A Simple P-matrix Linear Complementarity Problem for Discounted Games

Marcin Jurdziński   **Rahul Savani**

Department of Computer Science
University of Warwick

Motivation

- **Discounted games**
  - polytime reductions from **parity** and **mean-payoff** games
  - simple optimality equations give “transparent” reduction
- **P-matrix Linear Complementarity Problem**
  - well studied problem in mathematical programming
  - many algorithms known

Outline

- **P-matrix Linear Complementarity Problem**
  - **Unique sink orientations (USO) of cubes**

    combinatorial framework for strategy improvement algorithms
- **Discounted games**
  - Optimality equations characterize unique values
- **Reduction** from **discounted games** to **PLCP**
  - Connections between algorithms
- **Further research**

Linear Complementarity Problem (LCP)

Given: $q \in \mathbb{R}^n$, $M \in \mathbb{R}^{n\times n}$ Find: z, $w \in \mathbb{R}^n$ so that

$$z \geq 0 \quad \perp \quad w = q + Mz \geq 0$$

$\perp$ means orthogonal:

$$z^T w = 0$$

$$\Leftrightarrow \qquad z_i w_i = 0 \quad \text{all } i = 1, \ldots, n$$

Linear Complementarity Problem (LCP)

Given: $q \in \mathbb{R}^n$, $M \in \mathbb{R}^{n \times n}$ Find: $z, w \in \mathbb{R}^n$ so that

$$z \geq \mathbf{0} \quad \perp \quad w = q + Mz \geq \mathbf{0}$$

$\perp$ means orthogonal:

$$z^T w = 0$$

$$\Leftrightarrow \qquad z_i w_i = 0 \quad \text{all } i = 1, \ldots, n$$

If $q \geq \mathbf{0}$, the LCP has trivial solution $w = q$, $z = \mathbf{0}$.

LP in inequality form

primal : **max** $c^T x$

subject to $A x \leq b$

$x \geq 0$

dual : **min** $y^T b$

subject to $y^T A \geq c^T$

$y \geq 0$

LP in inequality form

primal : **max** $c^T x$

subject to $Ax \leq b$

$x \geq 0$

dual : **min** $y^T b$

subject to $y^T A \geq c^T$

$y \geq 0$

Weak duality: x, y feasible (fulfilling constraints)

$$\Rightarrow \quad c^T x \leq y^T A x \leq y^T b$$

LP in inequality form

primal : **max** $c^T x$

subject to $A x \leq b$

$x \geq 0$

dual : **min** $y^T b$

subject to $y^T A \geq c^T$

$y \geq 0$

Weak duality: x, y feasible (fulfilling constraints)

$$\Rightarrow \quad c^T x \leq y^T A x \leq y^T b$$

Strong duality: primal and dual feasible

$$\Rightarrow \exists \text{ feasible } x, y : \quad c^T x = y^T b \quad (x, y \text{ optimal})$$

LCP generalizes LP

LCP encodes **complementary slackness** of strong duality:

$$c^T x = \qquad y^T A x \qquad = y^T b$$

$$\Leftrightarrow \quad (y^T A - c^T) x = 0, \qquad y^T (b - A x) \qquad = 0.$$

$$\geq 0 \qquad \geq 0 \qquad\qquad \geq 0 \quad \geq 0$$

LCP generalizes LP

LCP encodes **complementary slackness** of strong duality:

$$
\begin{array}{llll}
 & c^T x = & y^T A x & = y^T b \\
\Leftrightarrow & (y^T A - c^T) x = 0, & y^T (b - A x) & = 0.
\end{array}
$$

$$\geq 0 \quad \geq 0 \qquad\qquad \geq 0 \quad \geq 0$$

LP $\Leftrightarrow$ LCP

$$
\underbrace{\begin{pmatrix} x \\ y \end{pmatrix}}_{z} \geq 0 \quad \perp \quad \underbrace{\begin{pmatrix} -c \\ b \end{pmatrix}}_{q} + \underbrace{\begin{pmatrix} 0 & A^T \\ -A & 0 \end{pmatrix}}_{M} \underbrace{\begin{pmatrix} x \\ y \end{pmatrix}}_{z} \geq 0
$$

LCPs and complementary cones

Given: $q \in \mathbb{R}^n$, $M \in \mathbb{R}^{n \times n}$ Find: $z \in \mathbb{R}^n$ so that

$$z \geq 0 \perp w = q + Mz \geq 0$$

LCPs and complementary cones

Given: $q \in \mathbb{R}^n$, $M \in \mathbb{R}^{n \times n}$ Find: $z \in \mathbb{R}^n$ so that

$$z \geq 0 \perp w = q + Mz \geq 0$$

$$\Leftrightarrow \quad z \geq 0 \perp w \geq 0 \quad \boxed{q = Iw - Mz}$$

LCPs and complementary cones

Given: $q \in \mathbb{R}^n$, $M \in \mathbb{R}^{n \times n}$ Find: $z \in \mathbb{R}^n$ so that

$$z \geq 0 \perp w = q + Mz \geq 0$$

$$\Leftrightarrow \quad z \geq 0 \perp w \geq 0 \quad \boxed{q = Iw - Mz}$$

$\Leftrightarrow$ q belongs to a **complementary cone:**

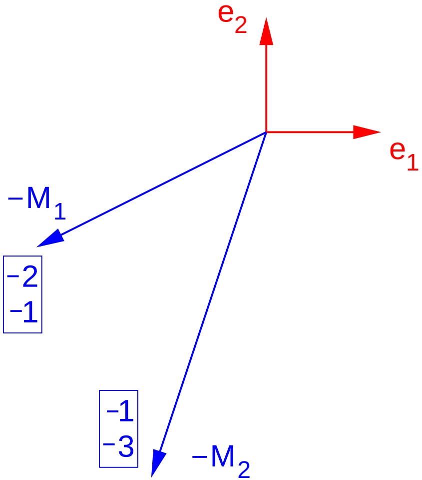
$$\boxed{q \in C(\alpha) = cone\,\{-M_i, e_j \mid i \in \alpha, j \notin \alpha\}}$$

