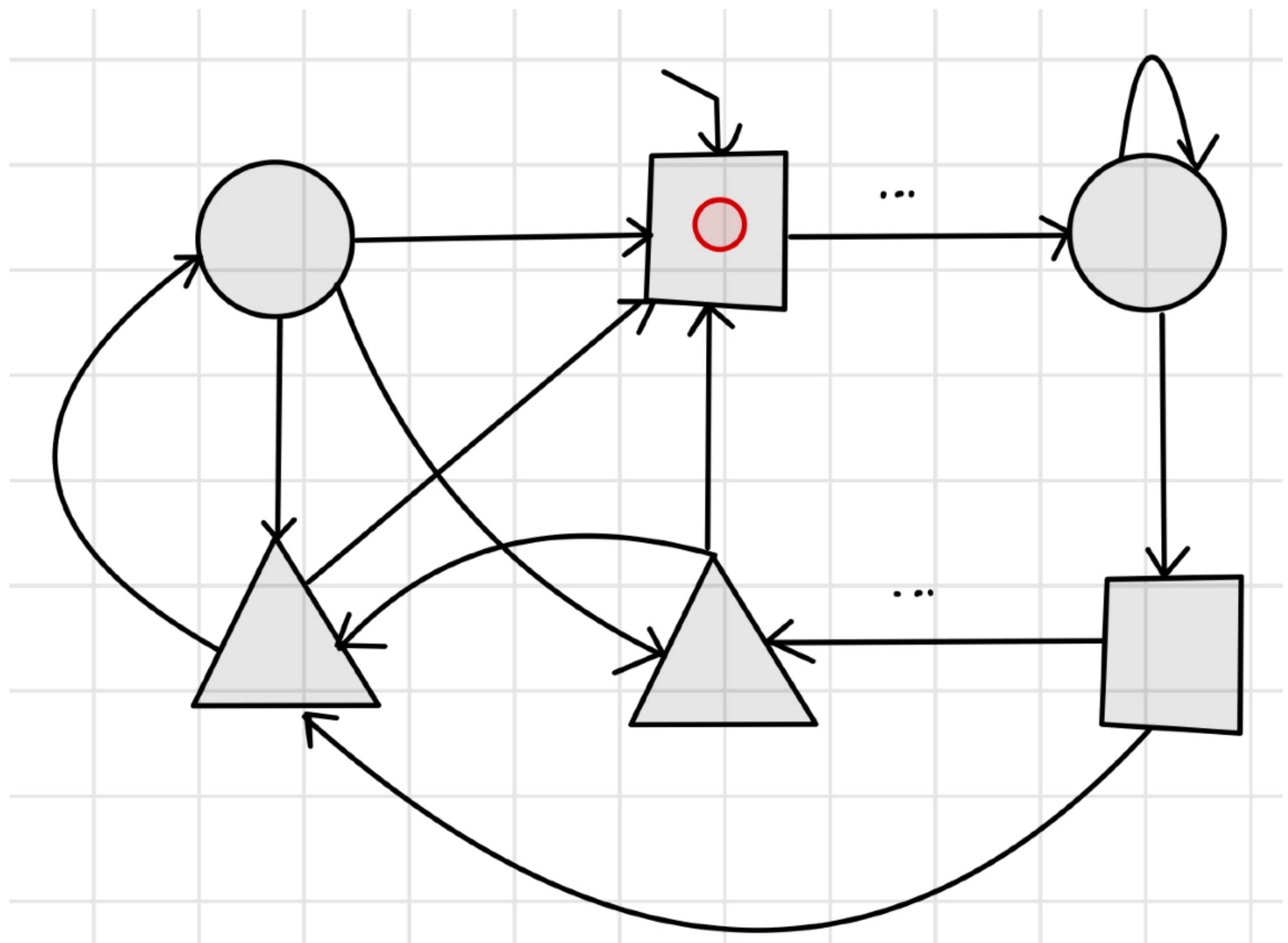
- **Subgame Perfect Equilibrium** to model **rationality** in sequential games (instead of Nash equilibrium)

- Expose new **algorithmic ideas** for SPE for N-player graph games with:
  - **Parity** objectives
  - **Mean-payoff** objectives

Setting

N player turn-based graph games
Game setting

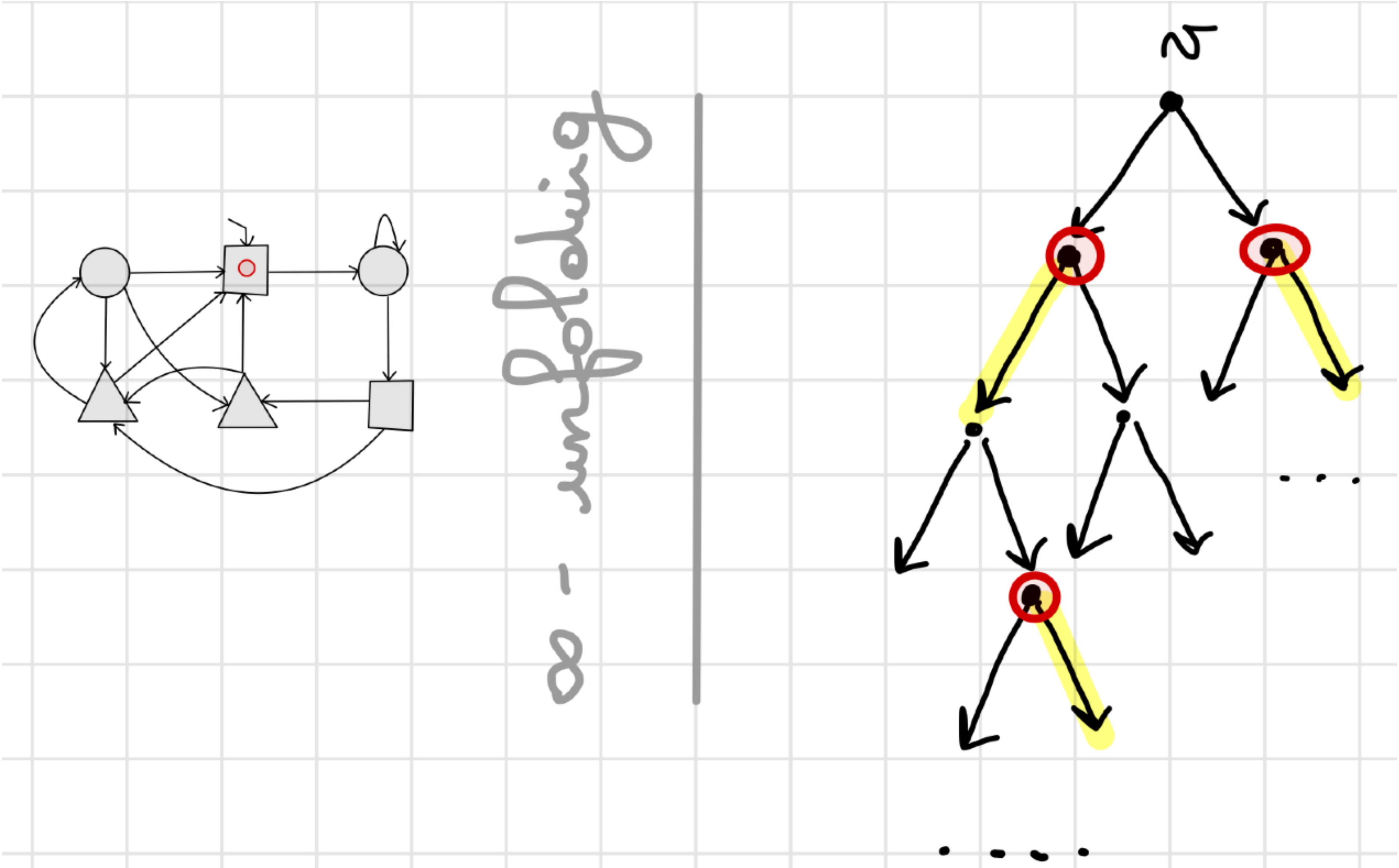
- Set of vertices **partitioned** according to players
- Players move a token. A **play** ρ is an **infinite path** in the graph (travel of the token)
- States annotated with vectors of colors ($\mathbb{N}$ for **parity**) or rewards ($\mathbb{Q}$ for **mean-payoff**), one dimension per player
- Each play ρ gives a **payoff** μ_i to each player:
  - Parity: $\mu_i(\rho) = \min\{\text{color}_i(v) \mid v \in \text{inf}(\rho)\}$ is even
  - Mean-payoff: $\mu_i(\rho) = \liminf_{j \to +\infty} \frac{\text{SumReward}_i(\rho(0..j))}{j}$
- **Rationality**: players want to maximize their own payoff



How do players play ?
Strategies, profiles, outcomes

- Players play strategies:

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

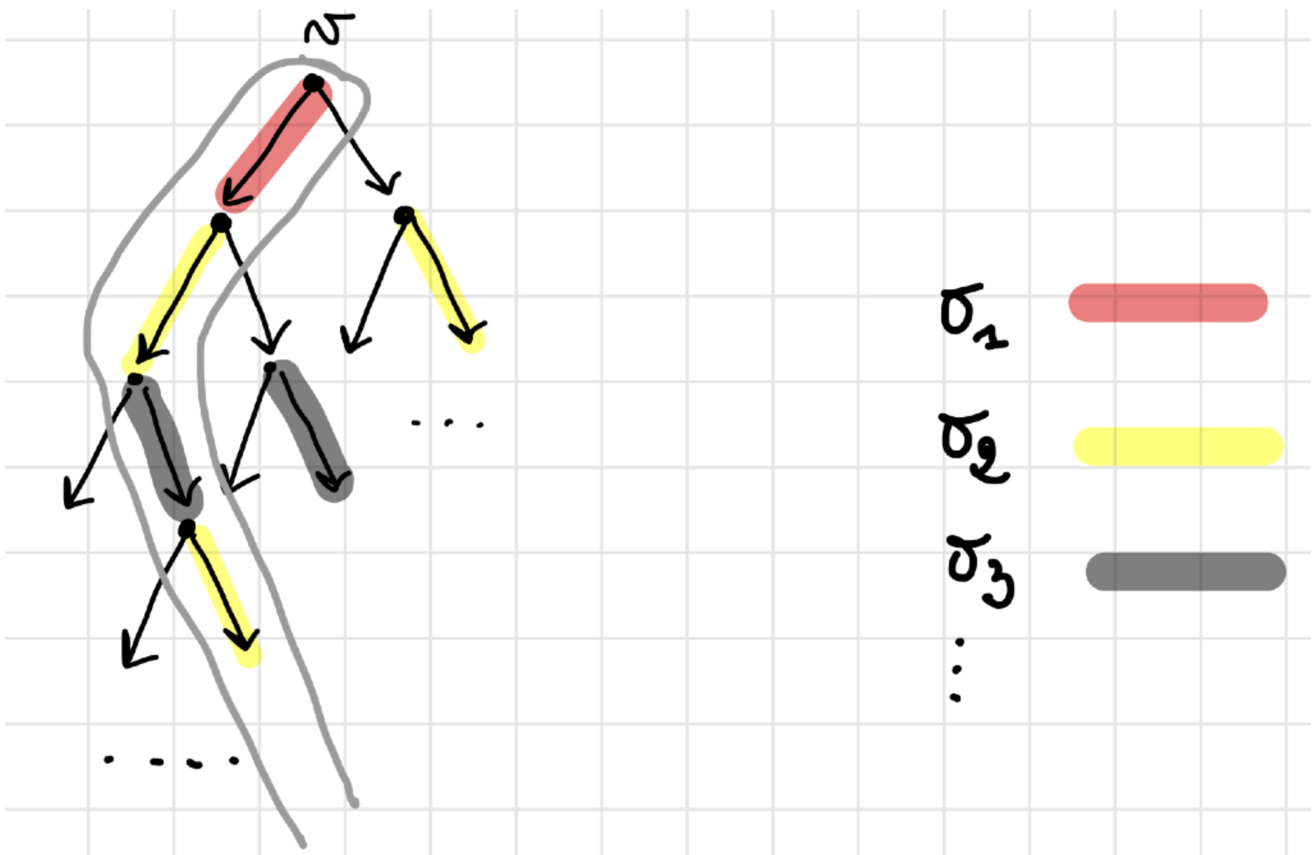


Σ_i = set of strategies of Player i

- Profiles of strategies:

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

Notation: (σ_i, σ_{-i})



Outcome$_v(\sigma_1, \sigma_2, \ldots, \sigma_n) = v_0 v_1 \ldots v_n \ldots = \rho$ such that $v = v_0 \wedge \forall j \geq 0 : v_j \in V_i \to v_{j+1} = \sigma_i(\rho(0 \ldots j))$.

Why to model rational agents/players ?

Assume turned based arena modeling a protocol to be used by **rational agents,** each having their **own** objectives.

Relevant questions:

- if agents resolve nondeterminism left in the protocol **rationally**, is it the case that some **good property emerges** ? do all rational executions satisfy ψ ?
- is there a **rational** behavior of the participants in which all participants gain at least c ?
  (if so, we could ask them to settle for this profile of behaviors)
- Is there **at least one** rational execution of the protocol ?
  Are all the possible executions of the protocol rational ?
- etc.

How to model rational agents/players ?

Different **solution concepts** used to predict how a game will be played:

- optimality (1-player/agent, e.g. shortest path)
- Pareto optimality (1-player/agent with several objectives)
- **NE, Admissible strategies, Dominent strategies, SPE** (when several agents are involved)
- ...

Rationality

When are players playing rationally?

Nash equilibrium

- A profile of strategies $(\sigma_1, \sigma_2, \ldots, \sigma_N)$ is a Nash Equilibrium (NE) in v_0 if

  $$\forall i \in [1,N] \cdot \forall \sigma'_i \in \Sigma_i : \mu_i(\text{Outcome}_{v_0}(\sigma_1, \ldots, \sigma'_i, \ldots, \sigma_N)) \leq \mu_i(\text{Outcome}_{v_0}(\sigma_1, \ldots, \sigma_i, \ldots, \sigma_N))$$

  i.e. no player has an incentive to deviate unitarily.

