Games with payoff

Wiesław Zielonka

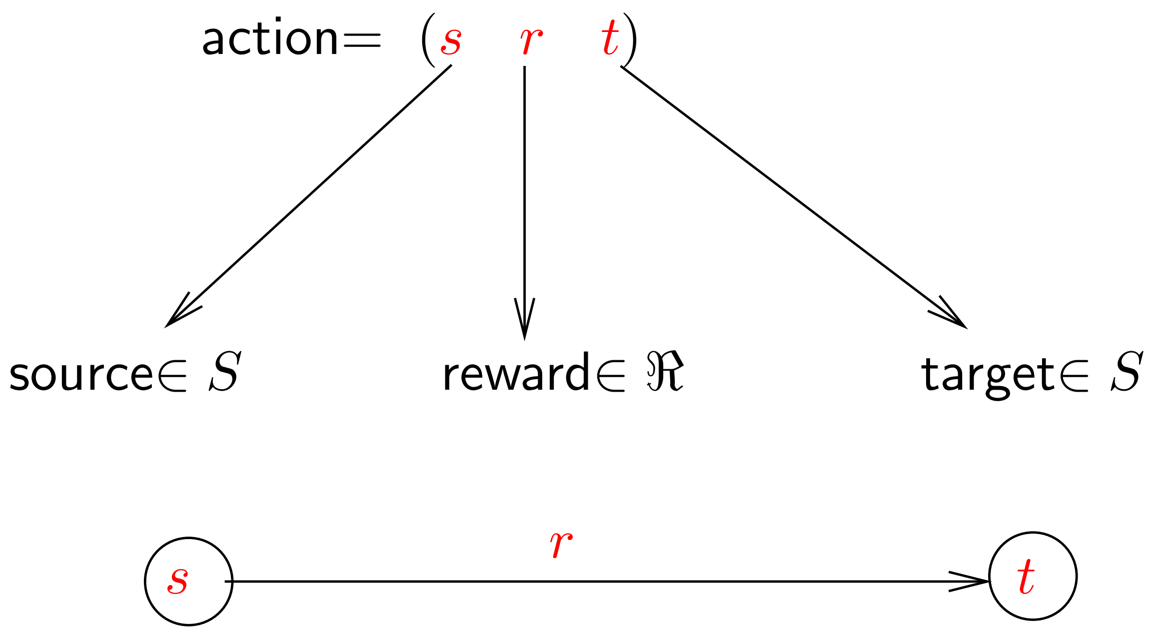
LIAFA, Université Paris 7 Denis Diderot and CNRS

Games

Players: 1 and 2.

Arenas:
S_1, S_2 — states controlled by players 1 and 2,
A – actions,
$\Re$ – rewards.

$S = S_1 \cup S_2$ – all states,
$A \subset S \times \Re \times S$.



$$A(s) = \{a \in A \mid \mathrm{source}(a) = s\}$$

actions *available* at $s \in S$.

Arena

$$\mathcal{A} = (S_1, S_2, A)$$

$S, A(s), \forall s \in S$ – finite, nonempty.

A *history* $h = a_1 a_2 \ldots \in A^{\infty}$,

$$\forall i, \mathrm{target}(a_i) = \mathrm{source}(a_{i+1}).$$

A *play* $p = a_1 a_2 \ldots$ — an infinite sequence of actions executed by the players (play=infinite history).

Preference relation

For a play $p = a_1 a_2 \ldots$

$$\mathrm{reward}(p) = \mathrm{reward}(a_1)\,\mathrm{reward}(a_2)\ldots \in \Re^\omega$$

an infinite finitely generated sequence of rewards.

A preference relation $\sqsubseteq$ complete, transitive and reflexive relation over $\Re^\omega$, $r, r' \in \Re^\omega$,

$$r \sqsubseteq r'$$

means that for player p the sequence r' is at least as valuable as the sequence r.

Payoff mapping

Payoff mapping

$$u : \Re^{\omega} \longrightarrow \mathbb{R} \cup \{-\infty, \infty\}$$

induces a preference relation $\sqsubseteq_u$:

$$r \sqsubseteq_u r' \quad \text{iff} \quad u(r) \leq u(r') \ .$$

Game

Game : arena $\mathcal{A}$ and preference relations $\sqsubseteq_1, \sqsubseteq_2$ of each player.

An antagonistic game: (zero-sum game) $\sqsubseteq_2 = (\sqsubseteq_1)^{-1}$.

$$\mathbf{G} = (\mathcal{A}, \sqsubseteq) \quad ,$$

$\mathcal{A}$ — an arena,
$\sqsubseteq$ — a preference relation for player 1.

Examples

Parity games

$\Re = \mathbb{N}$, rewards = priorities.

For $n = n_1 n_2 \ldots \in \mathbb{N}^\omega$ let

$$\text{priority}(n) = \limsup_{i \to \infty} n_i$$

maximal priority occurring infinitely often in n.

$m, n \in \mathbb{N}^\omega$,

$$n \sqsubseteq m \quad \text{if} \quad \text{priority}(n) \bmod 2 \leq \text{priority}(m) \bmod 2$$

Mean-payoff games

$\Re = \mathbb{R}$, for $r = r_1 r_2 \ldots \in \mathbb{R}^\omega$,

$$\overline{\mathrm{mean}}(r) = \limsup_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} r_i \quad .$$

Exercise

In this exercise we assume that in mean-payoff games both players have optimal positional strategies. (We shall prove it later.)

Show that parity games can be reduced to mean-payoff games.