for some $\alpha \subseteq \{1, \ldots, n\}$, $M = [M_1 M_2 \cdots M_n]$

$\alpha = \{i \mid z_i > 0\}$

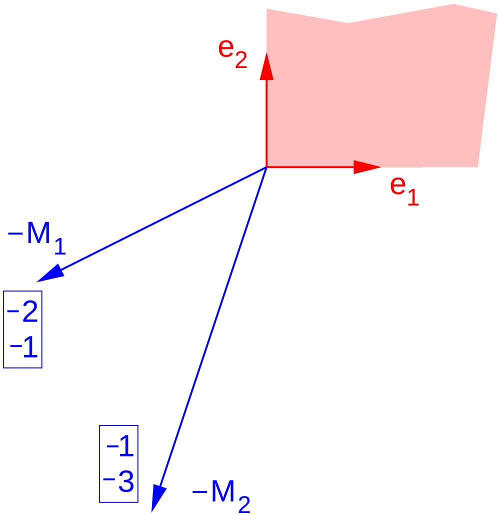
LCPs and complementary cones

$$M = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$$



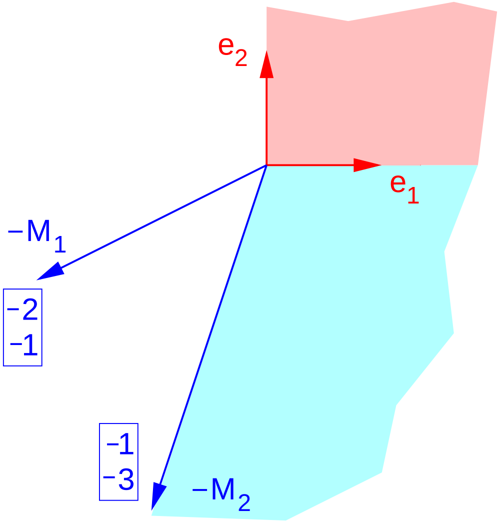
LCPs and complementary cones

$$M = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$$



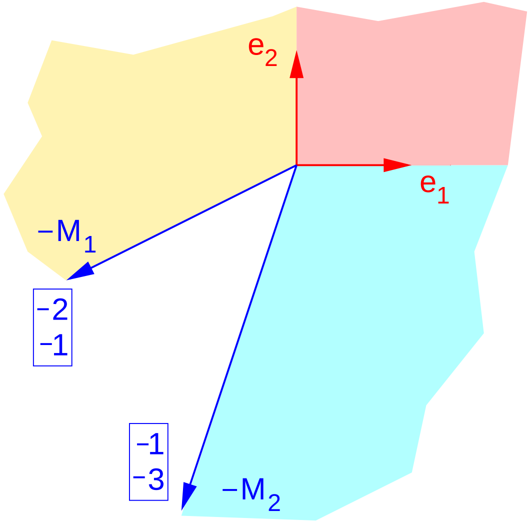
LCPs and complementary cones

$$M = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$$



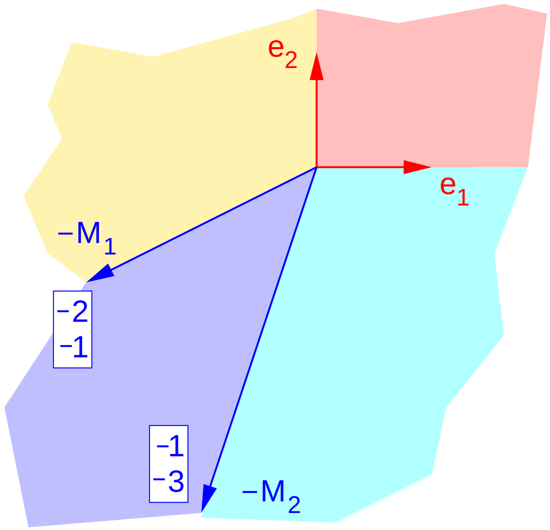
LCPs and complementary cones

$$M = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$$



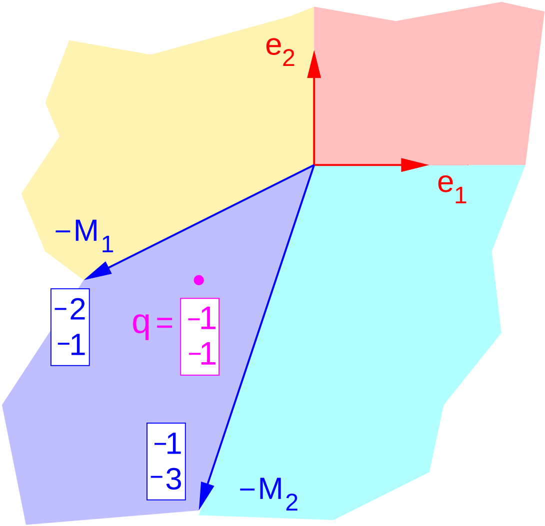
LCPs and complementary cones

$$M = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$$



LCPs and complementary cones

$$M = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$$



P-matrices

Def: $M \in \mathbb{R}^{n \times n}$ is a **P-matrix** if **all** its **principal minors** are **positive**.

Thm: M is a **P-matrix** $\Leftrightarrow$ LCP (M, q) has **unique solution** $\forall q \in \mathbb{R}^n$.

P-matrices

Def: $M \in \mathbb{R}^{n \times n}$ is a **P-matrix** if **all** its **principal minors** are **positive**.

Thm: M is a **P-matrix** $\Leftrightarrow$ LCP (M, q) has **unique solution** $\forall q \in \mathbb{R}^n$.

Example

$$M = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} \qquad M' = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$$

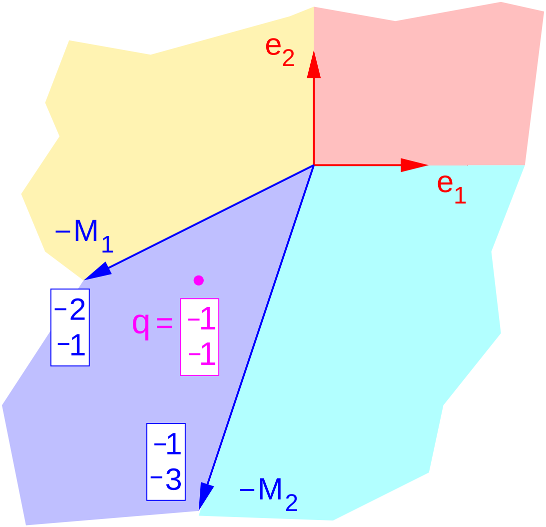
M is a P-matrix, as

$$\begin{aligned} det(M_{11}) &= 2 > 0 \\ det(M_{22}) &= 3 > 0 \\ det(M) &= 5 > 0 \end{aligned}$$

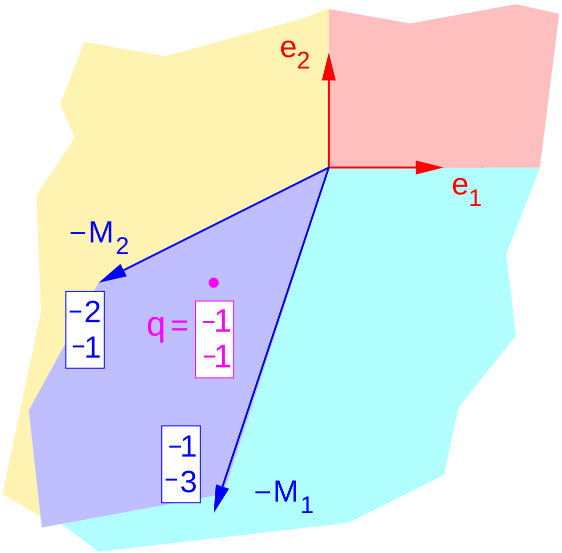
M' is not a P-matrix, as $det(M') = -5 < 0$

Complementary cones: P-matrix

$$M = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$$



Multiple solutions



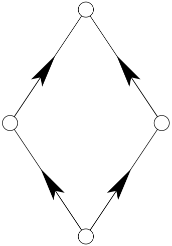
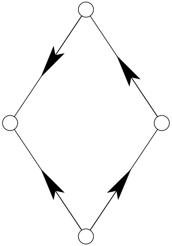
Unique sink orientations of cubes