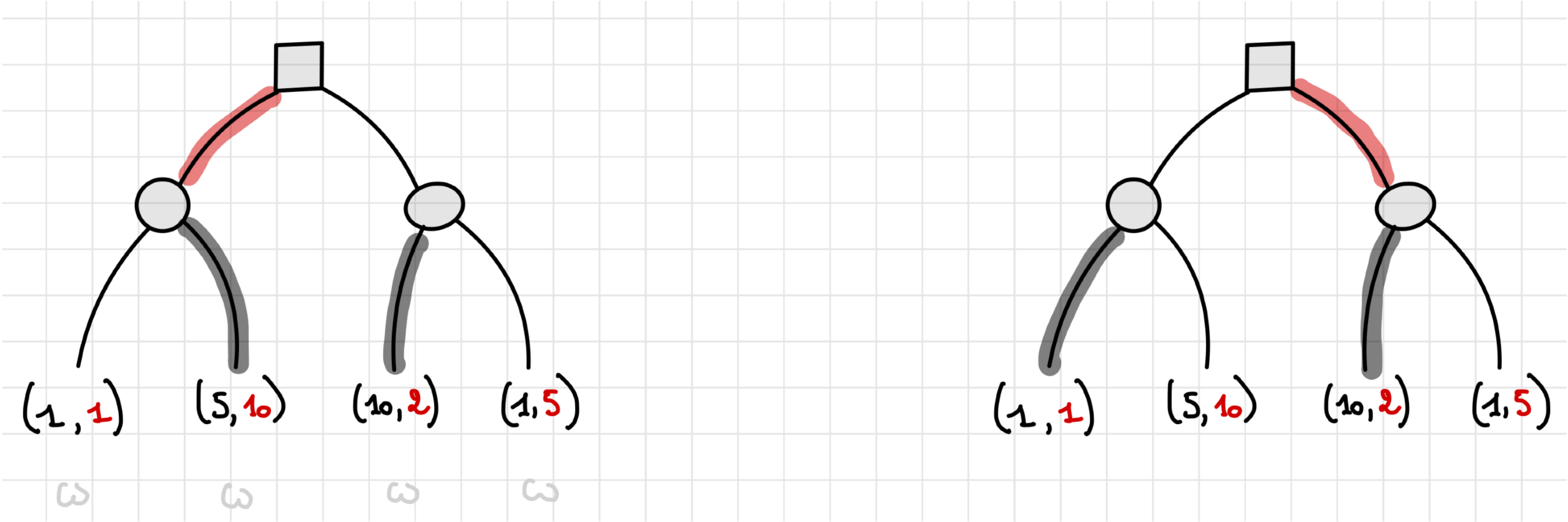
When are players playing rationally?

Nash equilibrium

- A profile of strategies $(\sigma_1, \sigma_2, \ldots, \sigma_N)$ is a Nash Equilibrium (NE) in v_0 if

$$\forall i \in [1,N] \cdot \forall \sigma_i' \in \Sigma_i : \mu_i(\text{Outcome}_{v_0}(\sigma_1, \ldots, \sigma_i', \ldots, \sigma_N)) \leq \mu_i(\text{Outcome}_{v_0}(\sigma_1, \ldots, \sigma_i, \ldots, \sigma_N))$$

i.e. no player has an incentive to deviate unitarily.



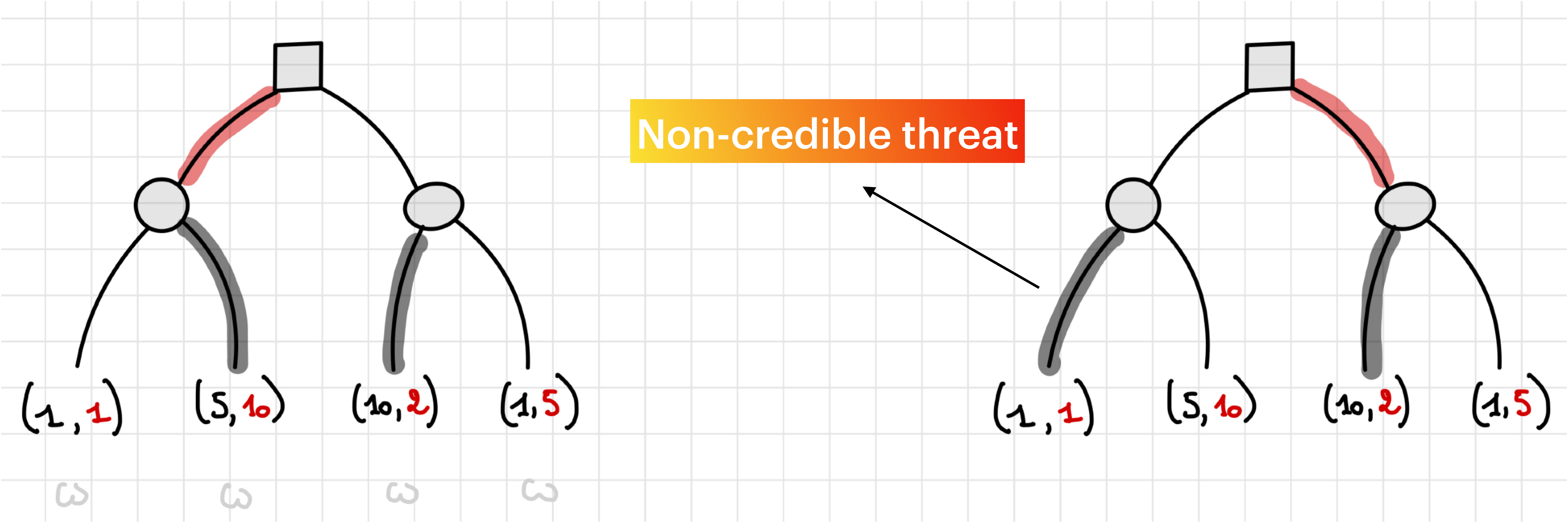
When are players playing rationally?

Nash equilibrium

- A profile of strategies $(\sigma_1, \sigma_2, \ldots, \sigma_N)$ is a Nash Equilibrium (NE) in v_0 if

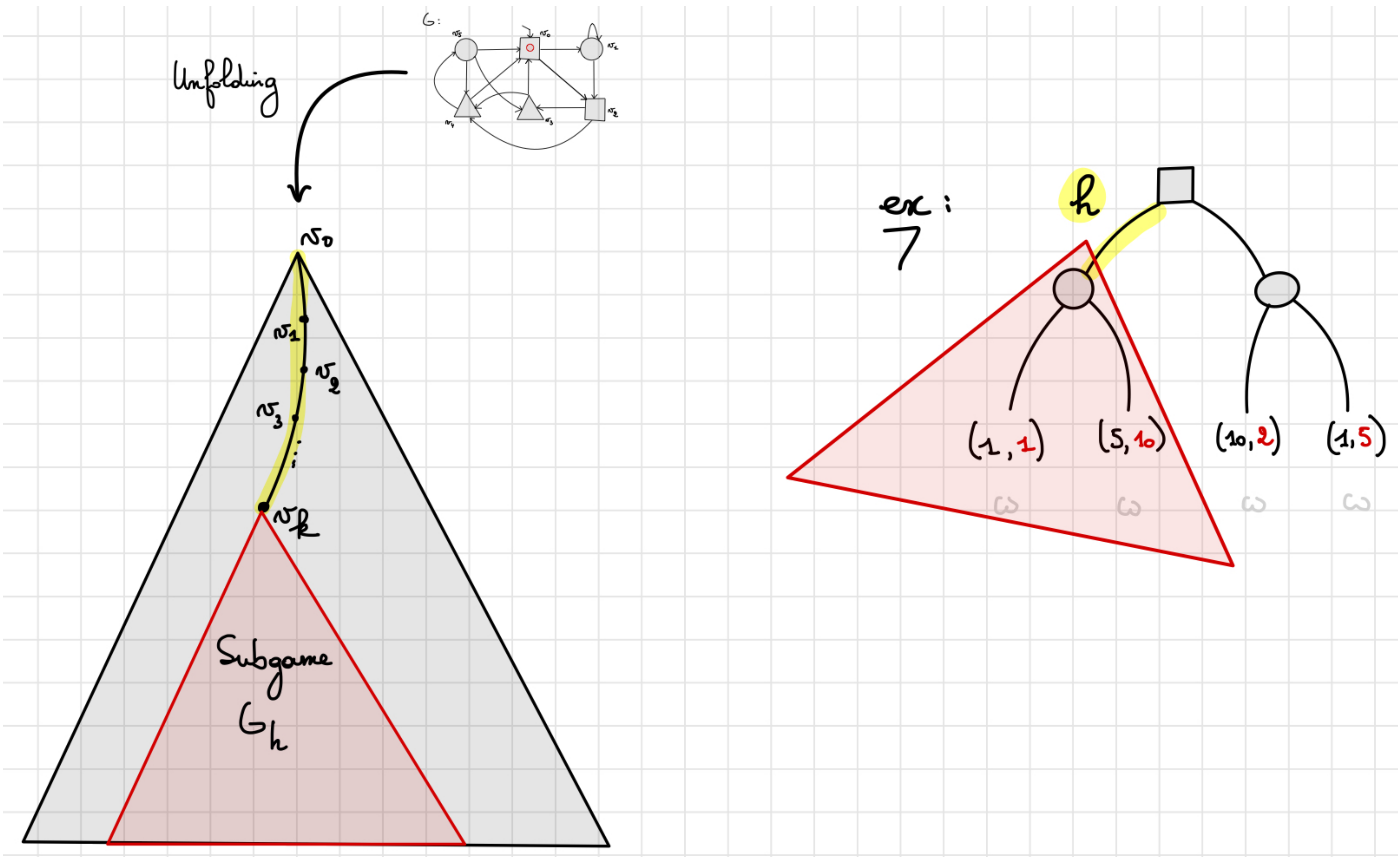
$$\forall i \in [1,N] \cdot \forall \sigma_i' \in \Sigma_i : \mu_i(\text{Outcome}_{v_0}(\sigma_1, \ldots, \sigma_i', \ldots, \sigma_N)) \leq \mu_i(\text{Outcome}_{v_0}(\sigma_1, \ldots, \sigma_i, \ldots, \sigma_N))$$

i.e. no player has an incentive to deviate unitarily.



When are players playing rationally?

Avoid non-credible threats: Subgame perfect equilibrium



Subgame G_h = game induced by history h

Players must be rational in all subgames !

When are players playing rationally?

Subgame perfect equilibrium

- A profile of strategies $(\sigma_1, \sigma_2, \ldots, \sigma_n)$ is a Subgame Perfect Equilibrium (SPE) in v_0 if

  $\forall i \in [1,N] \cdot$ $\forall$ histories h $\cdot \forall \sigma_i' \in \Sigma_i$:
  $\mu_i(\text{Outcome}_h(\sigma_1, \ldots, \sigma_i', \ldots, \sigma_N)) \leq \mu_i(\text{Outcome}_h(\sigma_1, \ldots, \sigma_i, \ldots, \sigma_N))$

  i.e. no player has an incentive to deviate unitarily in any subgame.
  Players are rational in all subgames (**no** non-credible threats.)

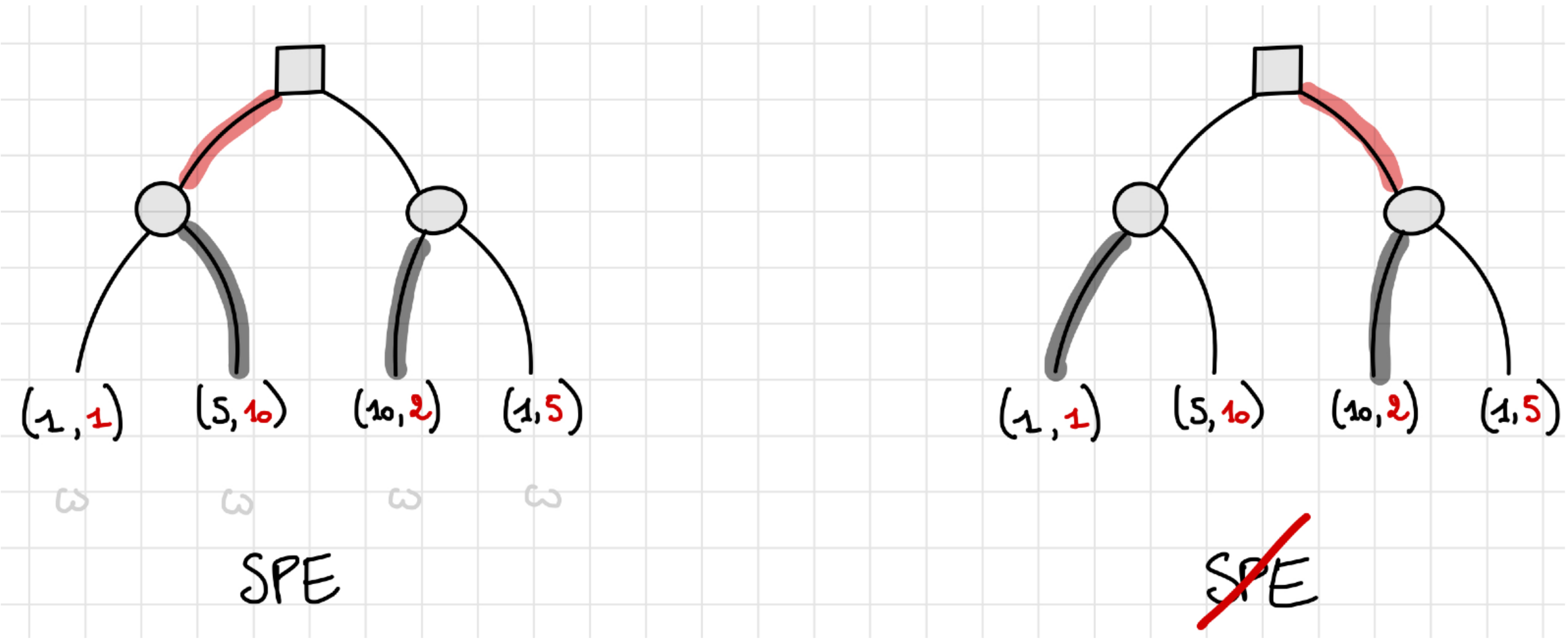
When are players playing rationally?