In other words, suppose that you have an algorithm solving mean-payoff games, i.e calculating the value of the mean-payoff game for each initial state and optimal positional strategies for both players. Then show how such an algorithm can be used to solve parity games.

Is your reduction parity$\Longrightarrow$mean-payoff polynomial?

Discounted games

$\Re = \mathbb{R}$, for $r = r_0 r_1 \ldots \in \mathbb{R}^\omega$,

$$\mathrm{disc}_\beta(r) = (1-\beta)\sum_{i=0}^{\infty} \beta^i r_i \ ,$$

$\beta \in (0,1)$ is a *discount factor*.

Strategies

A *strategy* for player 1

$$\sigma : \{h \mid h \text{ a finite history with } \mathrm{target}(h) \in S_1\} \longrightarrow A$$

where $\sigma(h) \in A(\mathrm{target}(h))$.

$\forall s \in S$, λ_s empty history $\mathrm{source}(\lambda_s) = \mathrm{target}(\lambda_s) = s$.

Histories consistent with a strategy

A history $h = a_1 a_2 \dots$ is *consistent* with a strategy σ of player 1 if for each $i < |h|$, if $h_i < h$ is the prefix of h of length i such that $\mathrm{target}(h_i) \in S_1$ then

$$a_{i+1} = \sigma(h_i) \ .$$

Notation:
σ and τ — strategies for players 1 and 2 respectively.

$$p_{\mathcal{A}}(s, \sigma, \tau)$$

unique play consistent with σ and τ with source s.

Optimal strategies

$\sigma^\sharp \in \mathrm{Strategy}_1$, $\tau^\sharp \in \mathrm{Strategy}_2$ are *optimal* if

$\forall s \in S$, $\forall \sigma \in \mathrm{Strategy}_1, \tau \in \mathrm{Strategy}_2,$

$$\mathrm{reward}(p_{\mathcal{A}}(s, \sigma, \tau^\sharp)) \sqsubseteq \mathrm{reward}(p_{\mathcal{A}}(s, \sigma^\sharp, \tau^\sharp)) \sqsubseteq \mathrm{reward}(p_{\mathcal{A}}(s, \sigma^\sharp, \tau)) \quad .$$

$\forall s \in S$, $\forall \sigma \in \mathrm{Strategy}_1, \tau \in \mathrm{Strategy}_2,$

$$u(\mathrm{reward}(p_{\mathcal{A}}(s, \sigma, \tau^{\sharp}))) \leq$$

$$u(\mathrm{reward}(p_{\mathcal{A}}(s, \sigma^{\sharp}, \tau^{\sharp})))$$

$$\leq u(\mathrm{reward}(p_{\mathcal{A}}(s, \sigma^{\sharp}, \tau)))$$

Value of the game at s

General value definition

$$\text{lower value} = \underline{\text{val}}(s) =$$

$$\sup_{\sigma} \inf_{\tau} u(\text{reward}(p_{\mathcal{A}}(s, \sigma^{\sharp}, \tau^{\sharp}))) \leq \inf_{\tau} \sup_{\sigma} u(\text{reward}(p_{\mathcal{A}}(s, \sigma^{\sharp}, \tau^{\sharp})))$$

$$= \overline{\text{val}}(s) = \text{upper value}$$

and

$$\text{value}(s) := \underline{\text{val}}(s) = \overline{\text{val}}(s) \quad .$$

A *positional* or *memoryless* strategy for player 1

$$\sigma : S_1 \to A$$

where $\sigma(s) \in A(s)$, $\forall s \in S_1$.

Basic questions of game theory

- Does there exist a value for a given game $\mathbf{G}$?
- The existence of optimal strategies for both players.
- The existence of "simple" optimal strategies.

Have the players optimal positional strategies?

More examples

Simple Priority Games

$\alpha : \mathbb{N} \longrightarrow \mathbb{R}$, a priority valuation.

Let $n = n_1 n_2 \ldots \in \mathbb{N}^\omega$. Then the payoff of *simple priority games*:

$$u_\alpha(n) = \alpha(\text{priority}(n)),$$

Exercise

Suppose that an arena $\mathcal{A}$ is colored with k different priorities. Show that the simple priority game on $\mathcal{A}$ can be solved by solving several (at most $k-1$) parity games and that both players have optimal positional strategies in simple priority games.

(This reduction will be valid also for infinite arenas.)

Mean-payoff Priority Games.

$\Re = \mathbb{N} \times \mathbb{R}$, $(n, r) \in \Re$, n — priority, r — reward.

$$x = (n_1, r_1), (n_2, r_2), \ldots \in \Re^\omega$$

$n = \mathrm{priority}(n_1 n_2 \ldots)$ priority of x,

$$x(n) = (n_{i_1}, r_{i_1}), (n_{i_2}, r_{i_2}), \ldots$$

where $n = n_{i_1} = n_{i_2} = \cdots$.

$$\boxed{\overline{\mathrm{mean}}(x) = \limsup_{k \to \infty} \frac{r_{i_1} + \cdots + r_{i_k}}{k} \ .}$$

Gambling Games.