[Szabó and Welzl (2001)] [Stickney and Watson (1978)]

=

- ***n***-dimensional hypercube
- edges oriented such that **every face** has a **unique sink**

The two USOs for $n = 2$:

USO for P-matrix LCP

LCP: $z \geq 0 \perp w \geq 0$, $\boxed{q = Iw - Mz}$

For every $\alpha \subseteq \{1, \ldots, n\}$, define $B^{\alpha} \in \mathbb{R}^{n \times n}$ by

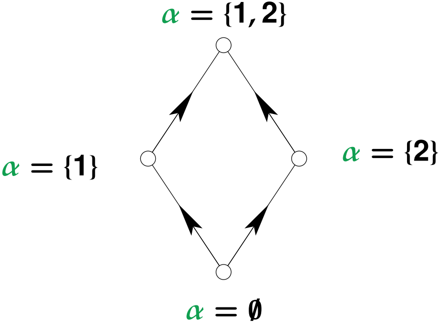
$$(B^{\alpha})_i = \begin{cases} -M_i, & i \in \alpha \\ e_i, & i \notin \alpha \end{cases}$$

Orient edges at vertex α oriented according to

$$sign\left((B^{\alpha})^{-1} q \right)$$

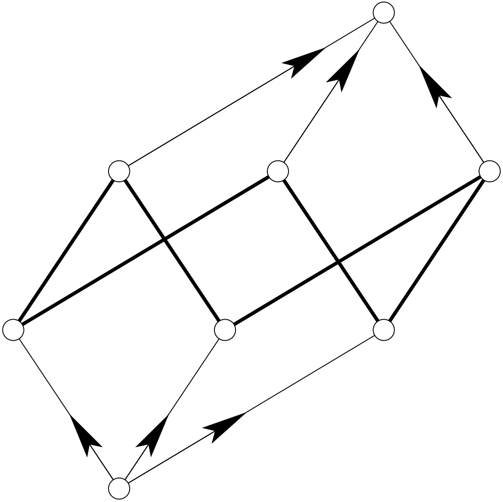
PLCP USO example

$$-1/5\begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix} z' + Iw' = q' = \begin{pmatrix} 2/5 \\ 1/5 \end{pmatrix} \geq 0$$



$$Iw - Mz = Iw - \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} z = q = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

Cyclic USO



Cyclic USOs can arise from P-matrix LCPs.

Murty's Least Index Method

Input: LCP(M, q) with P-matrix M.

Output: solution (w^*, z^*)

Start with $\alpha := \emptyset$, $\bar{q} := q$.

while $\bar{q} \not\geq 0$ do:

- Let $s = \min_{\{1,\dots,n\}}\{i \mid \bar{q}_i < 0\}$,
- Set $\alpha \leftarrow \alpha \oplus s$ and $\bar{q} \leftarrow (B^{\alpha})^{-1} q$.

Why does it work?

when $\bar{q}_1, \dots, \bar{q}_{n-1} \geq 0$, we determine $w_n^* = 0$ or $z_n^* = 0$.

Binary zero-sum discounted games

- Finite directed graph on states $S = \{1, \dots, n\}$
- Partition $S = S_{Max} \cup S_{Min}$

Binary zero-sum discounted games

- Finite directed graph on states $S = \{1, \ldots, n\}$
- Partition $S = S_{Max} \cup S_{Min}$
- Every state has a left successor $\lambda(s)$ and right successor $\rho(s)$

Binary zero-sum discounted games

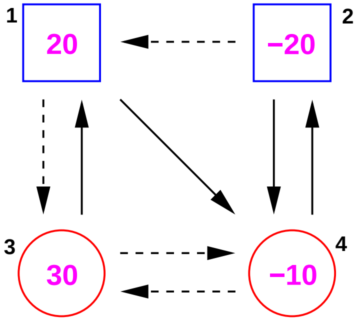
- Finite directed graph on states $S = \{1, \ldots, n\}$
- Partition $S = S_{Max} \cup S_{Min}$
- Every state has a left successor $\lambda(s)$ and right successor $\rho(s)$
- Every state has a reward - $r : S \mapsto \mathbb{Z}$

Binary zero-sum discounted games

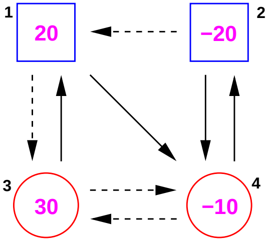
- Finite directed graph on states $S = \{1, \ldots, n\}$
- Partition $S = S_{Max} \cup S_{Min}$
- Every state has a left successor $\lambda(s)$ and right successor $\rho(s)$
- Every state has a reward - $r : S \mapsto \mathbb{Z}$
- Discount factor $\delta \in (0,1)$ (same for both players)

Binary zero-sum discounted games

- Finite directed graph on states $S = \{1, \ldots, n\}$
- Partition $S = S_{Max} \cup S_{Min}$
- Every state has a left successor $\lambda(s)$ and right successor $\rho(s)$
- Every state has a reward - $r : S \mapsto \mathbb{Z}$
- Discount factor $\delta \in (0,1)$ (same for both players)

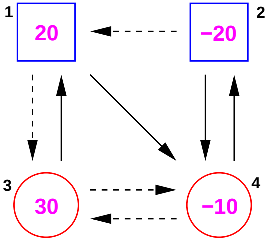


Player objectives



- A play is an infinite path $\pi = s_0, s_1, s_3, \ldots$
  - initial state s_0
  - owner of s_i chooses $s_{i+1} \in \{\ \lambda(s_i),\ \rho(s_i)\ \}$

Player objectives



- A play is an infinite path $\pi = s_0, s_1, s_3, \ldots$
  - initial state s_0
  - owner of s_i chooses $s_{i+1} \in \{ \lambda(s_i),\ \rho(s_i) \}$
- ***Max*** maximizes and ***Min*** minimizes

$$\sum_{i=0}^{\infty} \delta^i r(s_i)$$

Optimality equations

- Every state has a **value** $v(s)$ characterized by:

$$\forall s \in S_{Max} : \quad v(s) = \max_{t \in \{\lambda(s), \rho(s)\}} (r(s) + \delta v(t))$$

$$\forall s \in S_{Min} : \quad v(s) = \min_{t \in \{\lambda(s), \rho(s)\}} (r(s) + \delta v(t))$$

Optimality equations

- Every state has a **value** $v(s)$ characterized by:

$$\forall s \in S_{Max} : \quad v(s) = \max_{t \in \{\lambda(s), \rho(s)\}} (r(s) + \delta v(t))$$

$$\forall s \in S_{Min} : \quad v(s) = \min_{t \in \{\lambda(s), \rho(s)\}} (r(s) + \delta v(t))$$

- Proofs:
  - **Banach fixed point theorem** for **contraction** mappings
  - **Strategy improvement** algorithm (constructive)

Optimality equations

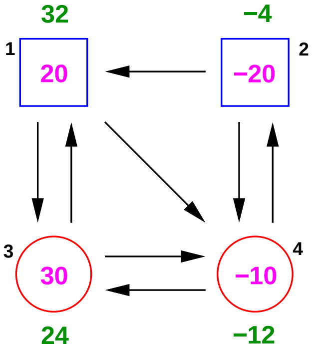
- Every state has a **value $v(s)$** characterized by:

$$\forall s \in S_{Max} : \quad v(s) = \max_{t \in \{\lambda(s), \rho(s)\}} (r(s) + \delta v(t))$$

$$\forall s \in S_{Min} : \quad v(s) = \min_{t \in \{\lambda(s), \rho(s)\}} (r(s) + \delta v(t))$$

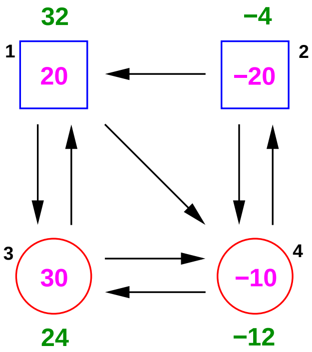
- Proofs:
  - **Banach fixed point theorem** for **contraction** mappings
  - **Strategy improvement** algorithm (constructive)
- Values give *pure* and *positional* optimal strategies:
  Max (***Min***) picks succesor with largest (smallest) value.