Subgame perfect equilibrium

- A profile of strategies $(\sigma_1, \sigma_2, \ldots, \sigma_n)$ is a Subgame Perfect Equilibrium (SPE) in v_0 if

  $\forall i \in [1,N] \cdot \forall$ histories $h \cdot \forall \sigma_i' \in \Sigma_i$:
  $\mu_i(\text{Outcome}_h(\sigma_1, \ldots, \sigma_i', \ldots, \sigma_N)) \leq \mu_i(\text{Outcome}_h(\sigma_1, \ldots, \sigma_i, \ldots, \sigma_N))$

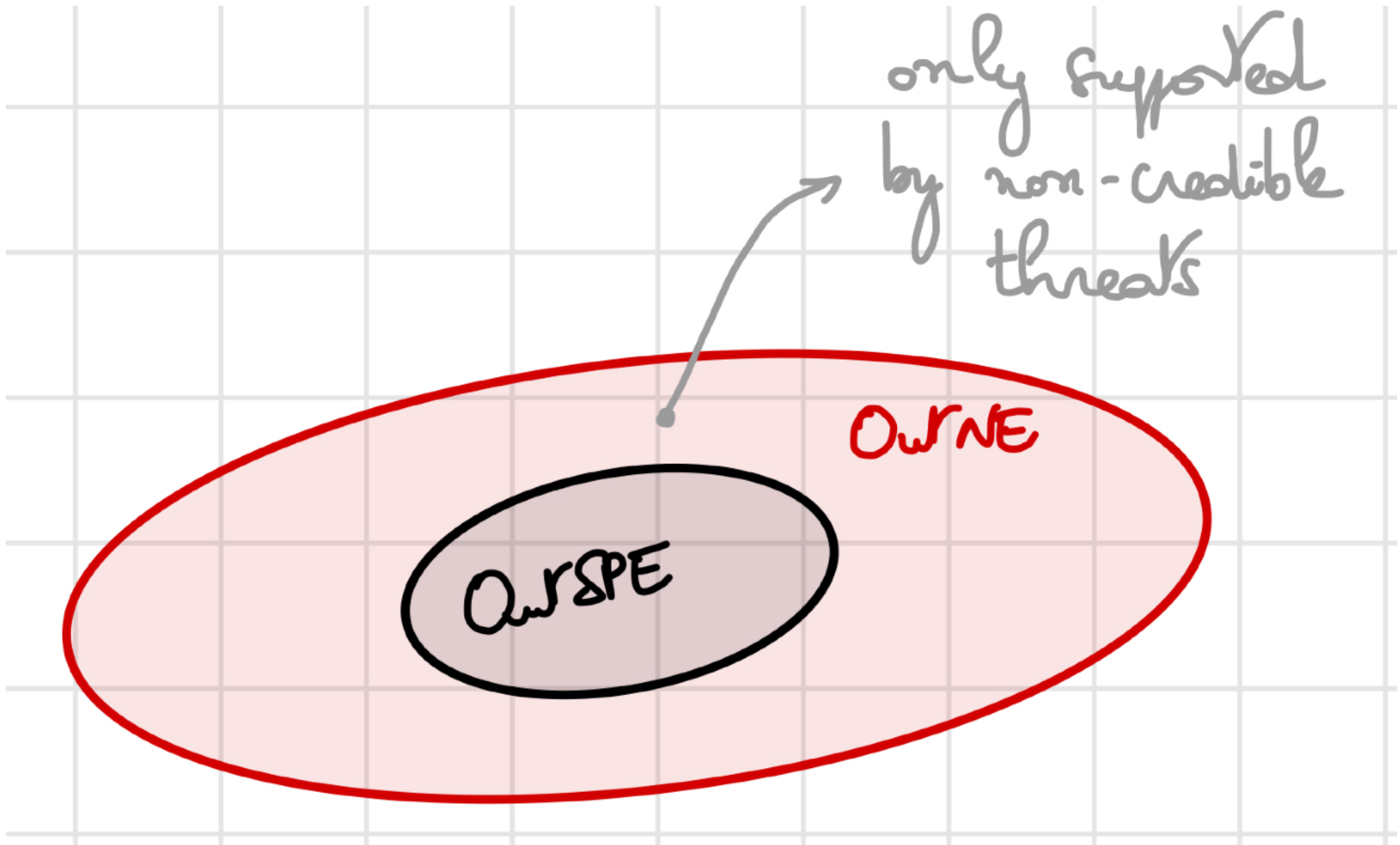
  i.e. in all subgames, we have NEs (**no** non-credible threats.)



Outcomes supported by equilibria

NE - SPE

- $\text{OutNE}(G) = \bigcup_{\bar{\sigma} \in NE} \text{Outcome}_{v_0}(\bar{\sigma})$
- $\text{OutSPE}(G) = \bigcup_{\bar{\sigma} \in SPE} \text{Outcome}_{v_0}(\bar{\sigma})$



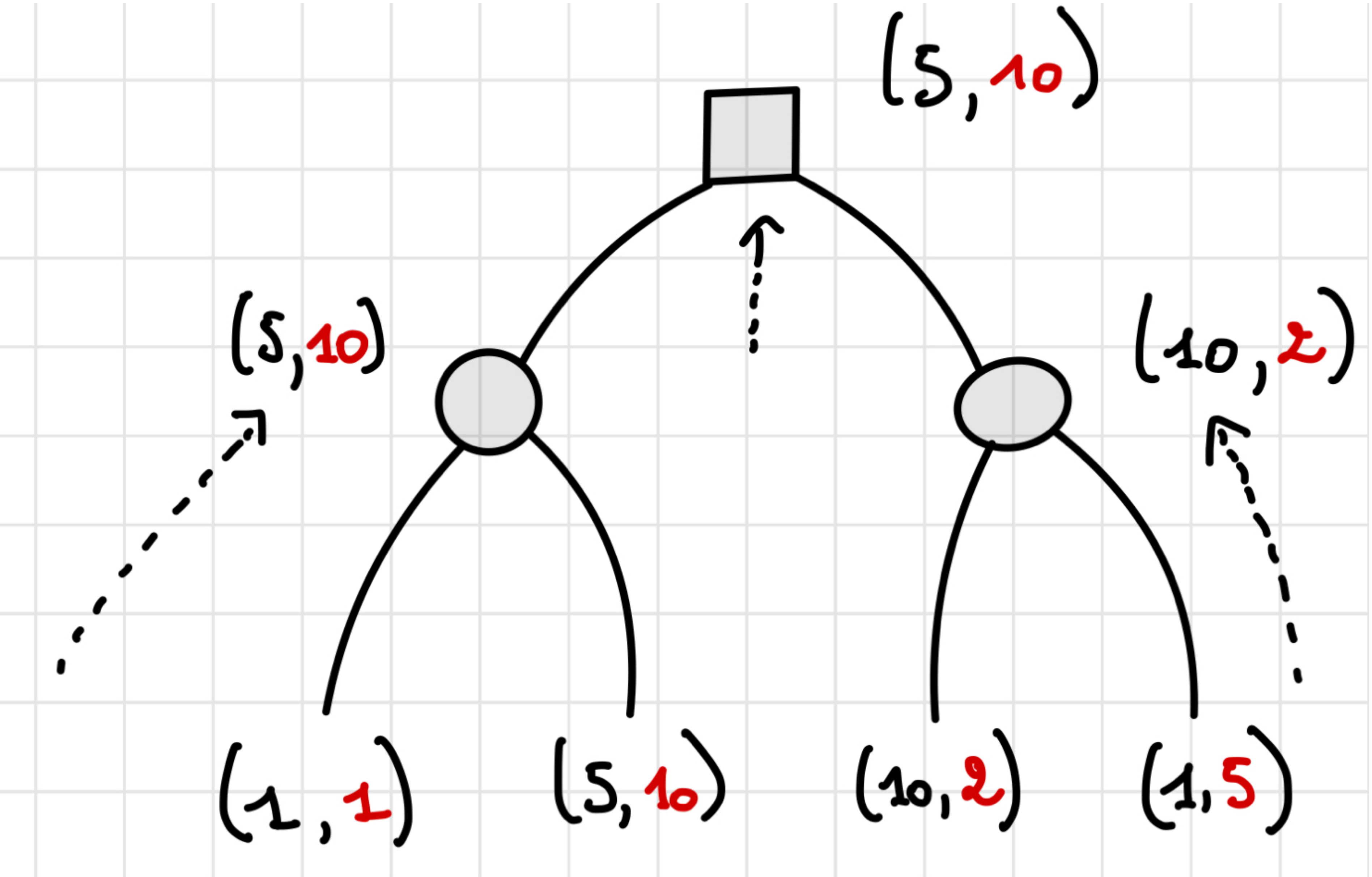
- **How to compute effective representations for those sets ?**
- **Why ?**
  - Existence problem: $\text{OutSPE}(G) \stackrel{?}{=} \varnothing$
    (while they always exists for parity games, it is not the case for MP games)
  - Rational verification:
    $(\exists)$ $\exists \rho \in \text{OutSPE}(G) : \rho \vDash \psi$?
    $(\forall)$ $\forall \rho \in \text{OutSPE}(G) : \rho \vDash \psi$?
  - Cooperative rational synthesis [Kuperfman et al.]: $\exists \rho \in \text{OutSPE(G)} : \rho \vDash p_0$? (parity obj. of Player 0 is true)

Algorithms

How to reason algorithmically on SPE?

Easy case: finite trees

- For finite trees: backward induction

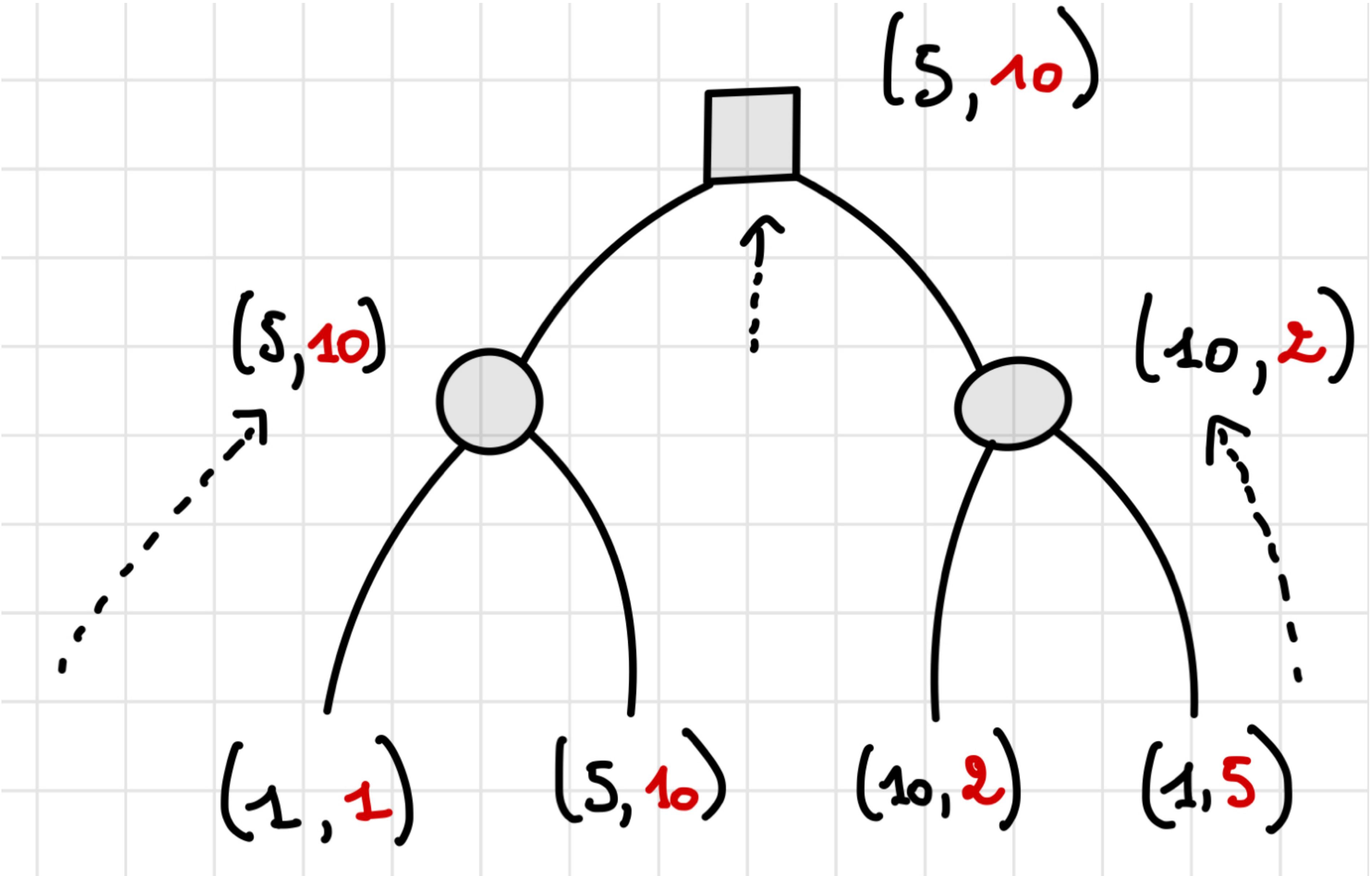


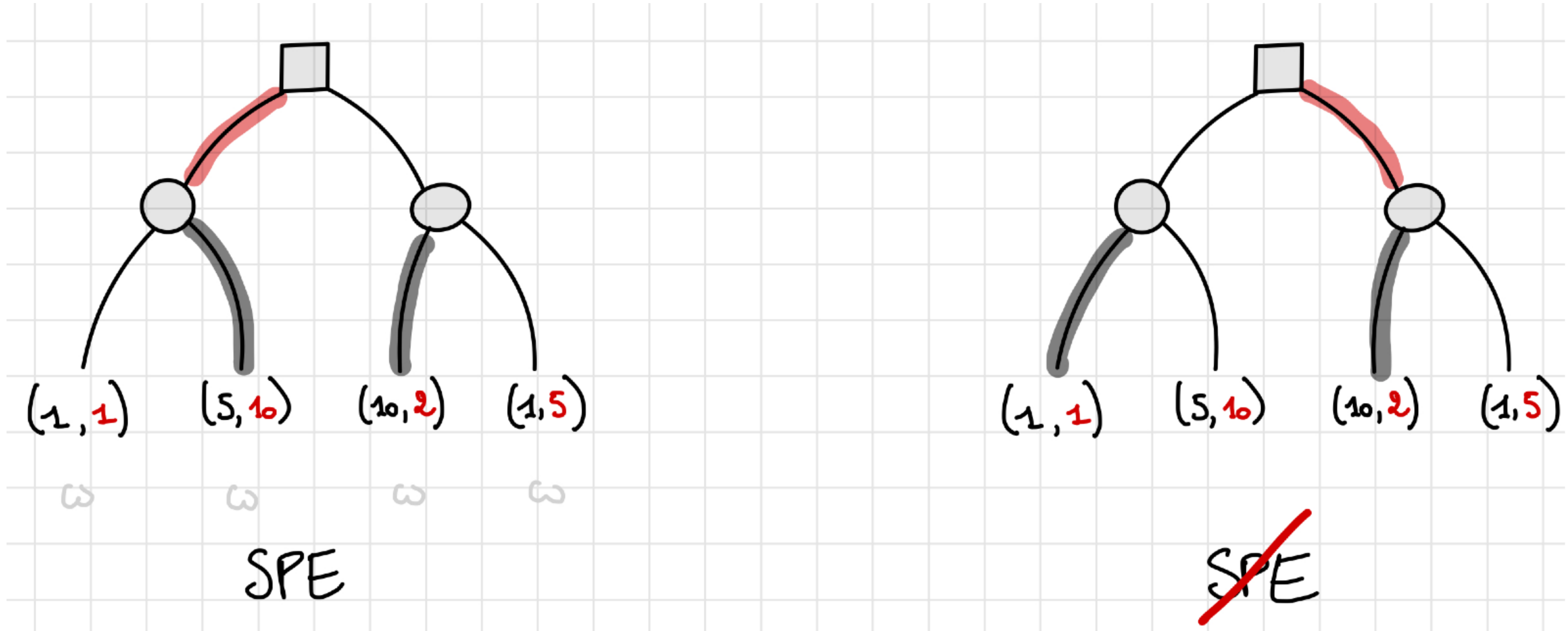
-

How to reason algorithmically on SPE?

