Graduate AI

Lecture 20: Game Theory II

Teachers:
Zico Kolter
Ariel Procaccia (this time)

A curious game

- Playing up is a dominant strategy for row player
- So column player would play left
- Therefore, (1,1) is the only Nash equilibrium outcome

1,1	3,0
0,0	2,1

Commitment is good

- Suppose the game is played as follows:
  - Row player commits to playing a row
  - Column player observes the commitment and chooses column
- Row player can commit to playing down!

1,1	3,0
0,0	2,1

Commitment is good

- By committing to a mixed strategy, row player can guarantee a reward of 2.5
- We assume that the follower breaks ties in favor of the leader
- This is called a strong Stackelberg equilibrium (SSE)

	0	1
.49	1,1	3,0
.51	0,0	2,1

Computing Stackelberg

- Theorem [Conitzer and Sandholm 2006]: In 2-player normal form games, an SSE can be found in polynomial time

Proof of theorem

- For each pure follower strategy s_2, we compute via the LP below a strategy x_1 for the leader such that
  - Playing s_2 is a best response for the follower
  - Under this constraint, x_1 is optimal
- Choose x_1^* that maximizes leader value

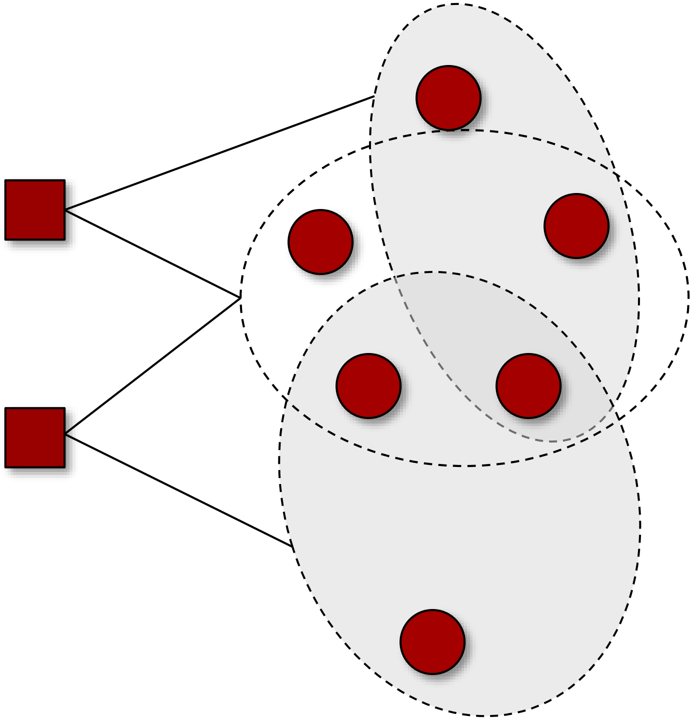
$$
\begin{aligned}
\textbf{max}\ & \textstyle\sum_{s_1 \in S} x_1(s_1) u_1(s_1, s_2) \\
\\
\textbf{s.t.}\ & \forall s_2' \in S,\ \textstyle\sum_{s_1 \in S} x_1(s_1) u_2(s_1, s_2) \geq \sum_{s_1 \in S} x_1(s_1) u_2(s_1, s_2') \\
& \textstyle\sum_{s_1 \in S} x_1(s_1) = 1 \\
& \forall s_1 \in S, x_1(s_1) \in [0,1]
\end{aligned}
$$

Application: security

Defender (leader) commits to a randomized allocation of security resources, attacker (follower) best responds

Security games

- Set of targets $T = \{1, \ldots, n\}$
- Set of m security resources R available to the defender (leader)
- Set of schedules $\Sigma \subseteq 2^T$
- Resource r can be assigned to one of the schedules in $A(r) \subseteq \Sigma$
- Attacker chooses one target to attack