Unique values for $\delta = 1/2$



$v(1) = 32 \quad = r(1) + \delta \max(v(3), v(4)) = \quad 20{+}1/2(24)$

Unique values for $\delta = 1/2$



$$v(1) = 32 \quad = r(1) + \delta \max(v(3), v(4)) = \quad 20 + 1/2(24)$$

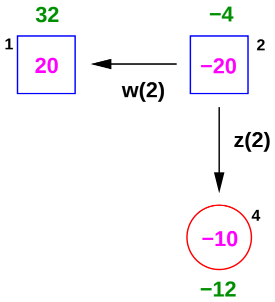
$$v(2) = -4 \quad = r(2) + \delta \max(v(1), v(4)) = \quad -20 + 1/2(32)$$

$$v(3) = 24 \quad = r(3) + \delta \min(v(1), v(4)) = \quad 30 + 1/2(-12)$$

$$v(4) = -12 \quad = r(4) + \delta \min(v(2), v(3)) = \quad -10 + 1/2(-4)$$

Nonnegative slacks and complementarity

$$v(2) = r(2) + \delta \max(v(1), v(4))$$



$$v(2) = w(2) + r(2) + \delta v(1)$$
$$v(2) = z(2) + r(2) + \delta v(4)$$

$$w(2), z(2) \geq 0, \quad w(2) \cdot z(2) = 0$$

Reduction to LCP

$$\forall s \in S_{Max} : \quad v(s) = \max_{t \in \{\lambda(s), \rho(s)\}} (r(s) + \delta v(t))$$

Replace **max/min** with **slacks** and **complementarity condition**

Reduction to LCP

$$\forall s \in S_{Max} : \quad v(s) = \max_{t \in \{\lambda(s), \rho(s)\}} (r(s) + \delta v(t))$$

Replace **max/min** with **slacks** and **complementarity condition**

$$\forall s \in S_{Max} : \quad v(s) = w(s) + r(s) + \delta v(\lambda(s))$$

$$v(s) = z(s) + r(s) + \delta v(\rho(s))$$

$$\forall s \in S : \quad w(s) \geq 0 \perp z(s) \geq 0$$

Reduction to LCP

$$\forall s \in S_{Max}: \quad v(s) = \max_{t \in \{\lambda(s), \rho(s)\}} (r(s) + \delta v(t))$$

$$\forall s \in S_{Min}: \quad v(s) = \min_{t \in \{\lambda(s), \rho(s)\}} (r(s) + \delta v(t))$$

Replace **max/min** with **slacks** and **complementarity condition**

$$\forall s \in S_{Max}: \quad v(s) = w(s) + r(s) + \delta v(\lambda(s))$$

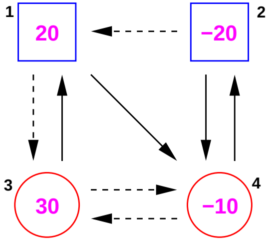
$$v(s) = z(s) + r(s) + \delta v(\rho(s))$$

$$\forall s \in S_{Min}: \quad v(s) = -w(s) + r(s) + \delta v(\lambda(s))$$

$$v(s) = -z(s) + r(s) + \delta v(\rho(s))$$

$$\forall s \in S: \quad w(s) \geq 0 \perp z(s) \geq 0$$

Example



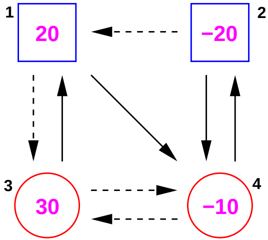
$\forall s \in S :$

$w(v) \geq 0 \perp z(v) \geq 0$

$$\begin{pmatrix} v(1) \\ v(2) \\ -v(3) \\ -v(4) \end{pmatrix} = \begin{pmatrix} w(1) \\ w(2) \\ w(3) \\ w(4) \end{pmatrix} + \begin{pmatrix} r(1) \\ r(2) \\ -r(3) \\ -r(4) \end{pmatrix} + \delta \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} v(1) \\ v(2) \\ v(3) \\ v(4) \end{pmatrix}$$

$$\begin{pmatrix} v(1) \\ v(2) \\ -v(3) \\ -v(4) \end{pmatrix} = \begin{pmatrix} z(1) \\ z(2) \\ z(3) \\ z(4) \end{pmatrix} + \begin{pmatrix} r(1) \\ r(2) \\ -r(3) \\ -r(4) \end{pmatrix} + \delta \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} v(1) \\ v(2) \\ v(3) \\ v(4) \end{pmatrix}$$

Example



$$w \geq 0 \perp z \geq 0$$

$$A := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$Av = w + Ar + \delta A \overbrace{\begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}}^{L} v$$

$$Av = z + Ar + \delta A \underbrace{\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}}_{R} v$$

Eliminate v

$$A(I - \delta L)v = w + Ar$$
$$A(I - \delta R)v = z + Ar$$

Eliminating v we get

$$w + Ar = A(I - \delta L)(A(I - \delta R))^{-1}(z + Ar)$$

$$\boxed{w = Mz + q}$$

$$\boxed{w \geq 0 \perp z \geq 0}$$

$$M = A(I - \delta L)(I - \delta R)^{-1}A, \quad q = (M - I)Ar$$

Example

$$\boxed{w = Mz + q}$$

$$\boxed{w \geq 0 \perp z \geq 0}$$

$$M = A(I - \delta L)(I - \delta R)^{-1}A, \quad q = (M - I)Ar$$

$$A(I - \delta L) = \begin{pmatrix} 1 & 0 & -\delta & 0 \\ -\delta & 1 & 0 & 0 \\ 0 & 0 & -1 & \delta \\ 0 & 0 & \delta & -1 \end{pmatrix} \quad A(I - \delta R) = \begin{pmatrix} 1 & 0 & 0 & -\delta \\ 0 & 1 & 0 & -\delta \\ \delta & 0 & -1 & 0 \\ 0 & \delta & 0 & -1 \end{pmatrix}$$

Theorem [Levy-Desplanques]

If $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ *is strictly diagonally dominant, i.e.,* $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$ *for all* i*, then* $\boldsymbol{A}$ *is non-singular.*

Theorem [Levy-Desplanques]

If $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is strictly diagonally dominant, i.e., $|a_{ii}| > \sum_{j\neq i} |a_{ij}|$ for all i, then $\boldsymbol{A}$ is non-singular.