Easy case: finite trees

- For finite trees: backward induction



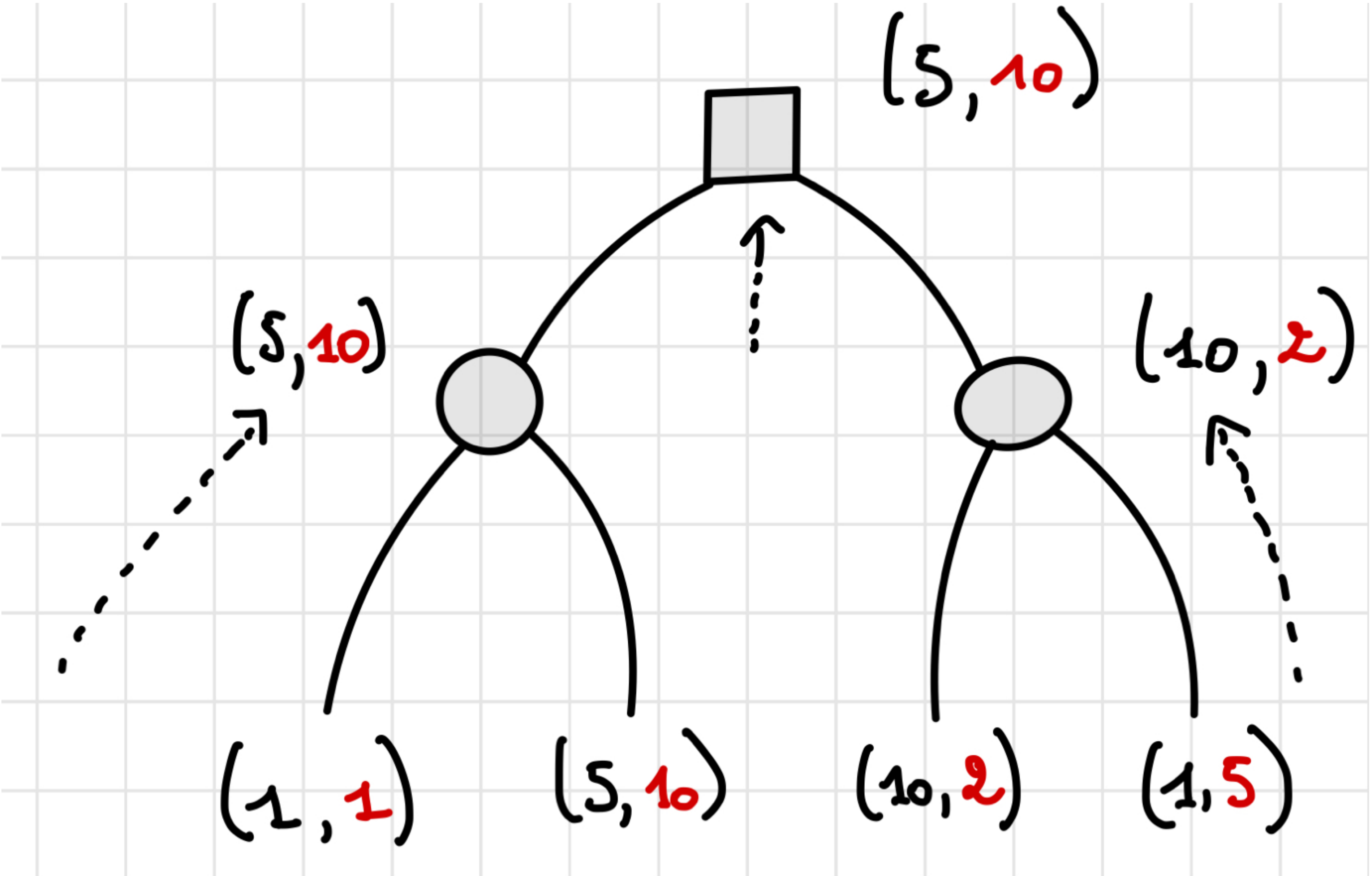


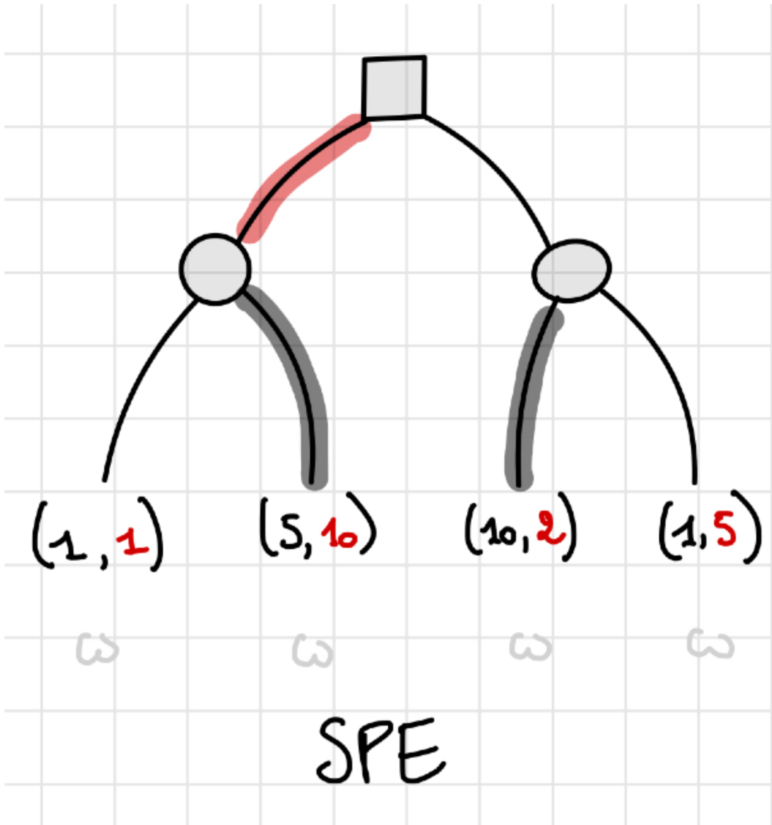
-

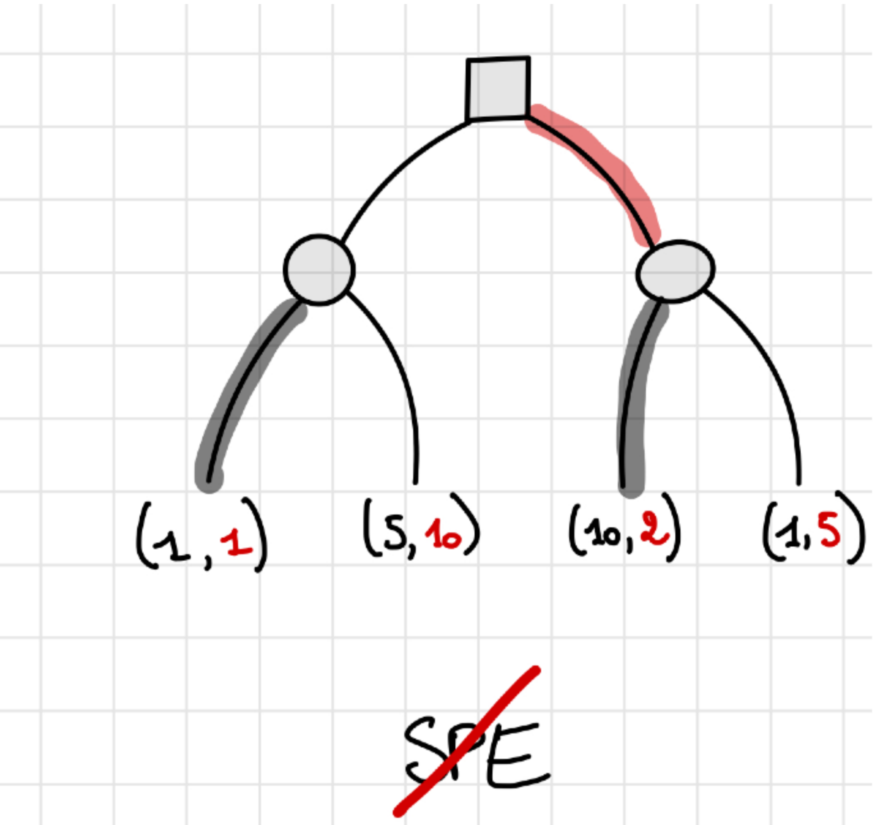
How to reason algorithmically on SPE?

Easy case: finite trees

- For finite trees: backward induction







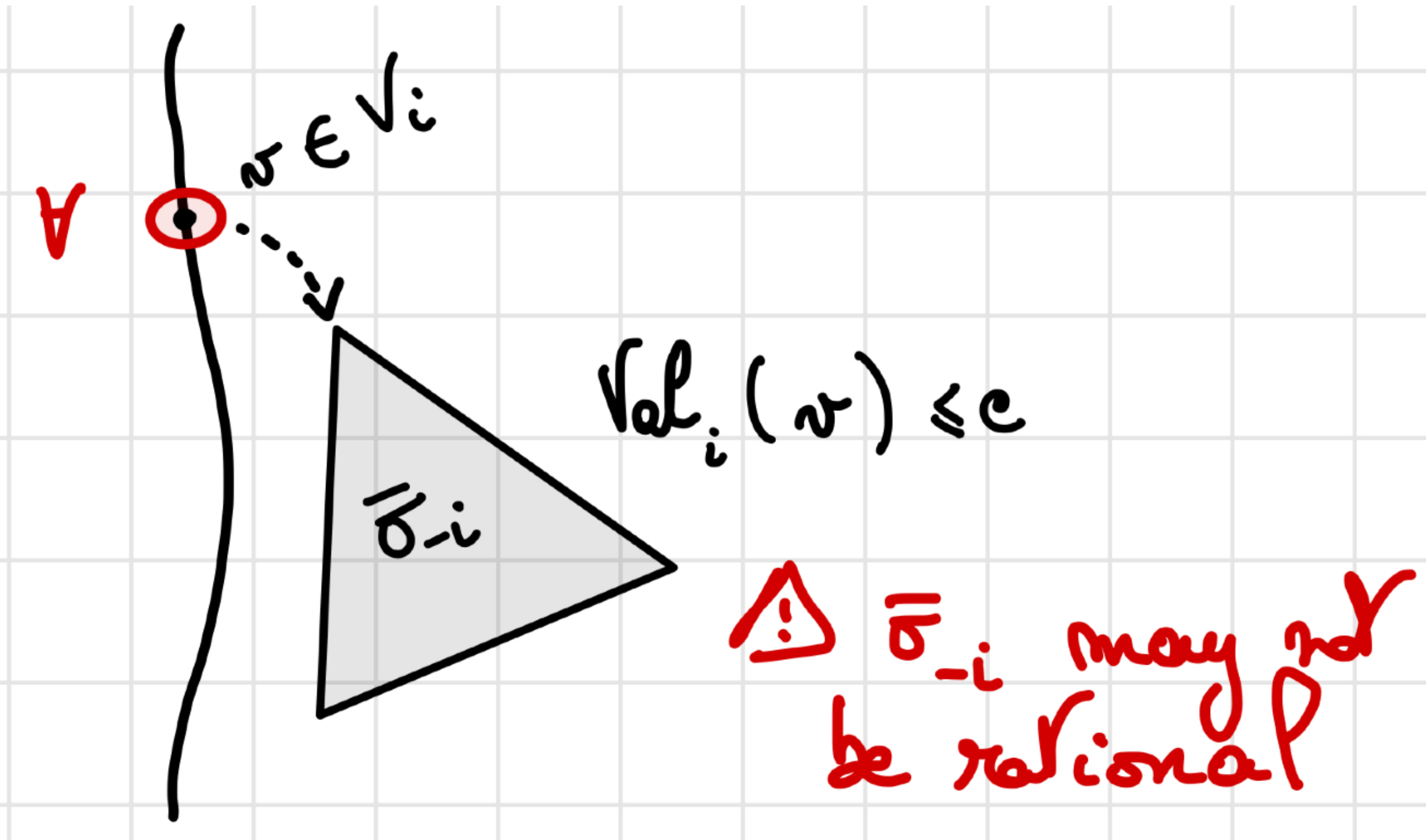
- Infinite trees: backward induction does not generalize…