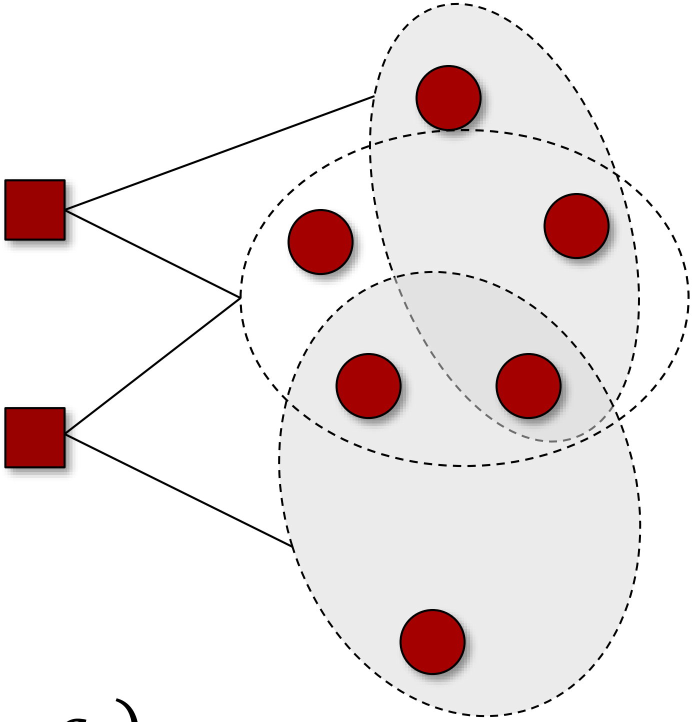


Security games

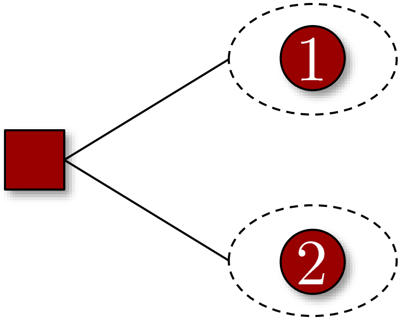
- For each target t, there are four numbers: $u_d^+(t) \geq u_d^-(t)$, and $u_a^+(t) \leq u_a^-(t)$
- Randomized defender strategy induces coverage probabilities $\boldsymbol{c} = (c_1, \ldots, c_n)$
- The utilities to the defender/attacker under **c** if target t is attacked are

$$u_d(t, \boldsymbol{c}) = u_d^+(t) \cdot c_t + u_d^-(t)(1 - c_t)$$

$$u_a(t, \boldsymbol{c}) = u_a^+(t) \cdot c_t + u_a^-(t)(1 - c_t)$$



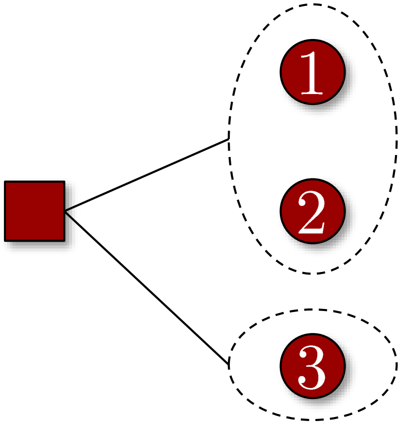
Example



	Attacker		Defender	
Target	Covered	Uncovered	Covered	Uncovered
1	0	1/2	0	-1
2	0	1	0	-1

Poll 1: What is the probability of assigning the resource to {1} in an SSE?

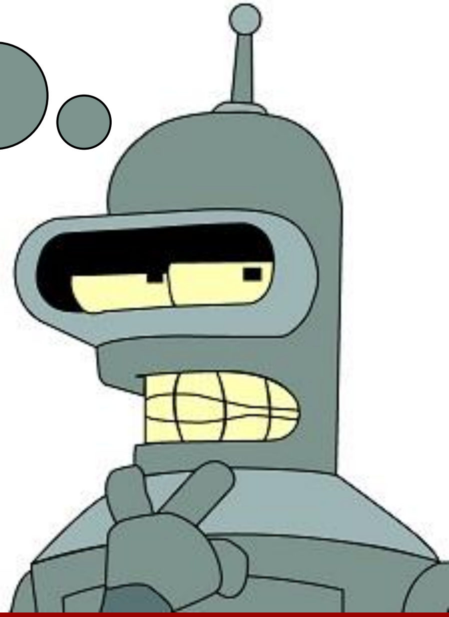
Another example



	Attacker		Defender	
Target	Covered	Uncovered	Covered	Uncovered
1	0	4	0	-2
2	0	3	0	-1
3	0	2	0	-5

Poll 2: What is the probability of assigning the resource to {1,2} in an SSE?