- $\boldsymbol{A}(\boldsymbol{I}-\delta\boldsymbol{L})$ and $\boldsymbol{A}(\boldsymbol{I}-\delta\boldsymbol{R})$ are strictly diagonally dominant. E.g.

$$\boldsymbol{A}(\boldsymbol{I}-\delta\boldsymbol{L}) = \begin{pmatrix} 1 & 0 & -\delta & 0 \\ -\delta & 1 & 0 & 0 \\ 0 & 0 & -1 & \delta \\ 0 & 0 & \delta & -1 \end{pmatrix} \quad \boldsymbol{A}(\boldsymbol{I}-\delta\boldsymbol{R}) = \begin{pmatrix} 1 & 0 & 0 & -\delta \\ 0 & 1 & 0 & -\delta \\ \delta & 0 & -1 & 0 \\ 0 & \delta & 0 & -1 \end{pmatrix}$$

- So $\boldsymbol{M} = \boldsymbol{A}(\boldsymbol{I}-\delta\boldsymbol{L})(\boldsymbol{I}-\delta\boldsymbol{R})^{-1}\boldsymbol{A}$ is well defined

Theorem [Johnson and Tsatsomeros (1995)]

Let $\boldsymbol{M} = \boldsymbol{B}\boldsymbol{C}^{-1}$*, where* $\boldsymbol{B}, \boldsymbol{C} \in \mathbb{R}^{n \times n}$*. Then,* $\boldsymbol{M}$ *is a P-matrix if* $\boldsymbol{T}\boldsymbol{C} + (\boldsymbol{I} - \boldsymbol{T})\boldsymbol{B}$ *is invertible for all* $\boldsymbol{T} \in [\boldsymbol{0}, \boldsymbol{I}]$.

Theorem [Johnson and Tsatsomeros (1995)]

Let $M = BC^{-1}$, where $B, C \in \mathbb{R}^{n \times n}$. Then, M is a P-matrix if $TC + (I - T)B$ is invertible for all $T \in [0, I]$.

$$w = Mz + q$$

$$w \geq 0 \perp z \geq 0$$

$$M = A(I - \delta L)(I - \delta R)^{-1} A, \quad q = (M - I)Ar$$

$B = A(I - \delta L)$ and $C = A(I - \delta R)$ are strictly diagonally dominant.

Thus, $TC + (I - T)B$ is s.d.d., and hence invertible, for all $T \in [0, I]$.

Thus, $M = BC^{-1}$ is a P-matrix.

Discounted game PLCP: What is q?

Considering just the right successors $\rho(s)$ we have

$$\begin{aligned} v^\rho &= r + \delta v^\rho(\rho(s)) \\ (I - \delta R) v^\rho &= r \\ v^\rho &= (I - \delta R)^{-1} r \end{aligned}$$

and rewriting q in terms of v^ρ

$$\begin{aligned} q &= (M - I)Ar \\ &= A\delta(R - L)v^\rho \end{aligned}$$

$$q_s = \begin{cases} +\delta[v^\rho(\rho(s)) - v^\rho(\lambda(s))], & s \in S_{Max} \\ -\delta[v^\rho(\rho(s)) - v^\rho(\lambda(s))], & s \in S_{Min} \end{cases}$$

Strategy improvement algorithm

Definition

State $s \in S_{Max}$ is switchable under strategy pair defined by ρ if

$$v^{\rho}(\lambda(s)) > v^{\rho}(\rho(s))$$

Algorithm [Strategy Improvement for *Max*]

*Start with ρ defined by arbitrary strategy **Max** and **Min**'s best response.*

loop:

Obtain ρ' from ρ by changing at all switchable $s \in S_{Max}$ under v^{ρ}.

*Obtain ρ'' from ρ' so **Min** plays best response.*

if $\rho'' \neq \rho'$ **repeat** *with* $\rho \leftarrow \rho''$.

Why does strategy improvement work?

Theorem [Global improvement from myopic improvement]

Let ρ and ρ' be the strategy pairs before and after some iteration. Then we have

$$\forall s \in S : v^{\rho'}(s) \geq v^{\rho}(s) ,$$

and

$$\exists s \in S : v^{\rho'}(s) > v^{\rho}(s) .$$

Finite number of pure positional **strategies** so algorithm terminates.

Global improvement

- Want: $\Delta = v^{\rho'}(s) - v^{\rho}(s) \geq 0$ for all s and $\Delta > 0$ for some s

Global improvement

- Want: $\boldsymbol{\Delta} = v^{\rho'}(s) - v^{\rho}(s) \geq 0$ for all s and $\boldsymbol{\Delta} > 0$ for some s
- Recall $v^{\rho} = r + \delta R v^{\rho}$.

Global improvement

- Want: $\boldsymbol{\Delta} = v^{\rho'}(s) - v^{\rho}(s) \geq 0$ for all s and $\boldsymbol{\Delta} > 0$ for some s
- Recall $v^{\rho} = r + \delta R v^{\rho}$.

$$\boldsymbol{\Delta} = \delta\left(R' v^{\rho'} - R v^{\rho}\right)$$

Global improvement

- Want: $\boldsymbol{\Delta} = \boldsymbol{v}^{\rho'}(\boldsymbol{s}) - \boldsymbol{v}^{\rho}(\boldsymbol{s}) \geq \mathbf{0}$ for all $\boldsymbol{s}$ and $\boldsymbol{\Delta} > \mathbf{0}$ for some $\boldsymbol{s}$
- Recall $\boldsymbol{v}^{\rho} = \boldsymbol{r} + \delta \boldsymbol{R} \boldsymbol{v}^{\rho}$.

$$\begin{aligned}\boldsymbol{\Delta} &= \delta\left(\boldsymbol{R}' \boldsymbol{v}^{\rho'} - \boldsymbol{R} \boldsymbol{v}^{\rho}\right) \\ &= \delta\left(\boldsymbol{R}' \boldsymbol{v}^{\rho'} - \boldsymbol{R}' \boldsymbol{v}^{\rho} + \boldsymbol{R}' \boldsymbol{v}^{\rho} - \boldsymbol{R} \boldsymbol{v}^{\rho}\right)\end{aligned}$$

Global improvement

- Want: $\boldsymbol{\Delta} = v^{\rho'}(s) - v^{\rho}(s) \geq 0$ for all s and $\boldsymbol{\Delta} > 0$ for some s
- Recall $v^{\rho} = r + \delta R v^{\rho}$.

$$\begin{aligned}
\boldsymbol{\Delta} &= \delta\left(R' v^{\rho'} - R v^{\rho}\right) \\
&= \delta\left(R' v^{\rho'} - R' v^{\rho} + R' v^{\rho} - R v^{\rho}\right) \\
&= \delta\left(\underbrace{R' v^{\rho'} - R' v^{\rho}}_{R'\boldsymbol{\Delta}} + \underbrace{R' v^{\rho} - R v^{\rho}}_{d}\right)
\end{aligned}$$

