Optimal Mixed Strategies for Cost-Adversarial Planning Games

R. Horčík [1] Á. Torralba [2] P. Rytíř [1] L. Chrpa [1] S. Edelkamp [1]

[1]*Czech Technical University in Prague* [2]Aalborg University

EUROPEAN UNION
European Structural and Investment Funds
Operational Programme Research,
Development and Education

MINISTRY OF EDUCATION, YOUTH AND SPORTS

2-Player Normal Form Games

✊	0,0	-1,1	1,-1
✋	1,-1	0,0	-1,1
✌	-1,1	1,-1	0,0

Formally, 2NFG is $\langle X, Y, u_1, u_2 \rangle$ where

X, Y are sets of Row's and Column's players pure strategies
$u_1 \colon X \times Y \to \mathbb{R}$, $u_2 \colon X \times Y \to \mathbb{R}$ utilities

Game is zero-sum if $u_1(x, y) + u_2(x, y) = 0$. It is almost zero-sum if $u_1(x, y) + u_2(x, y) = f(x)$.

Nash Equilibrium

Mixed strategy is a prob. distribution $p \colon X \to [0, 1]$ (resp. $q \colon Y \to [0, 1]$)

Best response w.r.t. q is p' such that

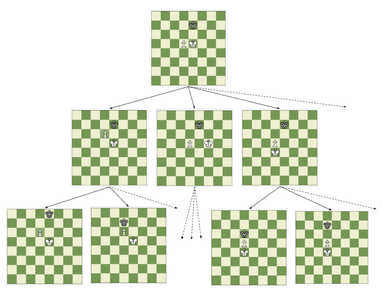
$$u_1(p', q) = \max_p u_1(p, q)$$

Definition: Pair of strategies $\langle p^*, q^* \rangle$ is Nash Equilibrium if p^* and q^* are mutually the best responses w.r.t. each other.

For zero-sum games NE can be computed in P-time by LP.

For large games, use an iterative method like Double Oracle based on the best response computations.

Extensive Form Games



Strategies as Plans

EFGs can be presented as NFGs.

Pure strategies are policies of the form $\pi \colon State \to Action$.

Observation

If we need to prepare the whole action sequence in advance, we can model pure strategies as plans!

The cost of a plan is the opposite of our utility and opponent's strategy influences the action costs.

Cost-Adversarial Planning Games

Π — planning task $\mathcal{P}(\Pi)$ — set of all its plans
c — base cost function $\mathcal{C}$ — finite set of cost functions

Cost-Adversarial Planning Game (CAPG)

is an almost zero-sum game $\langle \mathcal{P}(\Pi), \mathcal{C}, u_1, u_2 \rangle$ where

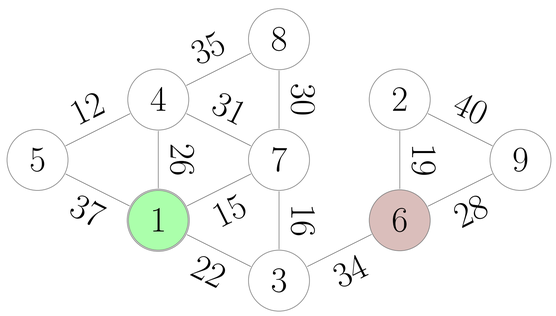
$$u_1(\pi, g) = -c(\pi) - g(\pi), \qquad u_2(\pi, g) = g(\pi)$$

for $\pi \in \mathcal{P}(\Pi)$ and $g \in \mathcal{C}$.

The best response w.r.t. a cost function is the optimal plan!

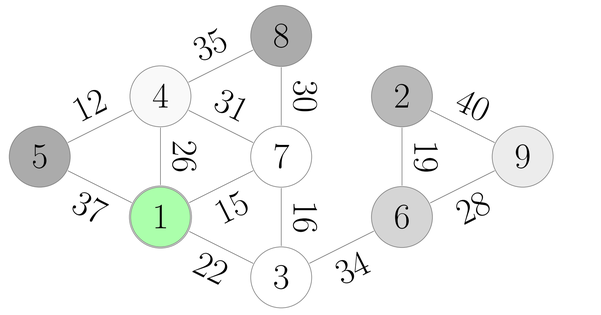
CAPGs can be solved by Double Oracle + optimal planner.

Patrolling Games

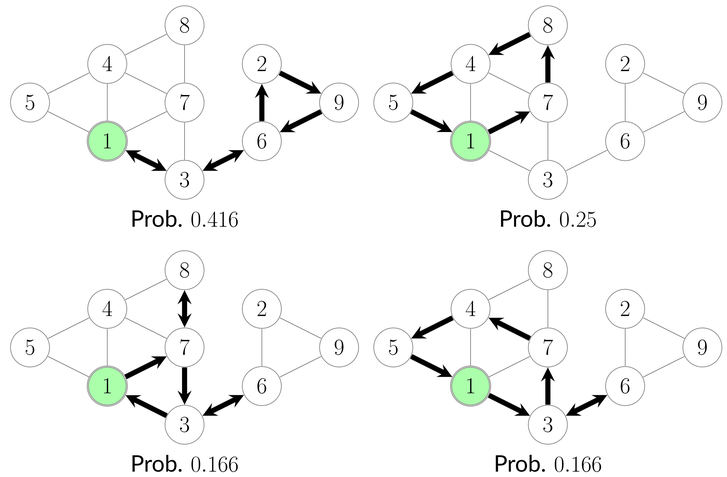


Location 6 — 70%, other locations — 30%

Poacher's Mixed Strategy



Guard's Mixed Strategy



Experiments

domain	ipdb NE	ipdb cov	lmcut NE	lmcut cov	ms NE	ms cov
patrol	**85**	**90**	71	**90**	81	**90**
transport	**13**	**14**	9	11	11	11
transport-road	**13**	**14**	10	11	11	11
data-network	**3**	**12**	**3**	**12**	**3**	10
visitall[11]	**11**	**16**	9	10	9	9

NE found in almost all cases when the optimal planner solved the base planning task.

{xhorcik, pavel.rytir, chrpaluk, edelkste}@fel.cvut.cz, alto@cs.aau.dk