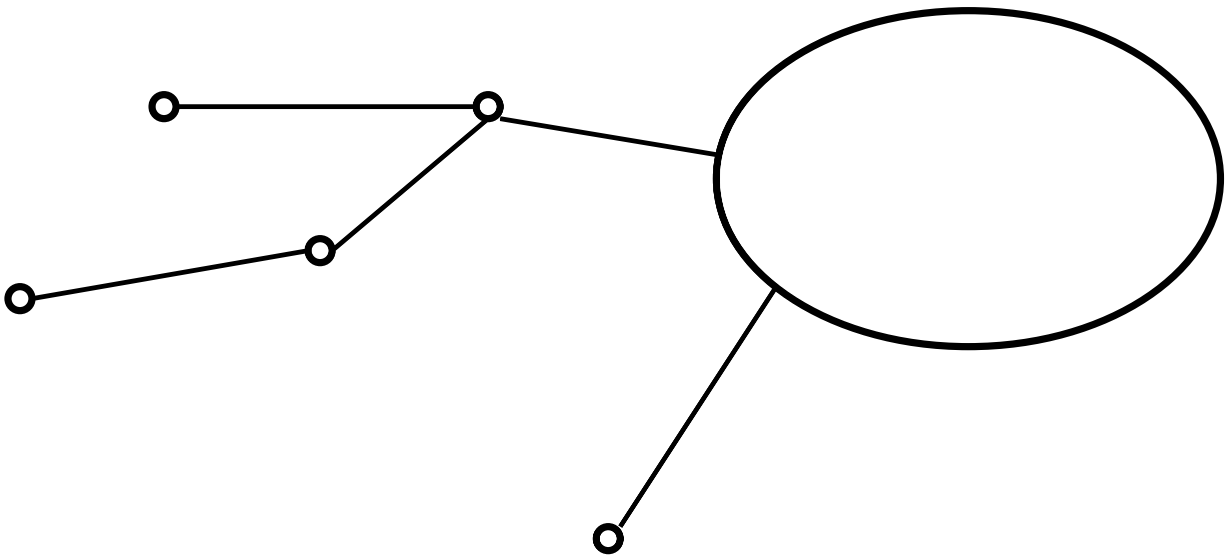
$\Re = \mathbb{R}$, $r_1 r_2 \ldots$

$$g_\lambda(r_1 r_2 \ldots) = (1-\lambda) \liminf_{i\to\infty} r_i + \lambda \limsup_{i\to\infty} r_i \quad .$$

Exercise

Show that parity games can be reduced to gabling gambling games with $\lambda = \frac{1}{2}$.

For all previous examples if all states are controlled by one player then this player has an optimal positional strategy.



Suppose that player 1 (maximizer) controls all states. For all games (except discounted) his optimal strategy for one-player game is the following:

- find the simple cycle in the arena $\mathcal{A}$ with the maximal payoff,

- go to this cycle and next go forever along this cycle.

Exercise

Show that the strategy described above is really optimal for one-player mean-payoff games.

Show the same for mean-payoff priority games and for gambling games.

From One-Player Games to Two-Player Games.

$\mathcal{A} = (S_1, S_2, A)$ is *controlled by player* i if $\forall s \in S_j$, $j \neq i$, $|A(s)| = 1$.

Theorem. *Fix*

- $\Re$ — *a set of rewards,*
- $\sqsubseteq$ — *a preference relation over* $\Re^\omega$.

Suppose that for each finite one-player arena $\mathcal{A}$ *the player controlling* $\mathcal{A}$ *has an optimal positional strategy in the game* $(\mathcal{A}, \sqsubseteq)$. *Then for all two-person games* $(\mathcal{A}, \sqsubseteq)$ *on finite arenas* $\mathcal{A}$ *both players have optimal positional strategies.*

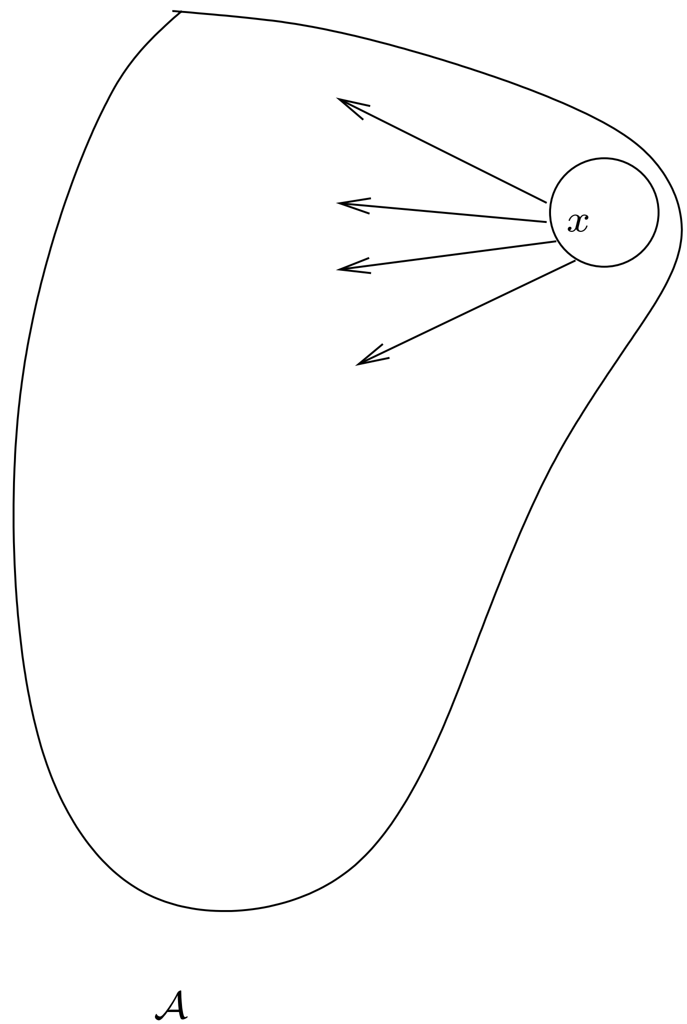
Proof $|A| - |S|$ the *rank* of $\mathcal{A}$.