Global improvement

- Want: $\Delta = v^{\rho'}(s) - v^{\rho}(s) \geq 0$ for all s and $\Delta > 0$ for some s
- Recall $v^{\rho} = r + \delta R v^{\rho}$.

$$\begin{aligned}
\Delta &= \delta\left(R' v^{\rho'} - R v^{\rho}\right) \\
&= \delta\left(R' v^{\rho'} - R' v^{\rho} + R' v^{\rho} - R v^{\rho}\right) \\
&= \delta\left(\underbrace{R' v^{\rho'} - R' v^{\rho}}_{R'\Delta} + \underbrace{R' v^{\rho} - R v^{\rho}}_{d}\right)
\end{aligned}$$

$$\Delta = (I - \delta R')^{-1} d$$

Global improvement

- Want: $\boldsymbol{\Delta} = v^{\rho'}(s) - v^{\rho}(s) \geq \mathbf{0}$ for all s and $\boldsymbol{\Delta} > \mathbf{0}$ for some s
- Recall $v^{\rho} = r + \delta R v^{\rho}$.

$$\begin{aligned}
\boldsymbol{\Delta} &= \delta\left(R' v^{\rho'} - R v^{\rho}\right) \\
&= \delta\left(R' v^{\rho'} - R' v^{\rho} + R' v^{\rho} - R v^{\rho}\right) \\
&= \delta\left(\underbrace{R' v^{\rho'} - R' v^{\rho}}_{R'\boldsymbol{\Delta}} + \underbrace{R' v^{\rho} - R v^{\rho}}_{d}\right)
\end{aligned}$$

$$\boldsymbol{\Delta} = (I - \delta R')^{-1} d$$

- $(I - \delta R')^{-1} \geq \mathbf{0}$ and $diag(I - \delta R')^{-1} > \mathbf{0}$.

Global improvement

- Want: $\Delta = v^{\rho'}(s) - v^{\rho}(s) \geq 0$ for all s and $\Delta > 0$ for some s
- Recall $v^{\rho} = r + \delta R v^{\rho}$.

$$
\begin{aligned}
\Delta &= \delta\left(R' v^{\rho'} - R v^{\rho}\right) \\
&= \delta\left(R' v^{\rho'} - R' v^{\rho} + R' v^{\rho} - R v^{\rho}\right) \\
&= \delta\left(\underbrace{R' v^{\rho'} - R' v^{\rho}}_{R'\Delta} + \underbrace{R' v^{\rho} - R v^{\rho}}_{d}\right)
\end{aligned}
$$

$$
\Delta = (I - \delta R')^{-1} d
$$

- $(I - \delta R')^{-1} \geq 0$ and $diag(I - \delta R')^{-1} > 0$.
- We want $d = R' v^{\rho} - R v^{\rho} \geq 0$ for all s and > 0 for some s.

- We want $\boldsymbol{d} = \boldsymbol{R}'\boldsymbol{v}^{\rho} - \boldsymbol{R}\boldsymbol{v}^{\rho} \geq \mathbf{0}$ for all $\boldsymbol{s}$ and $> \mathbf{0}$ for some $\boldsymbol{s}$.

- We want $\boldsymbol{d} = \boldsymbol{R'}\boldsymbol{v}^{\rho} - \boldsymbol{R}\boldsymbol{v}^{\rho} \geq \mathbf{0}$ for all $\boldsymbol{s}$ and $> \mathbf{0}$ for some $\boldsymbol{s}$.
-

$$\begin{aligned} \boldsymbol{d}_{\boldsymbol{s}} &= [\boldsymbol{R'}\boldsymbol{v}^{\rho}]_{\boldsymbol{s}} - [\boldsymbol{R}\boldsymbol{v}^{\rho}]_{\boldsymbol{s}} \\ &= \boldsymbol{v}^{\rho}(\rho'(\boldsymbol{s})) - \boldsymbol{v}^{\rho}(\rho(\boldsymbol{s})) \end{aligned}$$

- We want $\boldsymbol{d} = \boldsymbol{R}'\boldsymbol{v}^{\rho} - \boldsymbol{R}\boldsymbol{v}^{\rho} \geq \boldsymbol{0}$ for all $\boldsymbol{s}$ and $> \boldsymbol{0}$ for some $\boldsymbol{s}$.
-

$$
\begin{aligned}
\boldsymbol{d}_{\boldsymbol{s}} &= [\boldsymbol{R}'\boldsymbol{v}^{\rho}]_{\boldsymbol{s}} - [\boldsymbol{R}\boldsymbol{v}^{\rho}]_{\boldsymbol{s}} \\
&= \boldsymbol{v}^{\rho}(\rho'(\boldsymbol{s})) - \boldsymbol{v}^{\rho}(\rho(\boldsymbol{s}))
\end{aligned}
$$

- If $\rho'(\boldsymbol{s}) = \rho(\boldsymbol{s})$ then $\boldsymbol{d}_{\boldsymbol{s}} = \boldsymbol{0}$, so assume not. Then

- We want $\boldsymbol{d} = \boldsymbol{R}' \boldsymbol{v}^{\rho} - \boldsymbol{R} \boldsymbol{v}^{\rho} \geq \mathbf{0}$ for all $\boldsymbol{s}$ and $> \mathbf{0}$ for some $\boldsymbol{s}$.
- $$\begin{aligned} \boldsymbol{d}_{\boldsymbol{s}} &= [\boldsymbol{R}' \boldsymbol{v}^{\rho}]_{\boldsymbol{s}} - [\boldsymbol{R} \boldsymbol{v}^{\rho}]_{\boldsymbol{s}} \\ &= \boldsymbol{v}^{\rho}(\rho'(\boldsymbol{s})) - \boldsymbol{v}^{\rho}(\rho(\boldsymbol{s})) \end{aligned}$$
- If $\rho'(\boldsymbol{s}) = \rho(\boldsymbol{s})$ then $\boldsymbol{d}_{\boldsymbol{s}} = \mathbf{0}$, so assume not. Then
- If $\boldsymbol{s} \in \boldsymbol{S}_{\boldsymbol{Max}}$, because $\boldsymbol{s}$ was switched:

$$\boldsymbol{v}^{\rho}(\rho'(\boldsymbol{s})) > \boldsymbol{v}^{\rho}(\rho(\boldsymbol{s}))$$

- We want $\mathbf{d} = \mathbf{R}'\mathbf{v}^\rho - \mathbf{R}\mathbf{v}^\rho \geq \mathbf{0}$ for all $\mathbf{s}$ and $> \mathbf{0}$ for some $\mathbf{s}$.
-

$$\begin{aligned} \mathbf{d_s} &= [\mathbf{R}'\mathbf{v}^\rho]_\mathbf{s} - [\mathbf{R}\mathbf{v}^\rho]_\mathbf{s} \\ &= \mathbf{v}^\rho(\rho'(\mathbf{s})) - \mathbf{v}^\rho(\rho(\mathbf{s})) \end{aligned}$$