Better starting point:
Characterization of outcomes of NE

Characterizing outcomes of NE

Use adversarial values

if $\rho \in$ NE



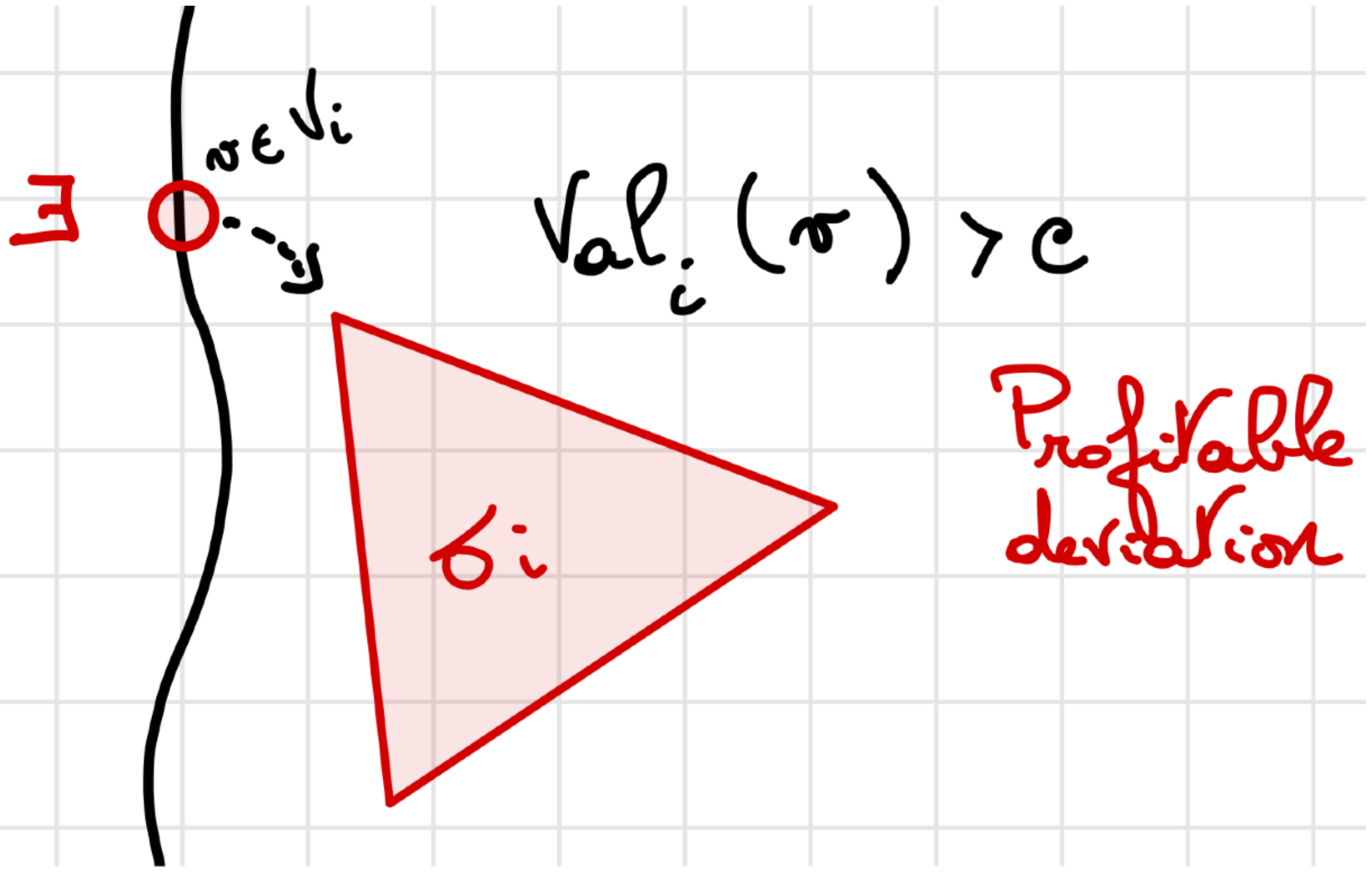
and $\mu_i(\rho) = c$

then

$$c \geq \inf_{\bar{\sigma}_{-i}} \cdot \sup_{\sigma_i} \cdot \mu_i(\text{Outcome}(\sigma_i, \bar{\sigma}_{-i})) = \text{Val}_i(v)$$

Player i has **no** incentive to deviate

if $\rho \notin$ NE



and $\mu_i(\rho) = c$

then

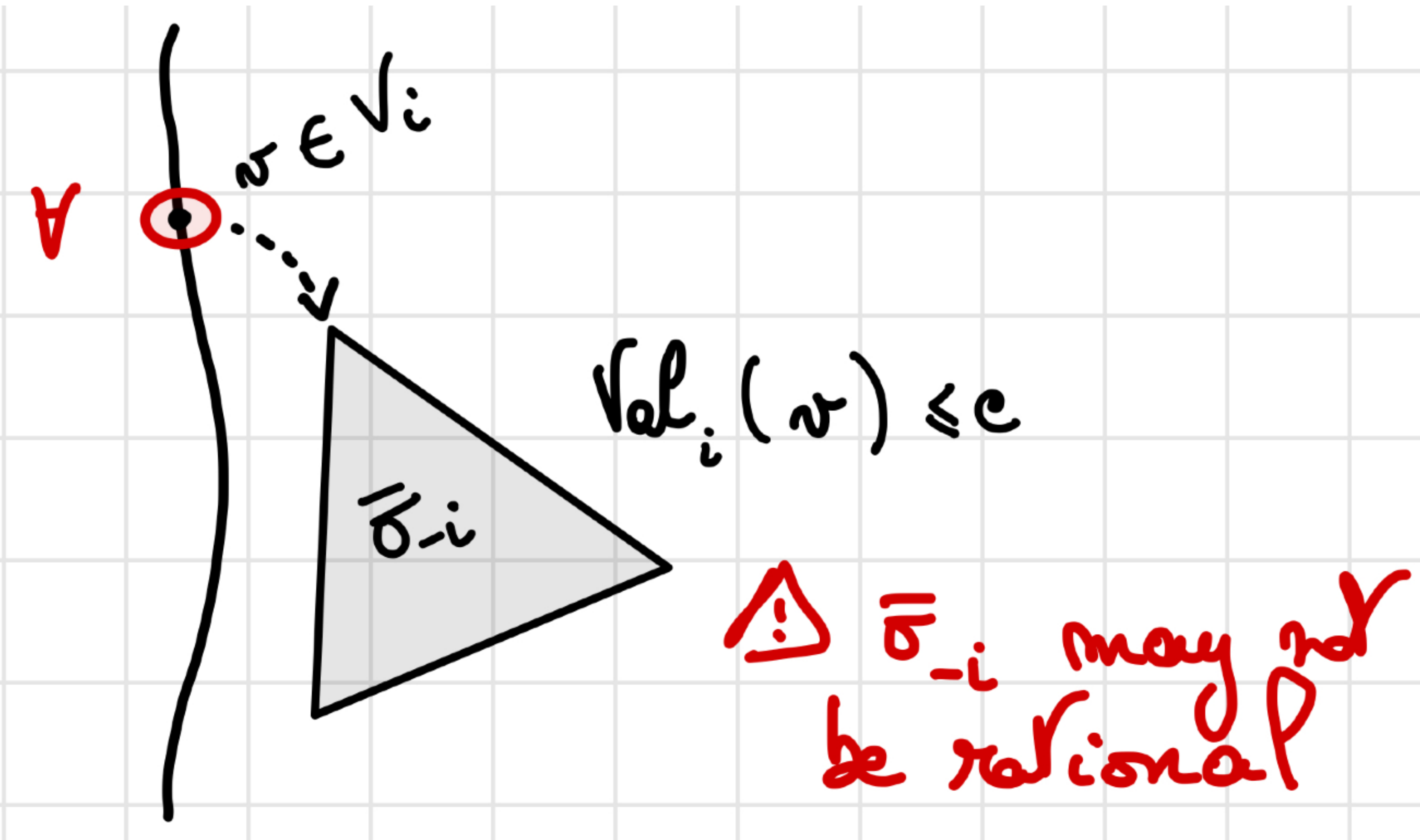
$$\text{Val}_i(v) = \sup_{\sigma_i} \cdot \inf_{\bar{\sigma}_{-i}} \cdot \mu_i(\text{Outcome}(\sigma_i, \bar{\sigma}_{-i})) > c$$

Player i has an incentive to **deviate**

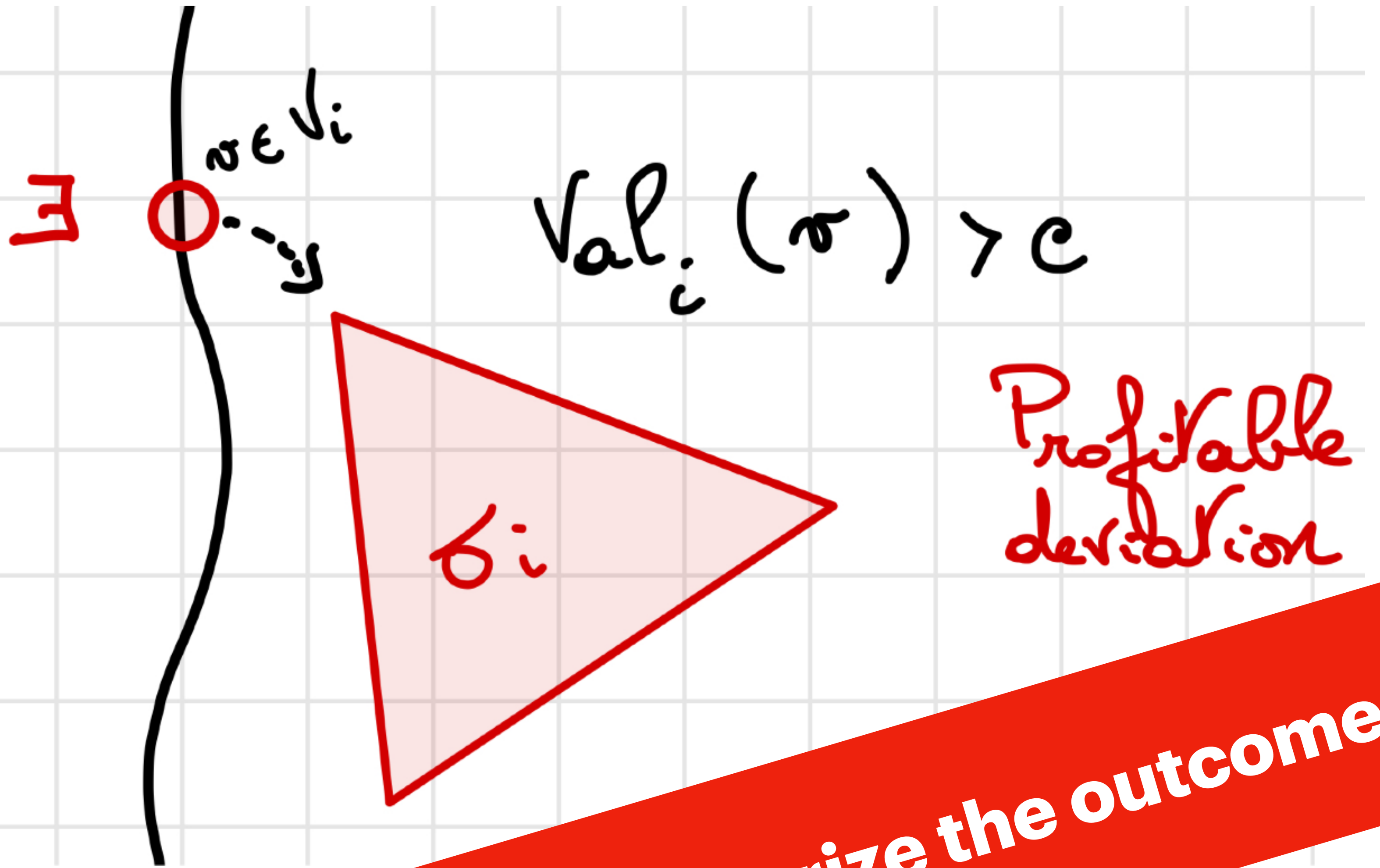
Characterizing outcomes of NE

Use adversarial values

if $\rho \in$ NE



if $\rho \notin$ NE



and $\mu_i(\rho) = c$

then

$c \geq \inf_{\bar{\sigma}_{-i}} \cdot \sup_{\sigma_i} \cdot \mu_i(\text{Outcome}(\sigma_i, \bar{\sigma}_{-i}))$

Player i has **no** incentive to deviate

$\text{Val}_i(v) = \sup_{\sigma_i} \cdot \inf_{\bar{\sigma}_{-i}} \cdot \mu_i(\text{Outcome}(\sigma_i, \bar{\sigma}_{-i})) > c$

Player i has an incentive to **deviate**

The zero-sum two player values (i,-i) characterize the outcomes of NE

Characterizing outcomes of NE

Use adversarial values

- A play $\rho = v_0 v_1 \ldots v_n \ldots$ is supported by a NE if $\forall i \in [1,N] \cdot \forall j \geq 0 : v_j \in V_i \cdot \mu_i(\rho) \geq \mathsf{Val}_i(v_j)$

$$\mathsf{Val}_i(v) = \inf_{\bar{\sigma}_{-i}} \cdot \sup_{\sigma_i} \cdot \mu_i(\mathsf{Outcome}(\sigma_i, \bar{\sigma}_{-i}))$$

- If $\mu_i(\cdot)$ is **prefix independent** (like parity or mean-payoff), this is equivalent to

$$\forall i \in [1,N] \cdot \mu_i(\rho) \geq \max_{v \in \mathsf{visit}(\rho) \cap V_i} \mathsf{Val}_i(v)$$

- So it is sufficient to compute for all $i \in [1,N]$ and vertex $v \in V_i$, the **worst-case value** $\mathsf{Val}_i(v)$ — this is equivalent to solving a two-player zero-sum game

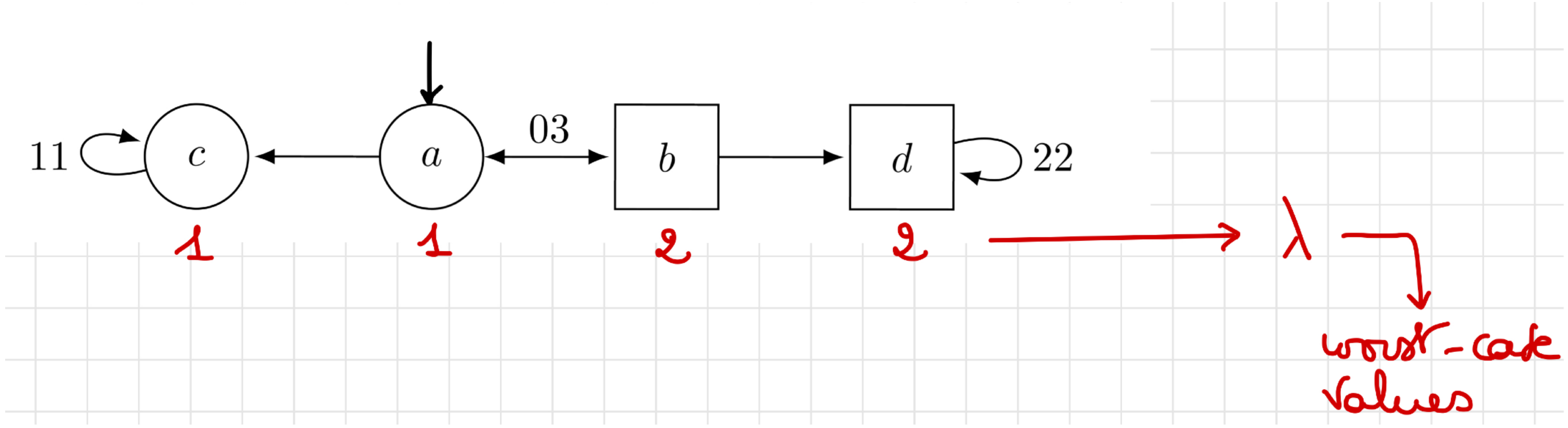
- Let $\lambda : V \to \mathbb{D}$, where $\mathbb{D} = \mathbb{B}$ or $\mathbb{D} = \mathbb{R}$, such that $\lambda(v) = \mathsf{Val}_i(v)$ for $v \in V_i$,