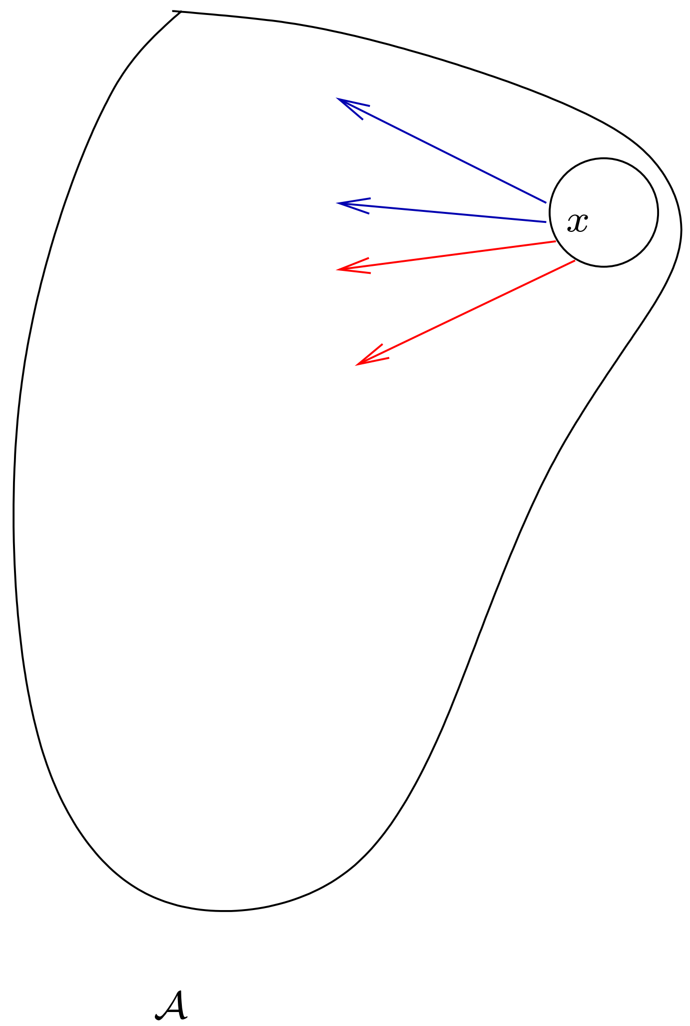
Proof: induction over the rank value.

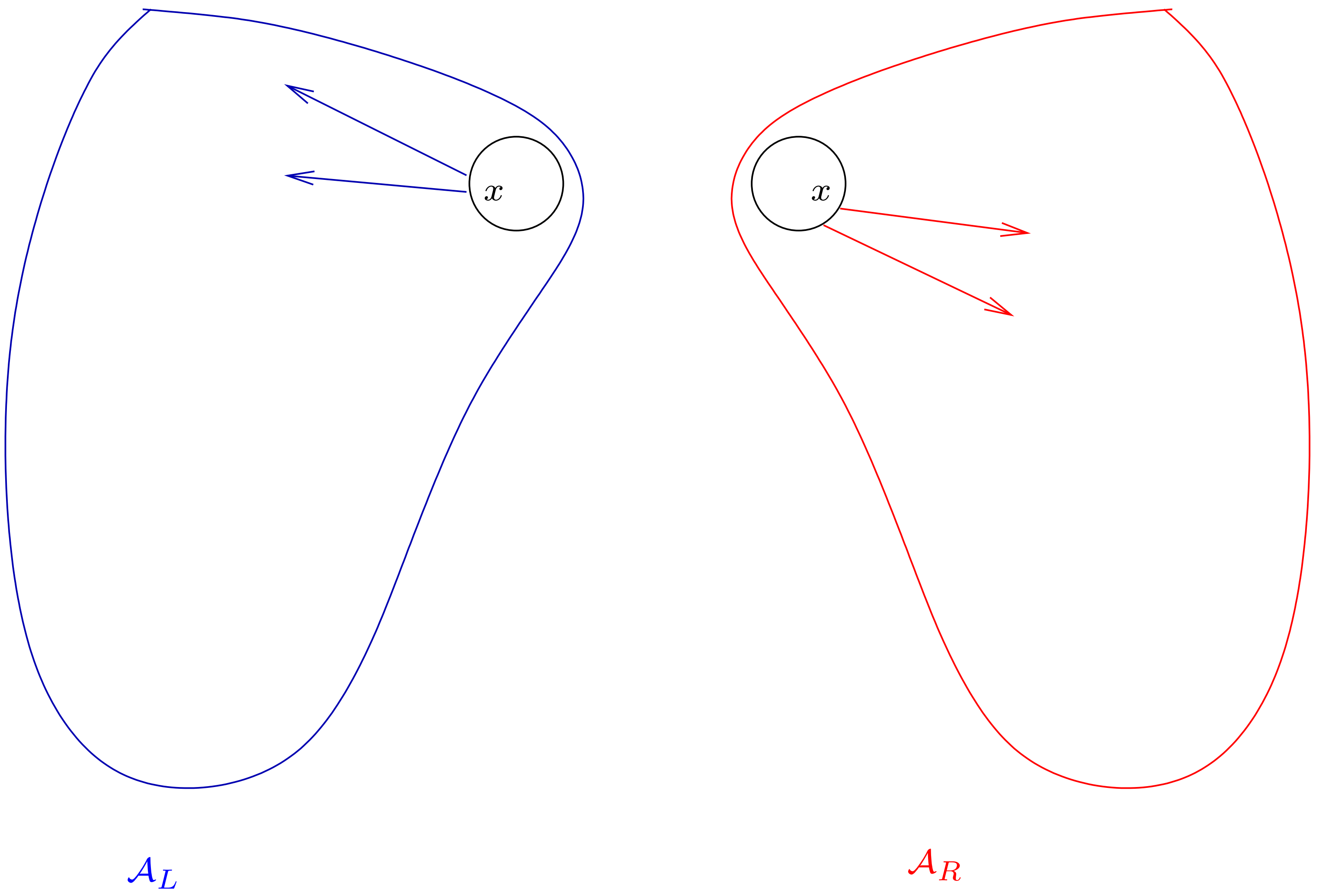
The pivot — a fixed state $x \in S_1$ such that $|A(x)| > 1$.