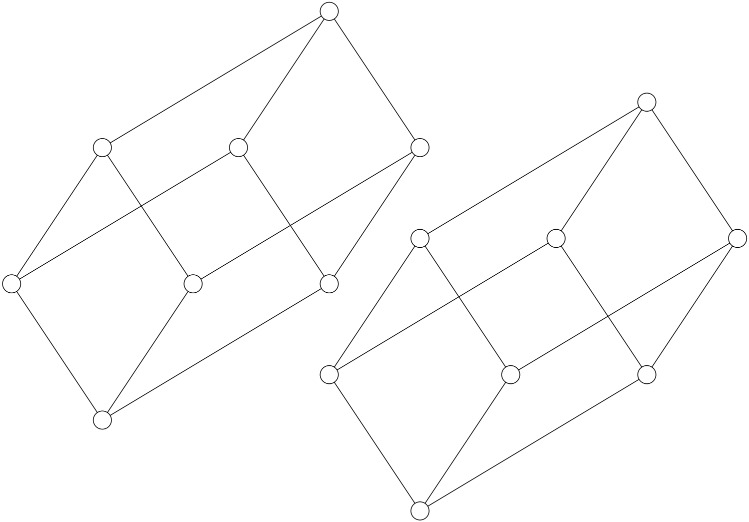
- If $\rho'(\mathbf{s}) = \rho(\mathbf{s})$ then $\mathbf{d_s} = \mathbf{0}$, so assume not. Then
- If $\mathbf{s} \in \mathbf{S}_{\mathbf{Max}}$, because $\mathbf{s}$ was switched:

$$\mathbf{v}^\rho(\rho'(\mathbf{s})) > \mathbf{v}^\rho(\rho(\mathbf{s}))$$

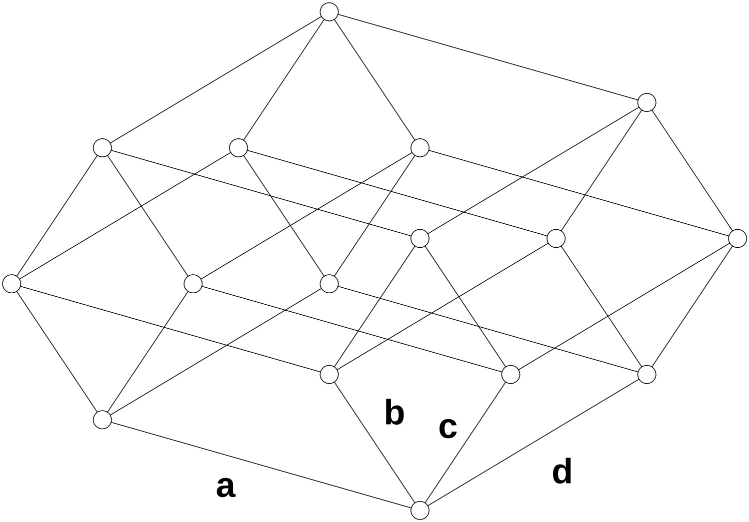
- If $\mathbf{s} \in \mathbf{S}_{\mathbf{Min}}$, because ***Min*** was playing a best response:

$$\mathbf{v}^\rho(\rho(\mathbf{s})) \leq \mathbf{v}^\rho(\rho'(\mathbf{s}))$$

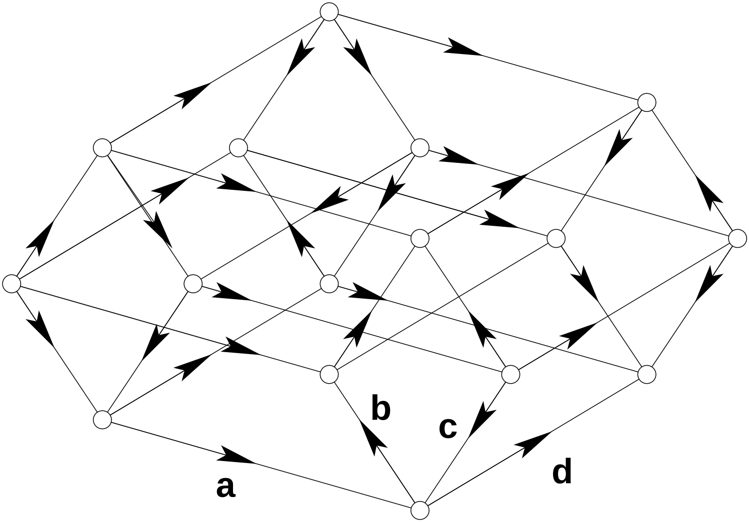
Inherited USO



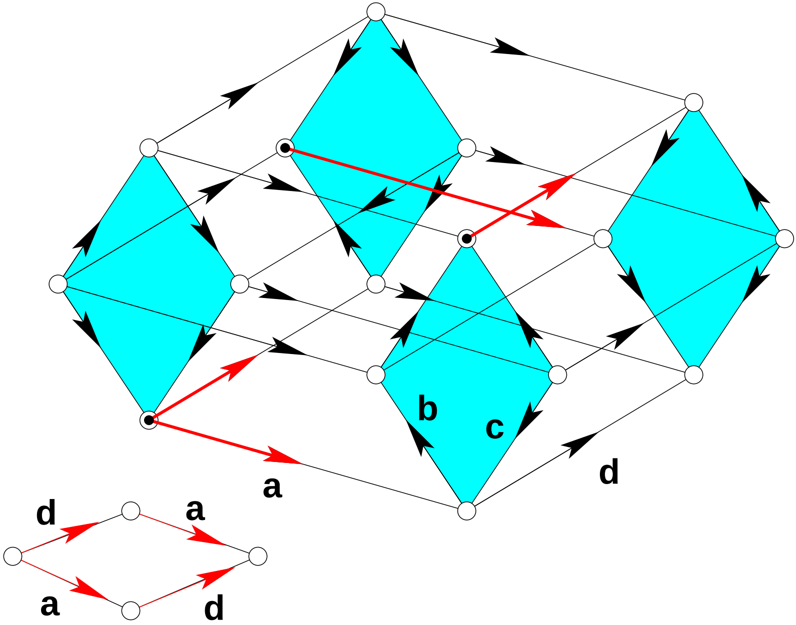
Inherited USO



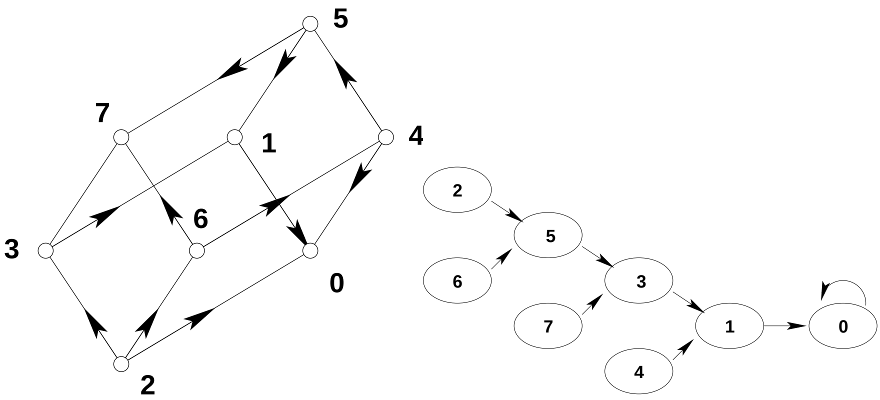
Inherited USO



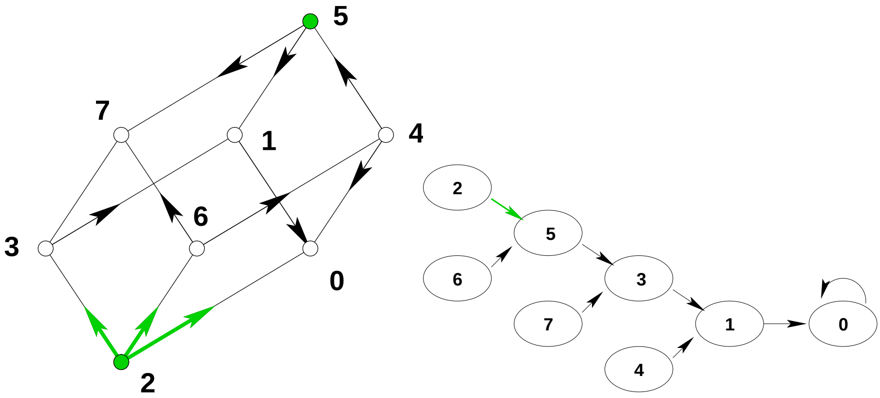
Inherited USO



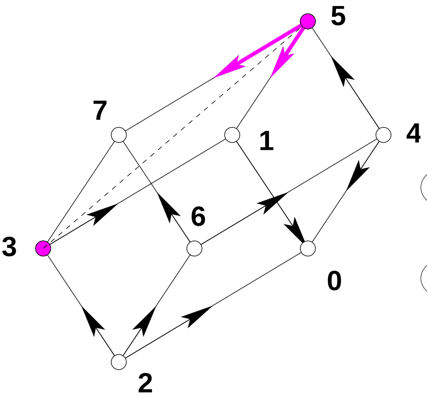
Strategy Improvement = Bottom-Antipodal

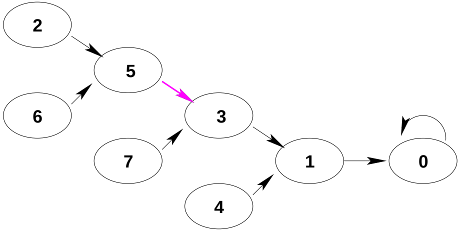


Strategy Improvement = Bottom-Antipodal

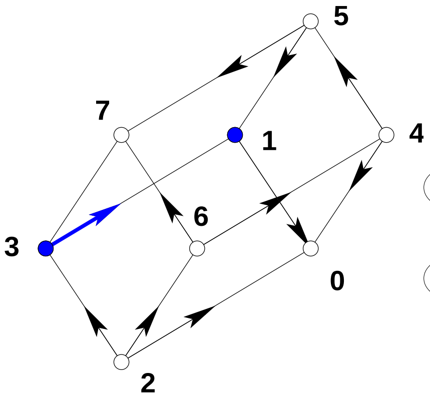


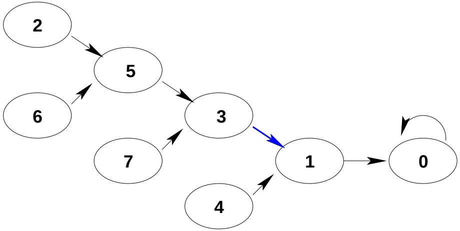
Strategy Improvement = Bottom-Antipodal



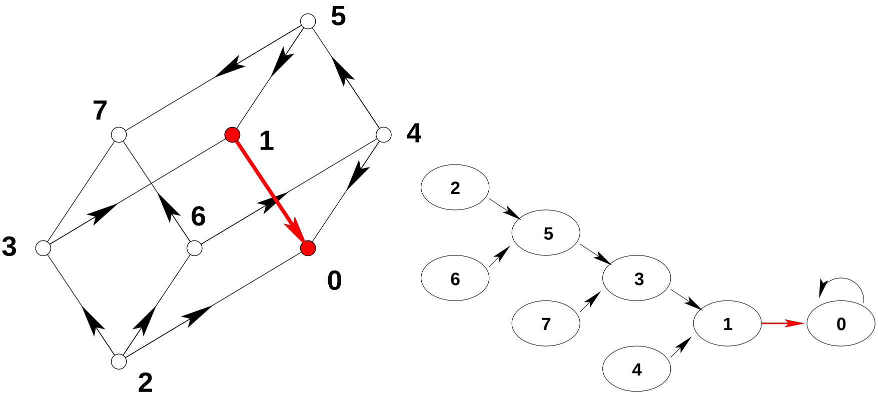


Strategy Improvement = Bottom-Antipodal

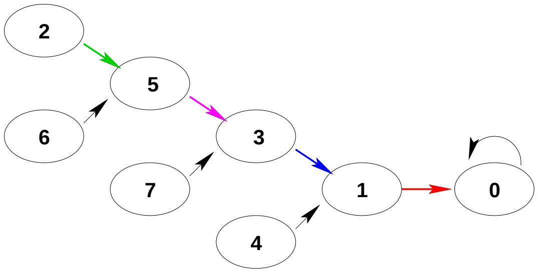




Strategy Improvement = Bottom-Antipodal



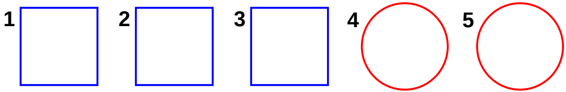
Strategy Improvement = Bottom-Antipodal



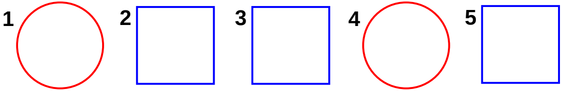
Interpretation of Murty’s least index method

Algorithm

Fix a permutation of states. Switch the switchable state with smallest index and repeat.



Max’s states come first - variant of strategy improvement.



This is a new algorithm.

Further research

- **Polynomial-time algorithms!**

Further research

- **Polynomial-time algorithms!**
- **Exponential examples** for (all-switching) **strategy improvement**
  - What strategy improvement (inherited) USOs arise from games?
  - Different from strategy improvement USOs for 1-player games?

Further research

- **Polynomial-time algorithms!**
- **Exponential examples** for (all-switching) **strategy improvement**
  - What strategy improvement (inherited) USOs arise from games?
  - Different from strategy improvement USOs for 1-player games?
- Can we identify **nontrivial classes** of games giving **nice** M's, e.g.,
  - **hidden-K matrices** (**LCP** solvable as an **LP**)
  - "**well-conditioned**": **polynomial PLCP interior point methods** **[Kojima, Megiddo, Ye (1988)]**

Further research

- **Polynomial-time algorithms!**
- **Exponential examples** for (all-switching) **strategy improvement**
  - What strategy improvement (inherited) USOs arise from games?
  - Different from strategy improvement USOs for 1-player games?
- Can we identify **nontrivial classes** of games giving **nice** ***M***'s, e.g.,
  - **hidden-K matrices** (**LCP** solvable as an **LP**)
  - "**well-conditioned**": **polynomial PLCP interior point methods** **[Kojima, Megiddo, Ye (1988)]**
- Study **other LCP and USO algorithms applied to games**, e.g., **Lemke's algorithm**, **Cottle-Dantzig principal pivoting**