$\rho = v_0 v_1 \ldots v_n \ldots$ is $\lambda -$ consistent iff $\forall i \in [1,N] \cdot \forall j \geq 0 \cdot \mu_i(\rho) \geq \lambda(v_j)$

- Such a function $\lambda : V \to \mathbb{D}$ is called a **requirement**

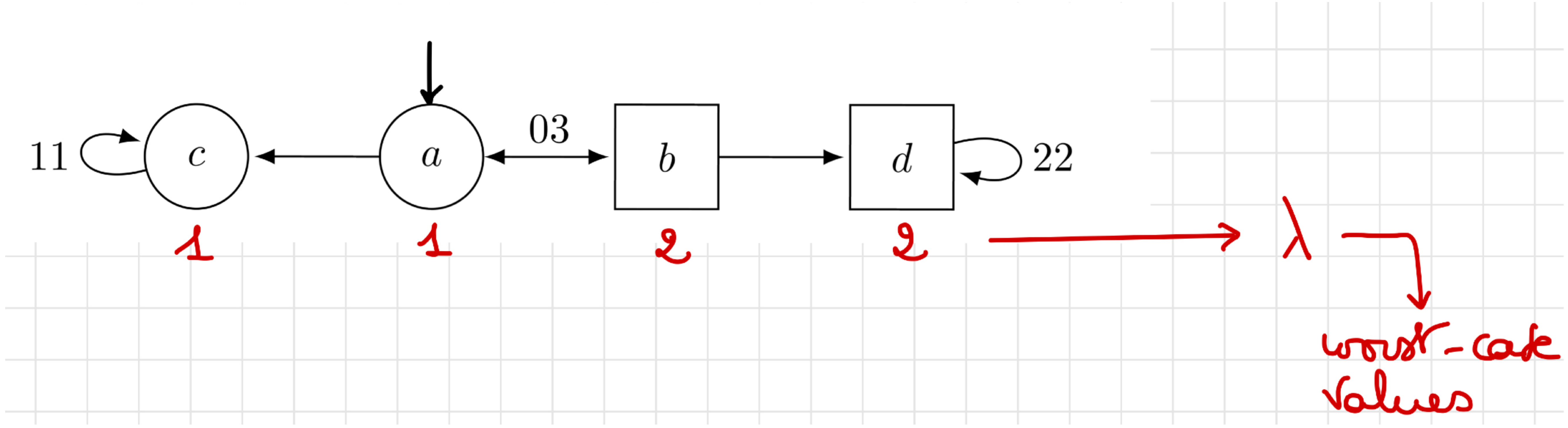
Set of outcomes supported by NE

Example Mean-payoff



Set of outcomes supported by NE

Example Mean-payoff



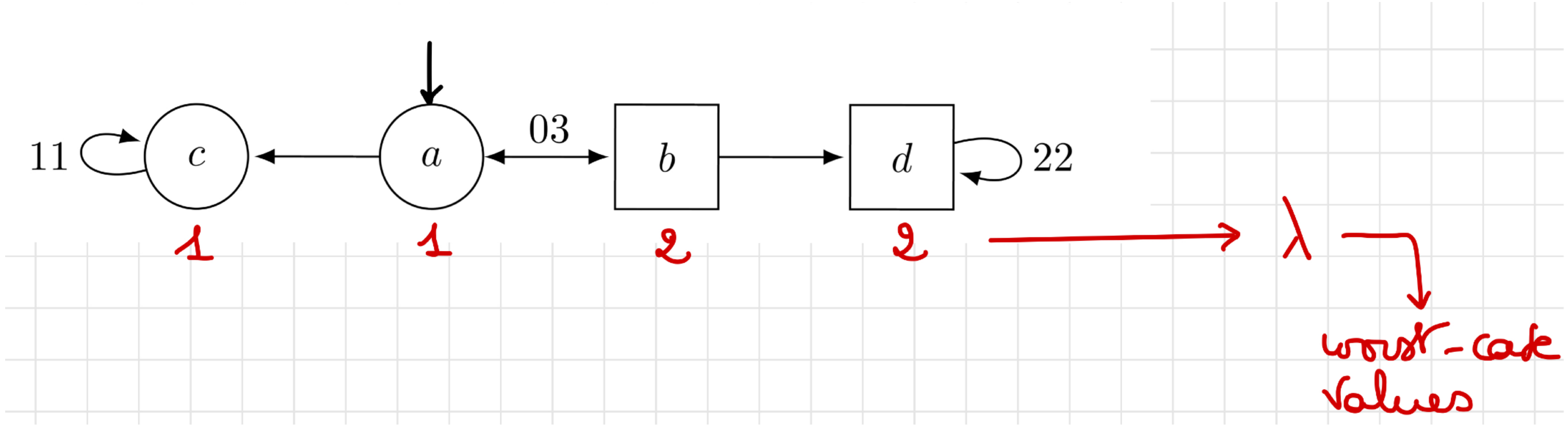
- the set of $\lambda -$ consistent paths in G are:

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

-
-

Set of outcomes supported by NE

Example Mean-payoff



- the set of $\lambda -$ consistent paths in G are:

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

- **Theorem** [Brihaye et al. 13]: $\rho = v_0 v_1 \ldots v_n \ldots \in \text{OutNE}(G)$ iff ρ is $\lambda -$ consistent.

Automaton for OutNE

Applications

- **Corollary (effectivity)**: the set of $\lambda -$ consistent paths is recognized, for **MP** by a multi-MP automaton (this language is not necessarily $\omega -$ regular), and for **Parity** by a Streett-automaton. In both cases, we can solve
  - Existence problem: OutNE$(G) =^? \varnothing$
    (trivial for NE - always non empty)
  - Rational verification (emergence of property ψ):
    $(\exists)$ $\exists \rho \in$ OutNE$(G) : \rho \models \psi$? $(\forall)$ $\forall \rho \in$ OutNE$(G) : \rho \models \psi$?
  - Cooperative rational synthesis [Kuperfman et al.]:
    $\exists \rho \in$ OutNE(G) : $\rho \models p_0$?   (parity obj. of Player 0 is true)
    $\exists \rho \in$ OutNE(G) : val$_0(\rho) \geq c$

Generalization to SPE

Relative worst-case value

- **Question**: given the requirement λ_1 defined by the worst-case values and a vertex $v \in V_i$, does player i have a strategy to **improve** the value that she can obtain from v if the other players are **not willing to give up** their worst-case value (as required by λ_1) ?
- **Can we compute a** $\lambda -$ relative **worst-case value ?**

Generalization to SPE

Relative worst-case value - The negotiation function

- $$\text{Nego} : [\lambda \to \mathbb{D} \cup \{+\infty\}] \to [\lambda \to \mathbb{D} \cup \{+\infty\}]$$

  where $\text{Nego}(\lambda)(v) = \inf_{\bar{\sigma}\_{-i} \in \lambda\text{Rat}} \cdot \sup_{\sigma_i \in \Sigma_i} \mu_i(\text{outcome}(\sigma_i, \bar{\sigma}\_{-i}))$

  i.e. it computes the worst-case value against λRat strategies, i.e. against players that do not want to trade off the value promised by λ.
- This can be reduced to a zero sum game (Prover/Challenger).

How to compute Nego(.)

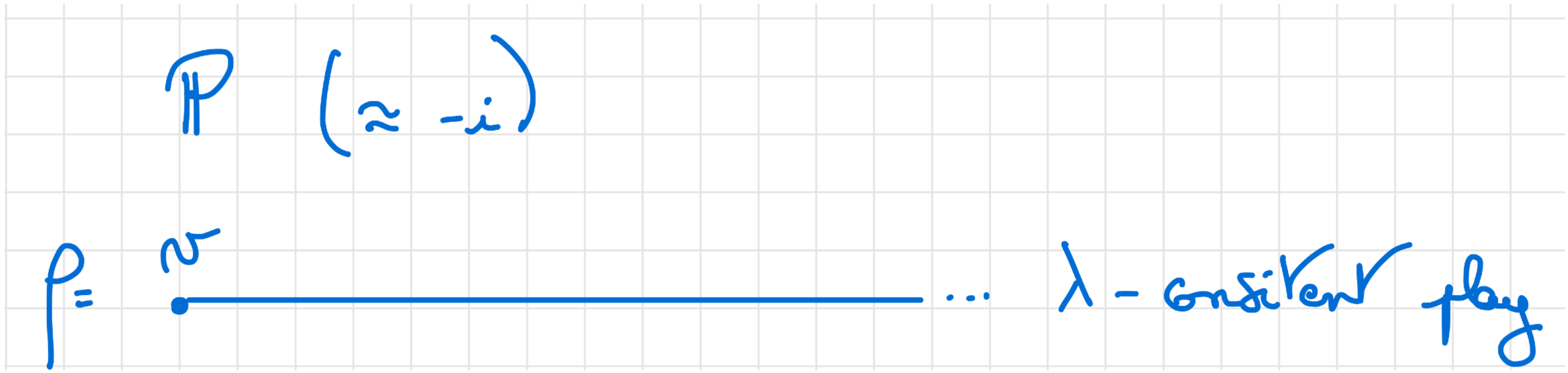
The abstract negotiation game between Prover and Challenger

- Let $v \in V_i$, $\mathbb{P} \approx -i$ want to prove that $\text{Nego}(\lambda)_i(v) \leq \alpha$ to $\mathbb{C} \approx i$

How to compute Nego(.)

The abstract negotiation game between Prover and Challenger

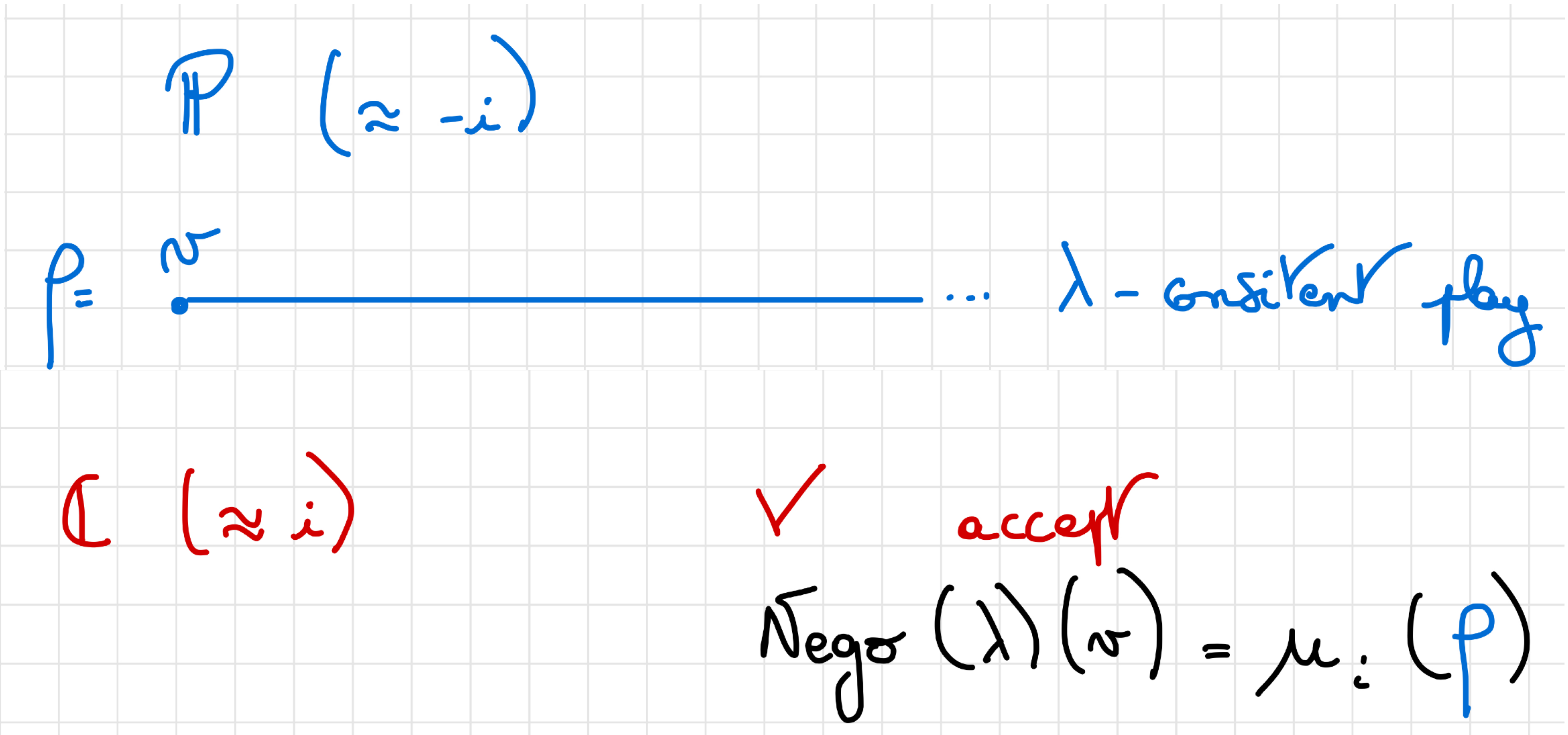
- Let $v \in V_i$, $\mathbb{P} \approx -i$ want to prove that $\text{Nego}(\lambda)_i(v) \leq \alpha$ to $\mathbb{C} \approx i$



How to compute Nego(.)