$$A(x) = A_L(x) \cup A_R(x)$$

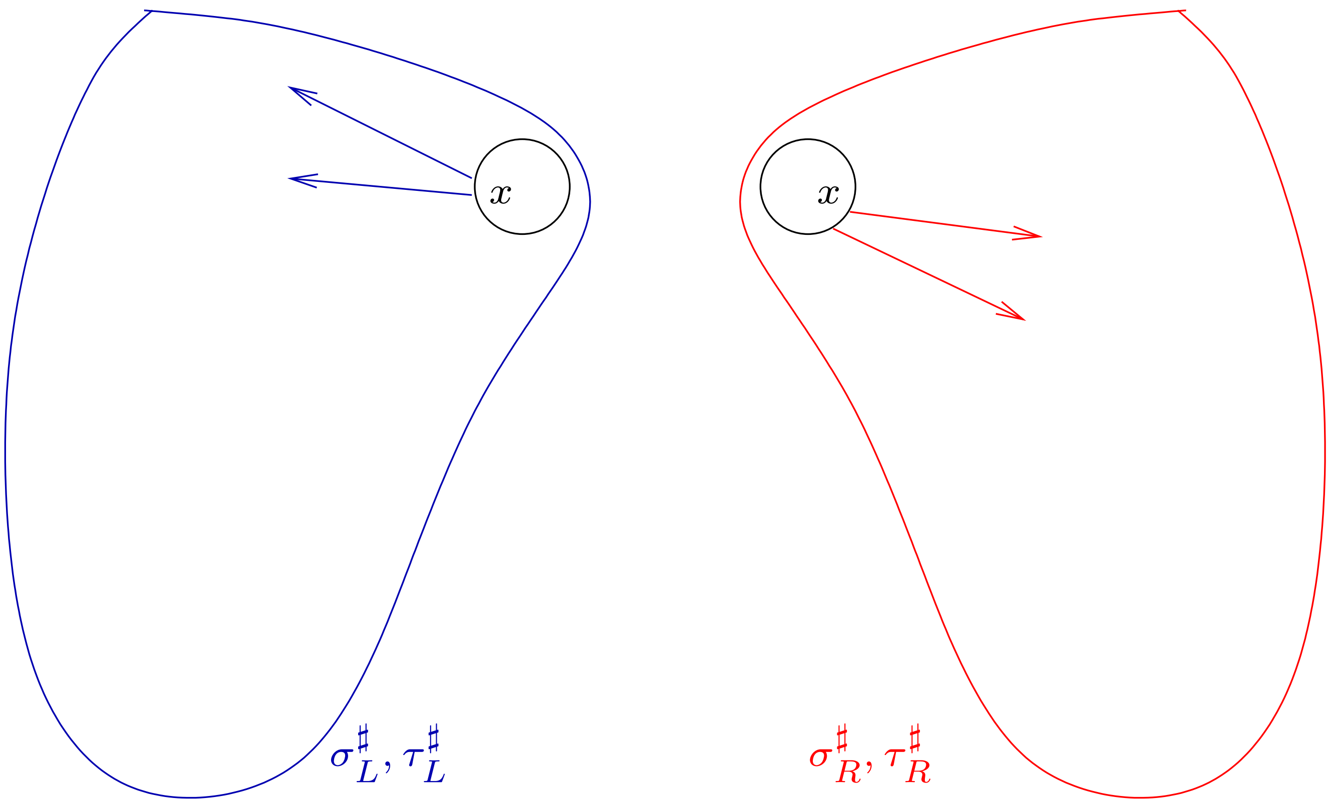
a non-trivial partition of $A(x)$ onto Left (bLue) actions and Right (Red) actions.


x
$\mathcal{A}$


x
$\mathcal{A}$



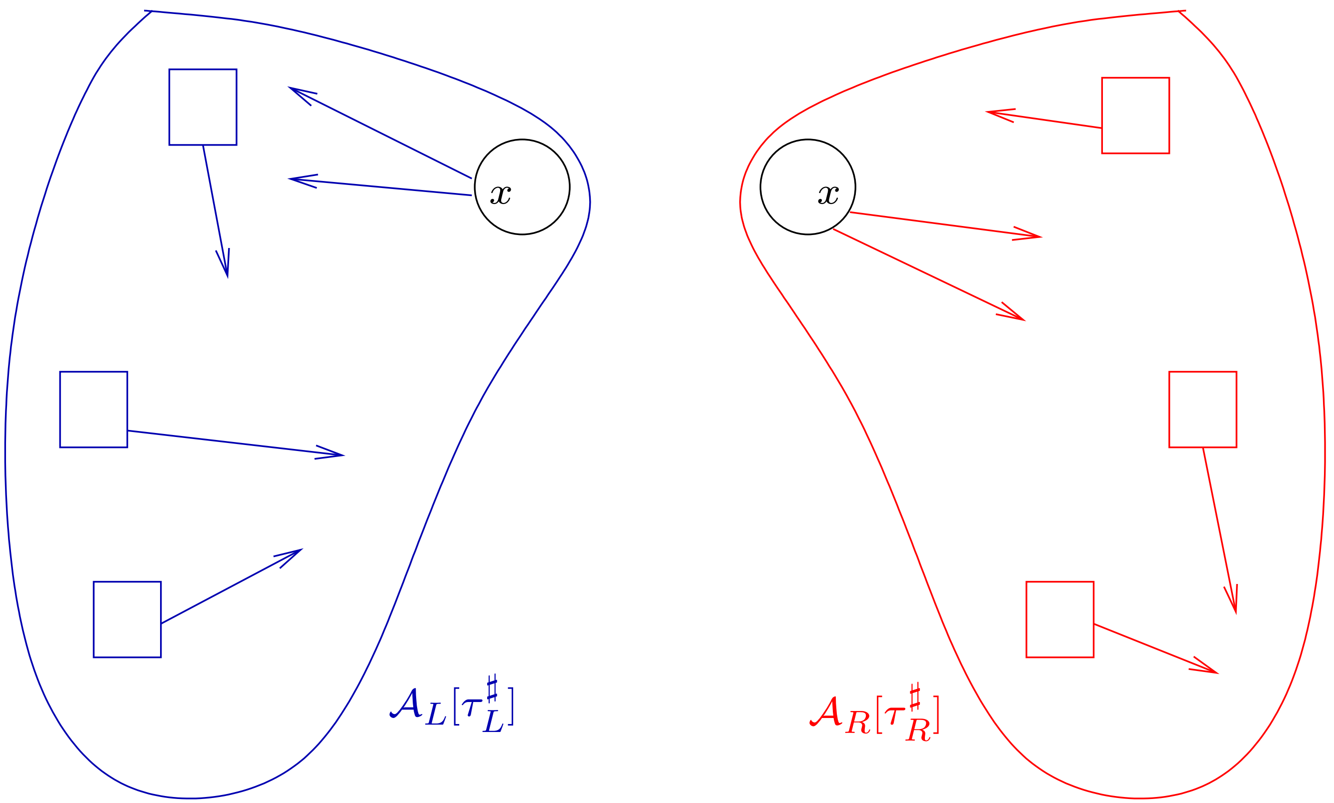
x
x
$\mathcal{A}_L$
$\mathcal{A}_R$



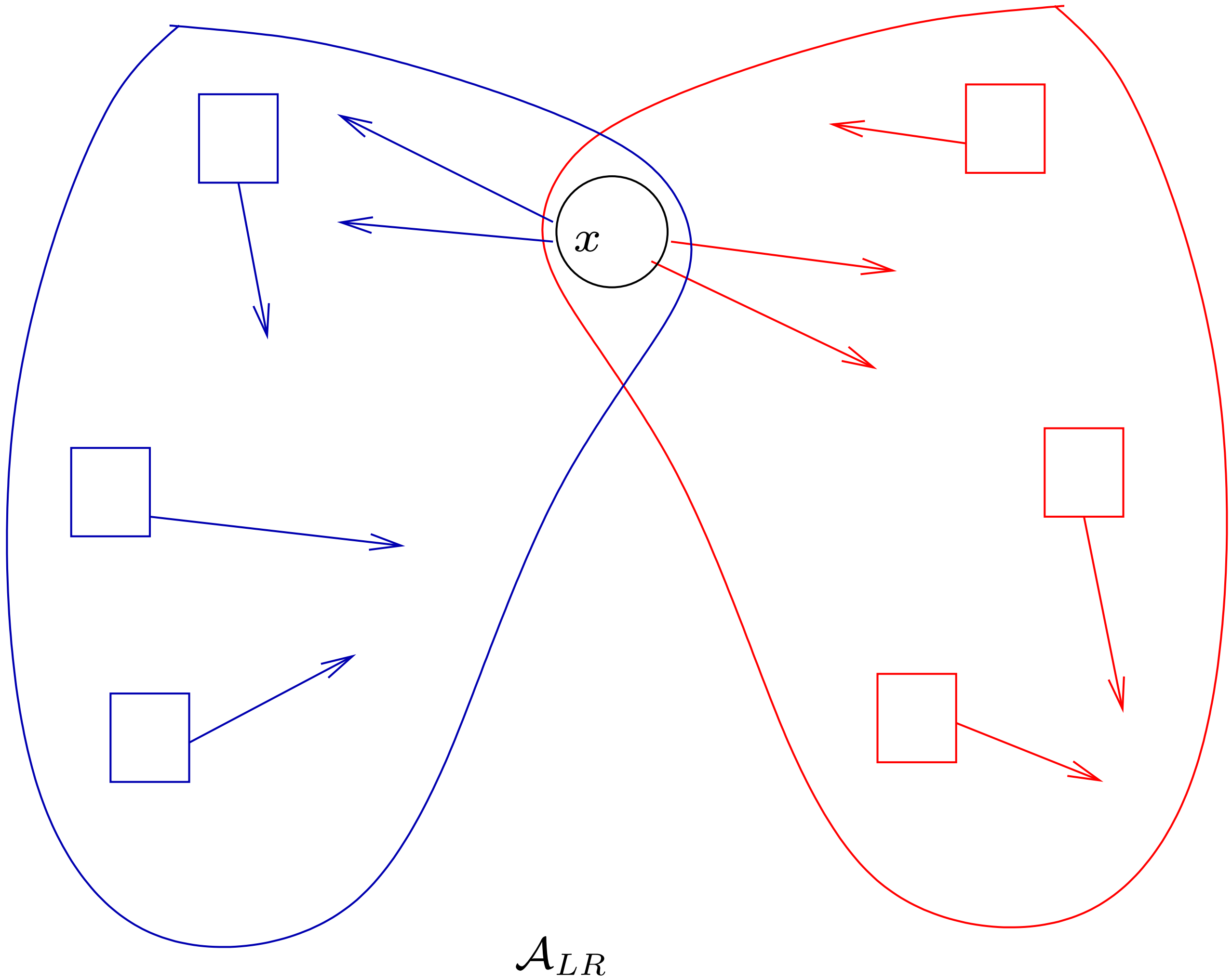
By induction there exist optimal positional strategies on $\mathcal{A}_L$ and $\mathcal{A}_R$