The abstract negotiation game between Prover and Challenger

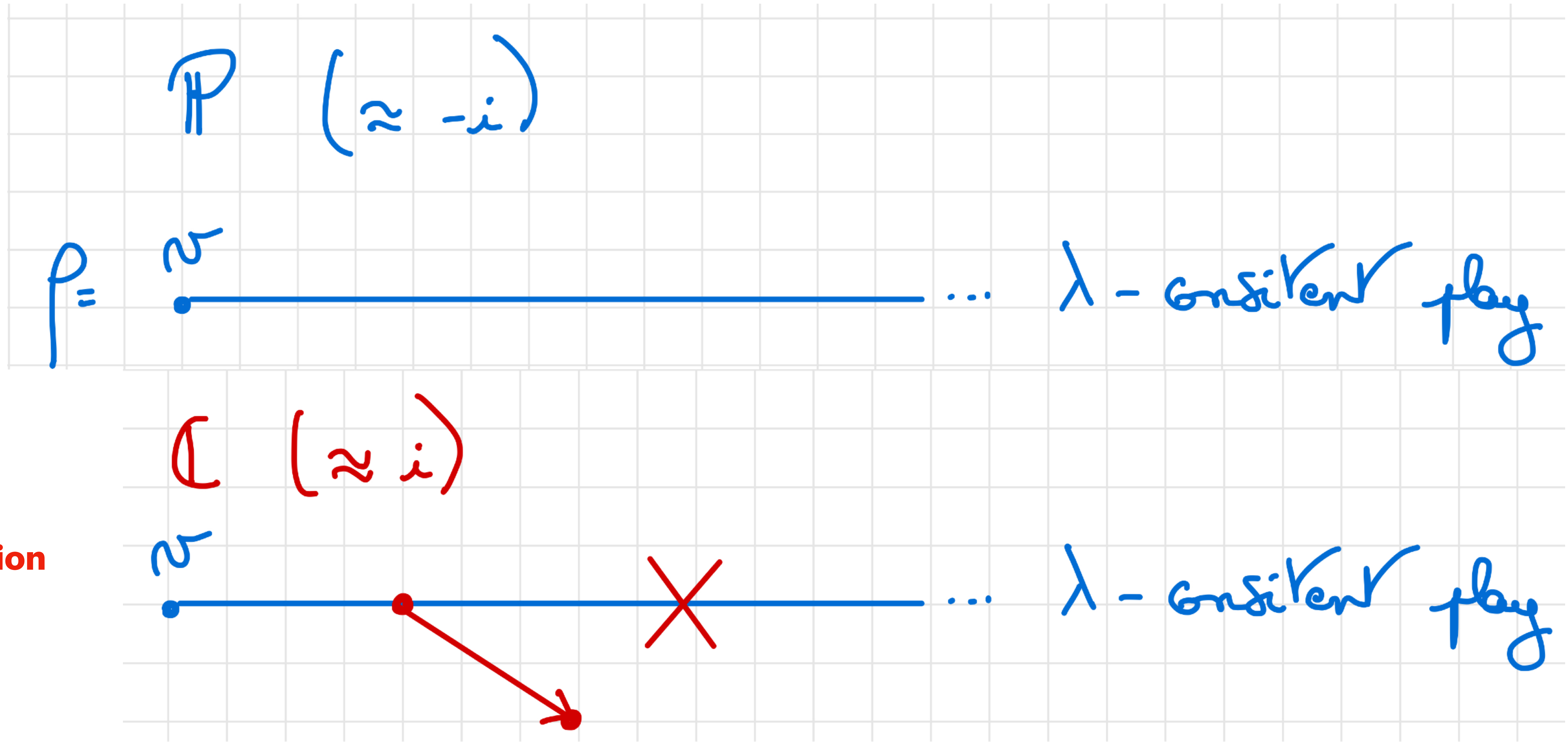
- Let $v \in V_i$, $\mathbb{P} \approx -i$ want to prove that $\text{Nego}(\lambda)_i(v) \leq \alpha$ to $\mathbb{C} \approx i$



How to compute Nego(.)

The abstract negotiation game between Prover and Challenger

- Let $v \in V_i$, $\mathbb{P} \approx -i$ want to prove that $\text{Nego}(\lambda)_i(v) \leq \alpha$ to $\mathbb{C} \approx i$

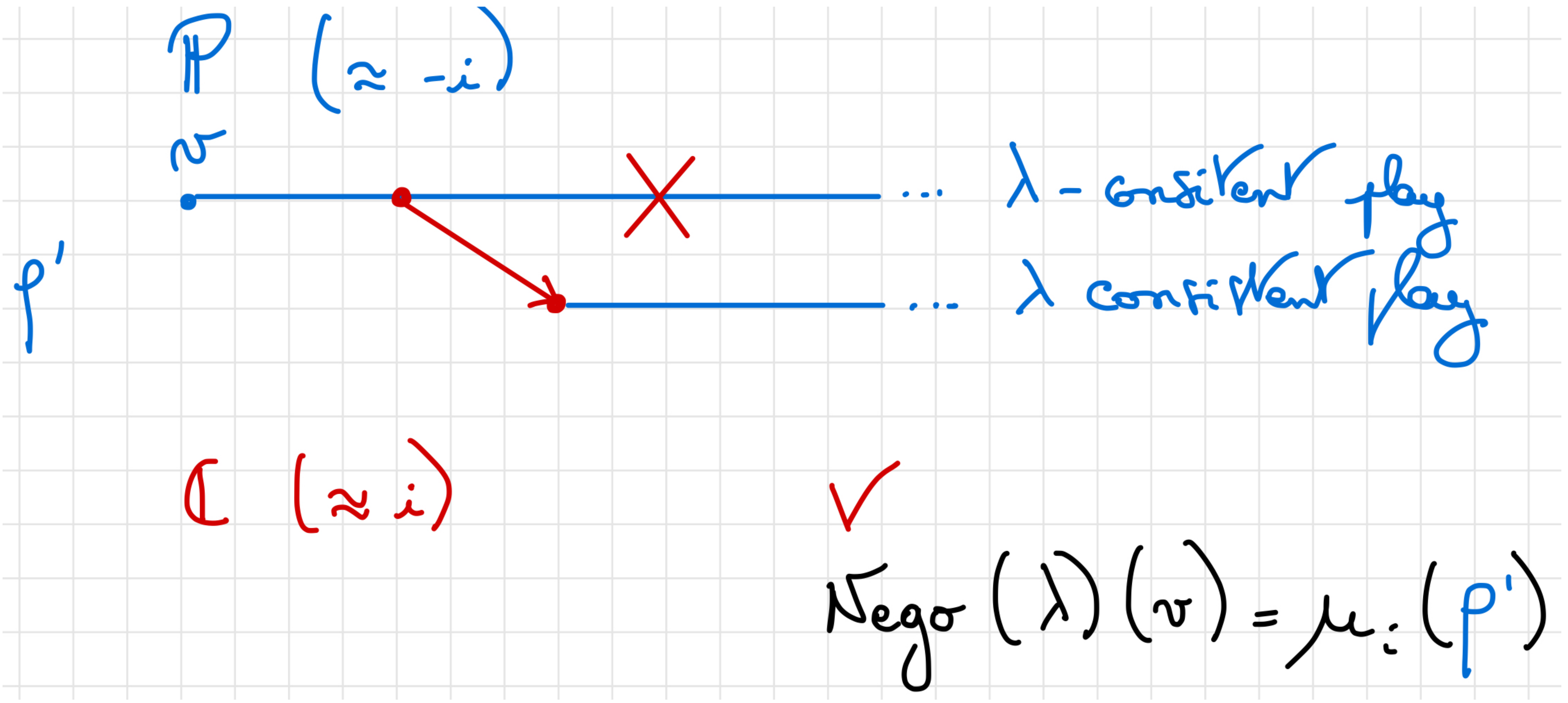


deviation

How to compute Nego(.)

The abstract negotiation game between Prover and Challenger

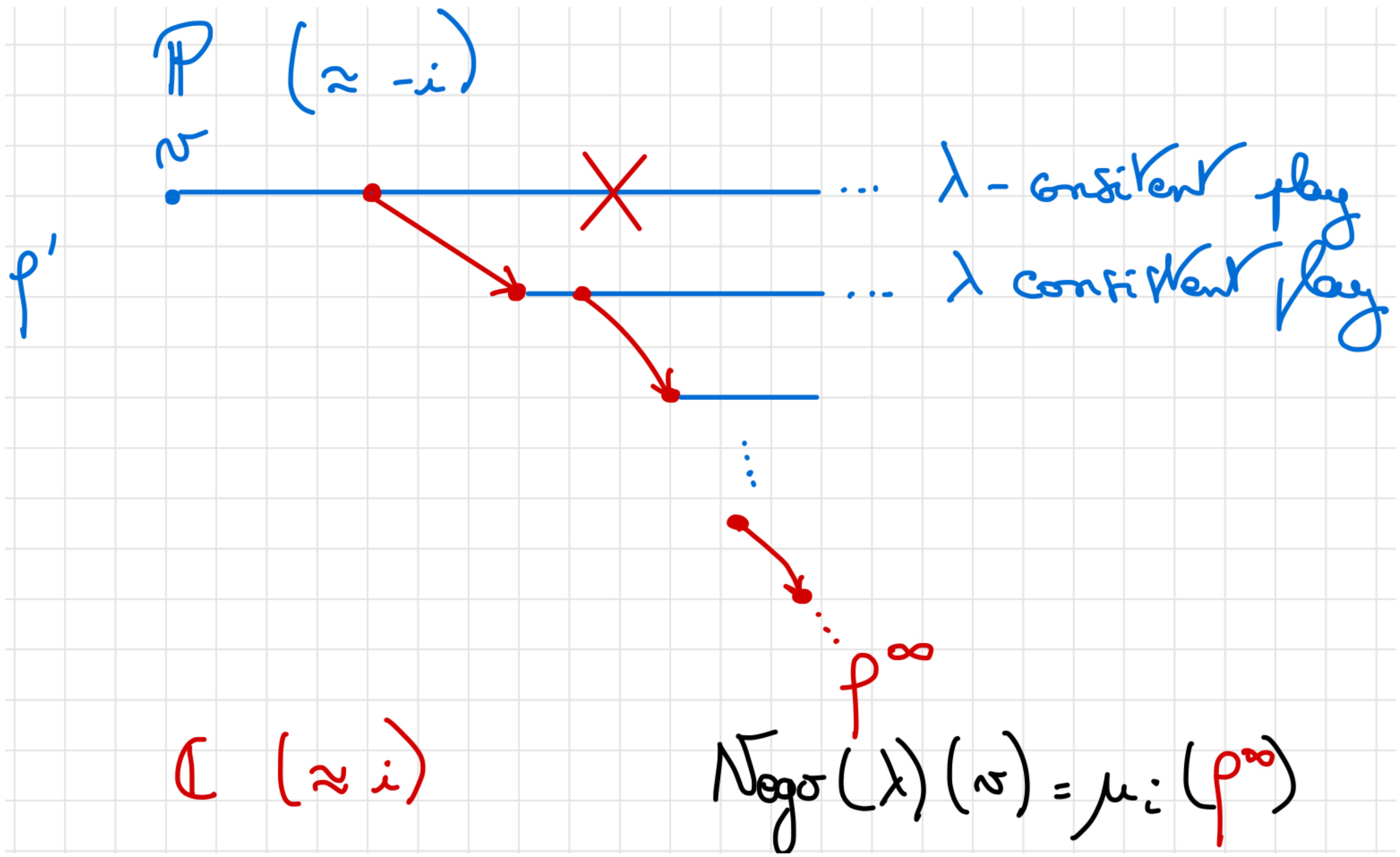
- Let $v \in V_i$, $\mathbb{P} \approx -i$ want to prove that $\text{Nego}(\lambda)_i(v) \leq \alpha$ to $\mathbb{C} \approx i$



How to compute Nego(.)

The abstract negotiation game between Prover and Challenger

- Let $v \in V_i$, $\mathbb{P} \approx -i$ want to prove that $\text{Nego}(\lambda)_i(v) \leq \alpha$ to $\mathbb{C} \approx i$



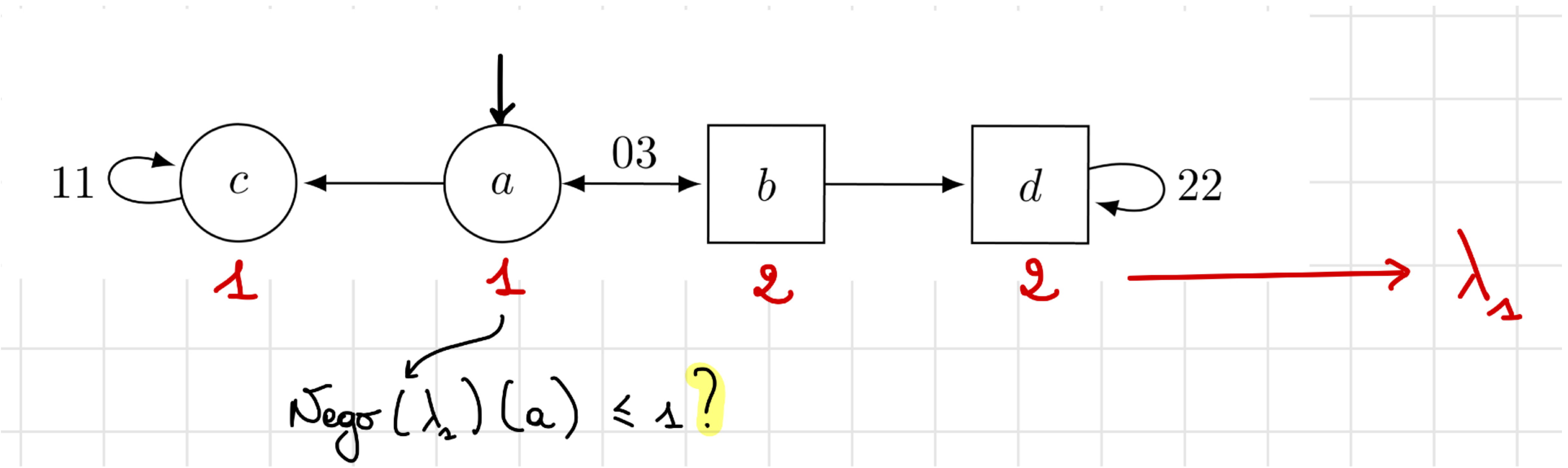
How to compute Nego(.)