We show that one of the strategies $\sigma_L^\sharp$, $\sigma_R^\sharp$ is optimal for player 1 in the initial game on $\mathcal{A}$.

Usually neither $\tau_L^\sharp$ nor $\tau_R^\sharp$ is optimal on $\mathcal{A}$ but we show how to construct an optimal strategy for player 2 on $\mathcal{A}$ using these two strategies. The strategy for player 2 that we will construct will use one bit of memory to choose between $\tau_L^\sharp$ nor $\tau_R^\sharp$ depending on the last movement of player 1 at the pivot state.



Restrict the movements of player 2 by allowing only actions imposed by strategies $\tau_L^\sharp$ and $\tau_R^\sharp$



$\sigma^{\sharp}_{LR}$ optimal positional strategy of player 1 on $\mathcal{A}_{LR}$

Optimal strategy of player 1 in $\mathcal{A}$

$$\sigma^\sharp = \begin{cases} \sigma^\sharp_L & \text{if } \sigma^\sharp_{LR}(x) \text{ is a blue action} \\ \\ \sigma^\sharp_R & \text{if } \sigma^\sharp_{LR}(x) \text{ is a red action} \end{cases}$$

Our choice of the optimal strategy of player 1 in $\mathcal{A}$

We assume that:

$$\sigma^{\sharp}_{LR}(x) \in A_L(x)$$

and

$$\sigma^{\sharp} := \sigma^{\sharp}_{L} \quad .$$

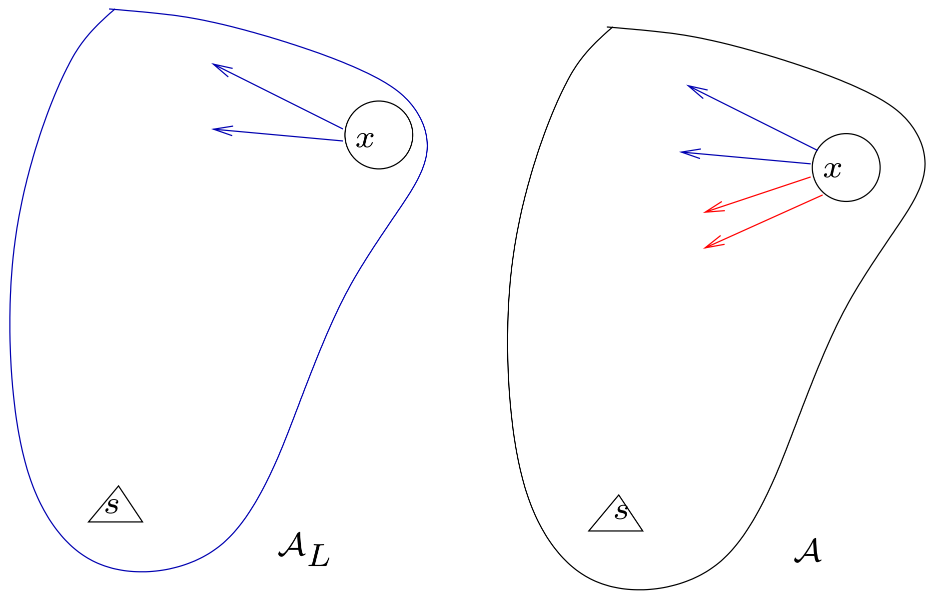
Optimal strategy for player 2 on $\mathcal{A}$

h be a finite history in $\mathcal{A}$ with $\mathrm{target}(h) \in S_2$.

$$\tau^\sharp(h) = \begin{cases} \tau^\sharp_L(\mathrm{target}(h)) & \text{if either } h \text{ does not contain any action with source } x \\ & \text{or the last such action belongs to } A_L(x), \\ \tau^\sharp_R(\mathrm{target}(h)) & \text{if } h \text{ contains at least one action with source } x \\ & \text{and the last such action belongs to } A_R(x). \end{cases}$$

We should prove

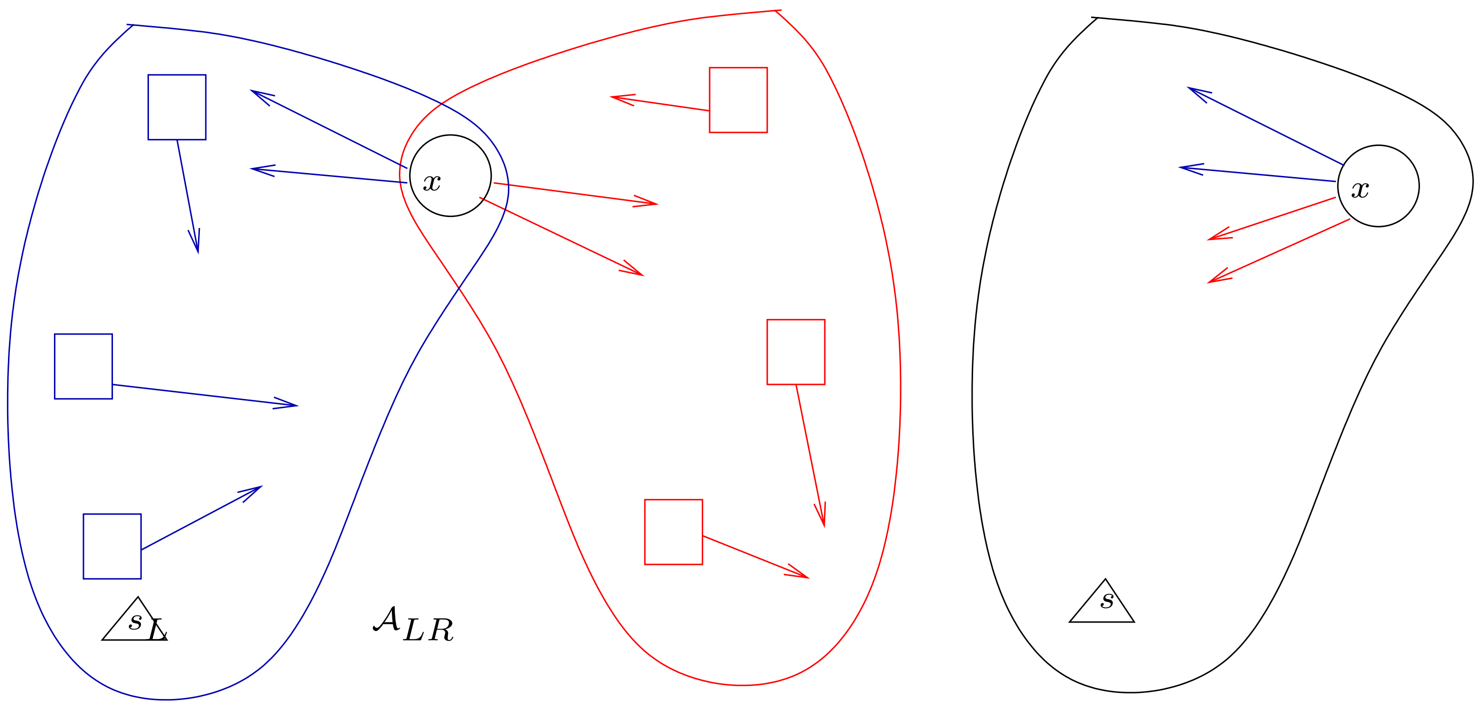
$$\mathrm{reward}(p_{\mathcal{A}}(s, \sigma, \tau^\sharp)) \sqsubseteq \mathrm{reward}(p_{\mathcal{A}}(s, \sigma^\sharp, \tau^\sharp)) \sqsubseteq \mathrm{reward}(p_{\mathcal{A}}(s, \sigma^\sharp, \tau)) \quad .$$



$$p_{\mathcal{A}}(s,\sigma^{\sharp},\tau^{\sharp}) = p_{\mathcal{A}}(s,\sigma^{\sharp}_{L},\tau^{\sharp}_{L}) =$$

$$p_{\mathcal{A}_L}(s,\sigma^{\sharp}_{L},\tau^{\sharp}_{L}) \sqsubseteq p_{\mathcal{A}_L}(s,\sigma^{\sharp}_{L},\tau) =$$

$$p_{\mathcal{A}}(s,\sigma^{\sharp}_{L},\tau) = p_{\mathcal{A}}(s,\sigma^{\sharp},\tau)$$



$$p_{\mathcal{A}}(s,\sigma,\tau^\sharp) = p_{\mathcal{A}_{LR}}(s_L,\sigma,\cdot) \sqsubseteq p_{\mathcal{A}_{LR}}(s_L,\sigma^\sharp_{LR},\cdot) = p_{\mathcal{A}_L}(s,\sigma^\sharp_{LR},\tau^\sharp_L) \sqsubseteq p_{\mathcal{A}_L}(s,\sigma^\sharp_L,\tau^\sharp_L) = p_{\mathcal{A}}(s,\sigma^\sharp,\tau^\sharp)$$

Exchangeability property for optimal strategies

$(\sigma^\sharp, \tau^\sharp)$ and $(\sigma^\ddagger, \tau^\ddagger)$ pairs of optimal strategies.

Then $(\sigma^\ddagger, \tau^\sharp)$ and $(\sigma^\sharp, \tau^\ddagger)$ are also optimal.

Moreover,

$$\mathrm{reward}(p_{\mathcal{A}}(s, \sigma^\ddagger, \tau^\ddagger)) \sqsubseteq \mathrm{reward}(p_{\mathcal{A}}(s, \sigma^\sharp, \tau^\sharp)) \sqsubseteq \mathrm{reward}(p_{\mathcal{A}}(s, \sigma^\ddagger, \tau^\ddagger)) \ .$$