The abstract negotiation game between Prover and Challenger

- **Theorem** [Concur'21]: Nego$(\lambda)(v)$ is equal to the value of the abstract negotiation game.
- **Theorem** [Concur'21]: The abstract negotiation game for MP can be transformed into a **finite state multi-mean payoff parity game.**
- **Theorem** [CSL'22]: The abstract negotiation game for Parity can be transformed into a **Streett game**.

How to compute Nego(.)

... an example



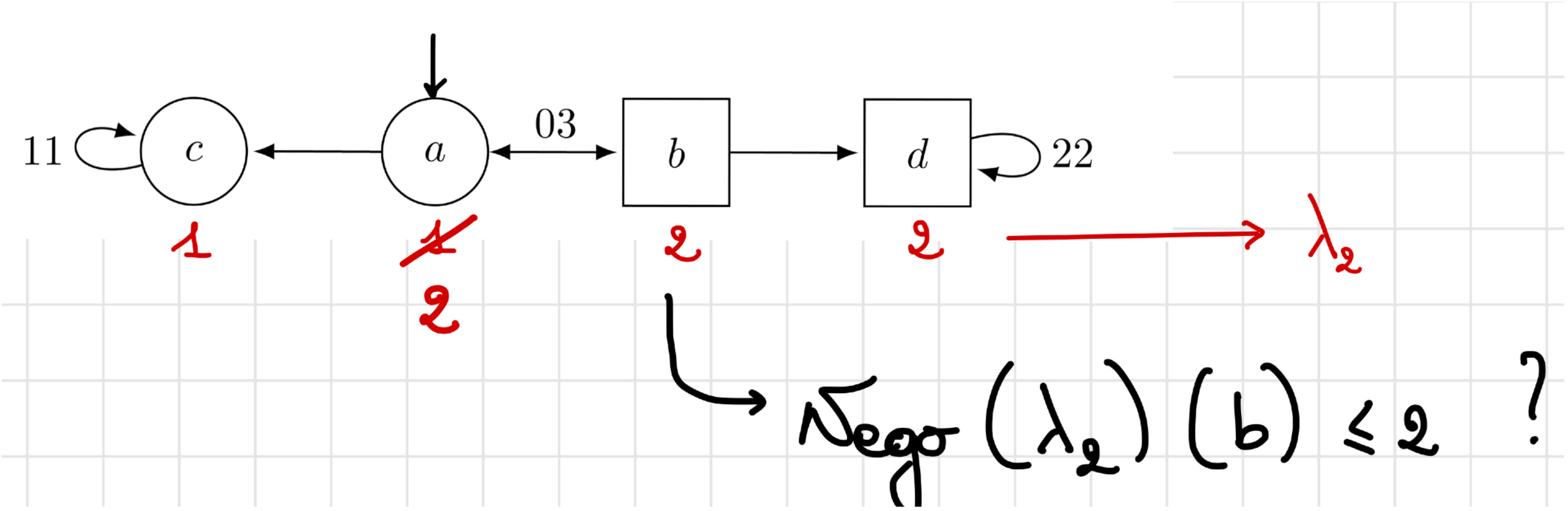
- $\mathbb{P}$: $a \to c^{\omega}$ (this path is λ_1 – consistent)
- $\mathbb{C}$: deviation $a \to b$
- $\mathbb{P}$: from b, the only λ_1 – consistent paths are in $(b \to a)^{\star} \to d^{\omega}$

(even if $(a \to b \to)^{\omega}$ is tempting but it fails to give 1 to Player 1)

  - As $\mathrm{MP}_1((b \to a)^{\star} \to d^{\omega}) = 2$, Nego$(\lambda_1)(a) = 2$.

How to compute Nego(.)

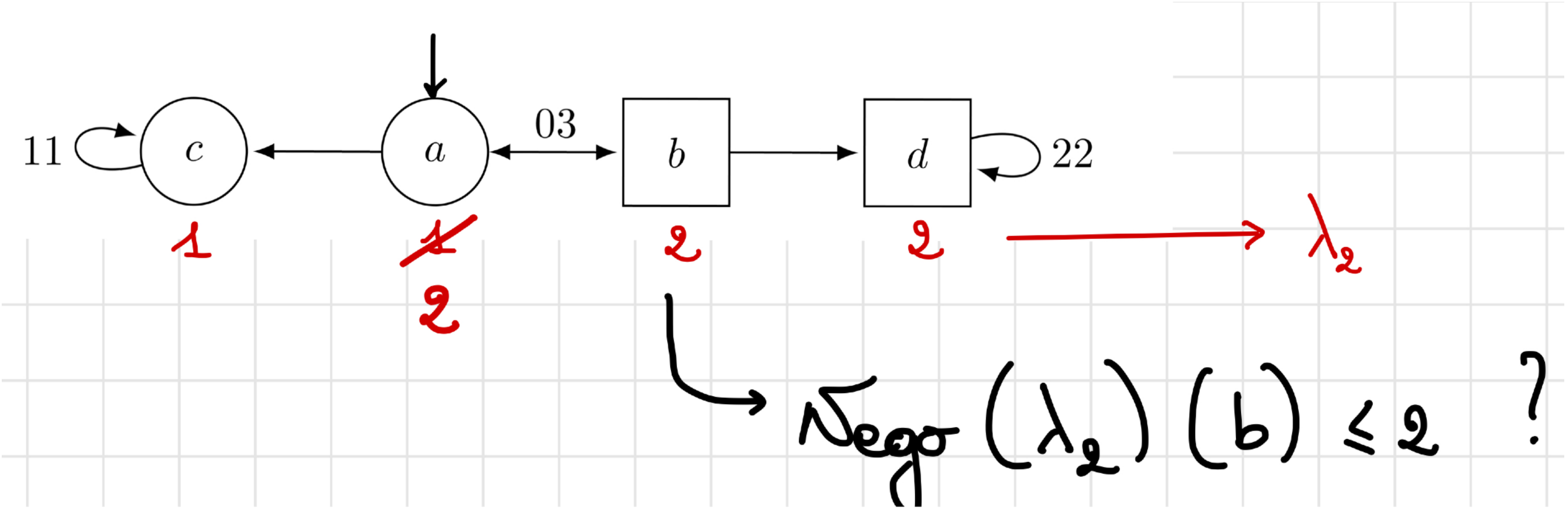
... an example



- $\mathbb{P}$: $(b \to d)^\omega$
- $\mathbb{C}$: deviation $b \to a$
- $\mathbb{P}$: $a \to (b \to d)^\omega$
- $\mathbb{C}$: deviation $b \to a$
- if we repeat the last two steps for ω rounds, we get $\rho = (b \to a \to)^\omega$ and so Nego$(\lambda_2)(b) = \text{MP}_2(\rho) = 3.$

How to compute Nego(.)

… an example



- $\mathbb{P}$: $(b \to d)^\omega$
- $\mathbb{C}$: deviation $b \to a$
- $\mathbb{P}$: $a \to (b \to d)^\omega$
- $\mathbb{C}$: deviation $b \to a$
- if we repeat the last two step

  and so $\text{Nego}(\lambda_2)(b) = \text{MP}_2($

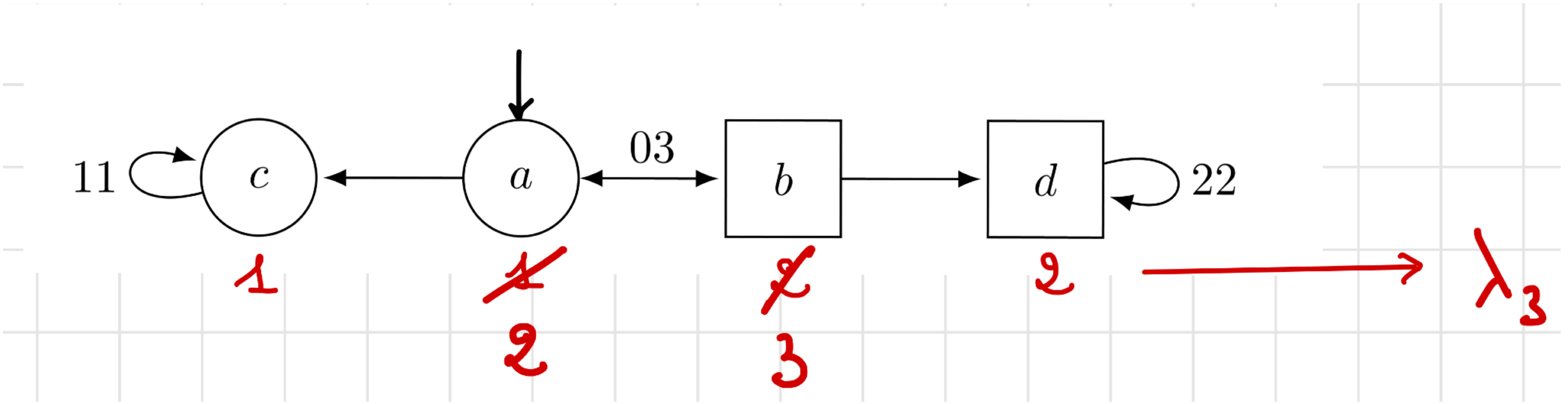
When should we stop ?

Iterate up to (least) fixed point !

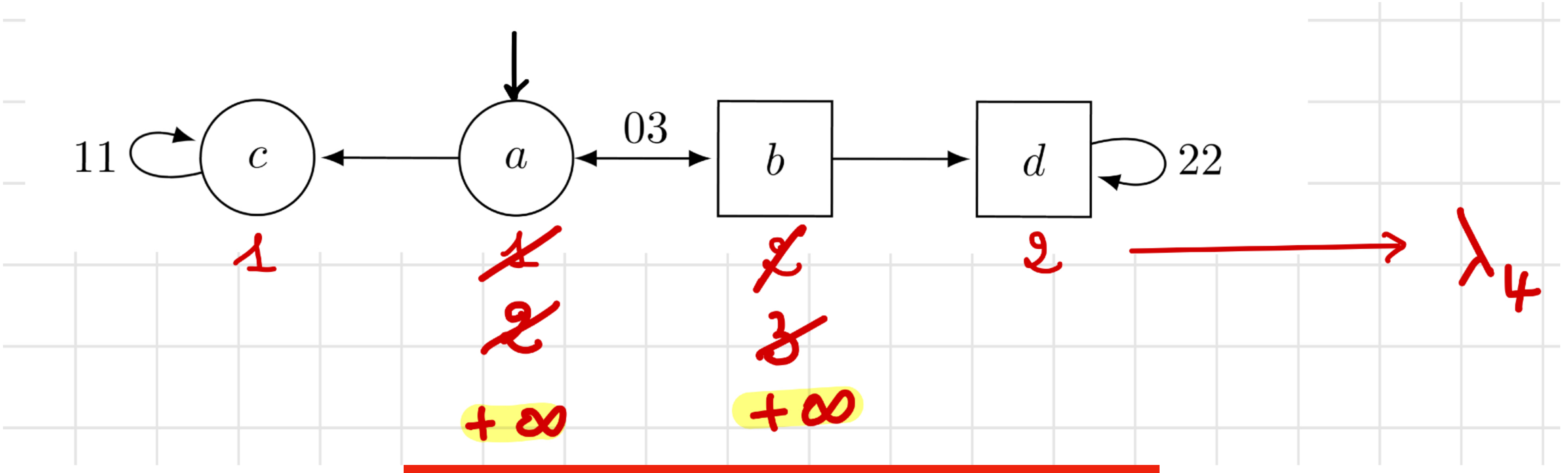
The least fixed point characterizes all the outcomes of SPEs !

Least fixed point characterizes all SPEs

… an example



- There is no λ_3 – consistent path from a, nor from b !



No SPE starting from a or b !

Least fixed point characterization

Prefix independent objectives

- **Theorem [Concur'21]**: For *prefix independent* (including MP and parity) objectives, the set of all outcomes of SPEs is characterized by the least fixed point $\lambda^\star$ of $\mathrm{Nego}(\,\cdot\,)$, i.e.:

$$\mathrm{OutSPE}_{v_0}(G) = \bigcup_{\bar{\sigma} \in SPE} \mathrm{Outcome}_{v_0}(\bar{\sigma}) = \{\rho \mid \rho \text{ is } \lambda^\star - \text{consistent}\}$$

- For Parity objectives, $\lambda^\star$ is reached within $|V|$ steps
- For MP objectives, reaching $\lambda^\star$ may require transfinite number of iterations (but the complexity of $\lambda^\star$ can be bounded)

Complexity
CSL'22

- **Theorem** [CSL'22]: Given a N-player **parity** game G:
  - Constrained existence problem: - existence is always guaranteed
    given two vectors $u, v \in \mathbb{B}^N$, deciding if there exists a SPE $\bar{\sigma}$ such that $u \leq \mu(\text{outcome}(\bar{\sigma})) \leq v$ is NP – Complete.
    —The notion of witness is non trivial
    —This was previously now to be in ExpTime (using alternating automata constructions)
  - Least fixed point checking problem:
    given a vector $\lambda \in \mathbb{B}^N$, deciding if $\lambda = \lambda^\star$ is BH_2 – complete.
  - LTL verification problem:
    given a LTL formula ψ, deciding if **for all** SPE $\bar{\sigma}$, we have that $\text{outcome}(\bar{\sigma}) \vDash \psi$, ie. checking if $\text{OutSPE}(G) \vDash \psi$, is PSpace – Complete.

Complexity

Mean-Payoff (ICALP'22)

- **Theorem** [ICALP'22]: Given a N-player mean-payoff game G:
  - Constrained SPE existence problem: given $u, v \in \mathbb{Q}^N$, deciding if there exists a SPE $\bar{\sigma}$ s.t. $u \leq \mu(\text{outcome}(\bar{\sigma})) \leq v$ is $\text{NP} - \text{Complete}$.
  - The "plain" existence problem is also $\text{NP} - \text{Complete}$.

—The notion of witness is non trivial
—We know that the least fixed point is the solution of a set of linear equations for which we can bound the size of solutions - and so we can guess it
—The decidability status of this problem was left open in the literature

Summary - Conclusion

- SPE is a **natural** solution concept to model rationality in multi-player graph games, and SPE is better suited than NE for sequential games (non-credible threats)
- We have described new algorithmic ideas to compute an effective representation of the set of **outcomes supported by a SPE** of a N-player non-zero game graph (for parity and mean-payoff). This is relevant to solve rational verification problems and cooperative rational synthesis problems
- We have characterized the complexity of the **(constrained) existence problems** for SPE in N-player non-zero sum games played on graphs with mean-payoff and parity objectives (both are **NP-complete** problems)