This is a 2-player Stackelberg game, so we can compute an optimal strategy for the defender in polynomial time...?



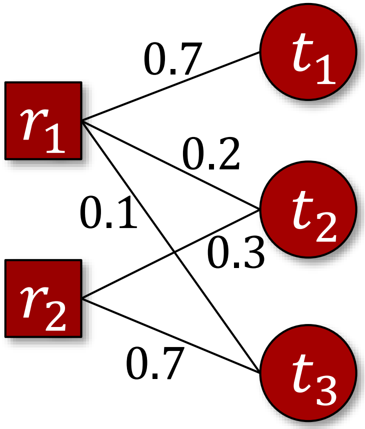
Solving security games

- Consider the case of $\Sigma = T$, i.e., resources are assigned to individual targets, i.e., schedules have size 1
- Nevertheless, number of leader strategies is exponential
- Theorem [Korzhyk et al. 2010]: Optimal leader strategy can be computed in poly time

A compact LP

- LP formulation similar to previous one
- Advantage: logarithmic in #leader strategies
- Question: Do probabilities correspond to strategy?
- Answer: Yes! Example in next slide

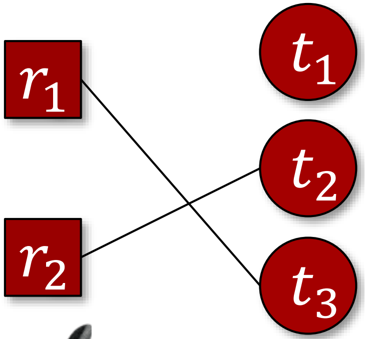
$$
\begin{aligned}
\textbf{max}\ & u_d(t^*, \boldsymbol{c}) \\
\textbf{s.t.}\ & \forall r \in R, \forall t \in A(r), 0 \leq c_{r,t} \leq 1 \\
& \forall t \in T, c_t = \sum_{r \in R : t \in A(r)} c_{r,t} \leq 1 \\
& \forall r \in R, \sum_{t \in A(r)} c_{r,t} \leq 1 \\
& \forall t \in T, u_a(t, \boldsymbol{c}) \leq u_a(t^*, \boldsymbol{c})
\end{aligned}
$$



	t_1	t_2	t_3
r_1	0.7	0.2	0.1
r_2	0	0.3	0.7

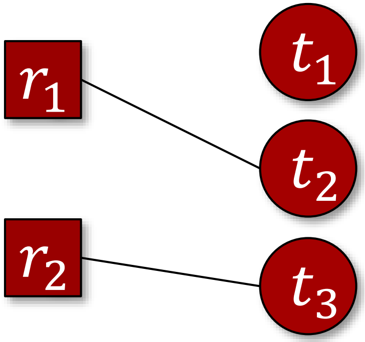
<u>0.1</u>

	t_1	t_2	t_3
r_1	0	0	1
r_2	0	1	0



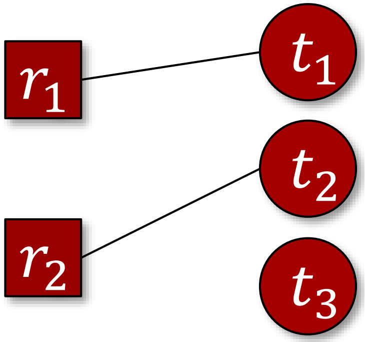
<u>0.2</u>

	t_1	t_2	t_3
r_1	0	1	0
r_2	0	0	1



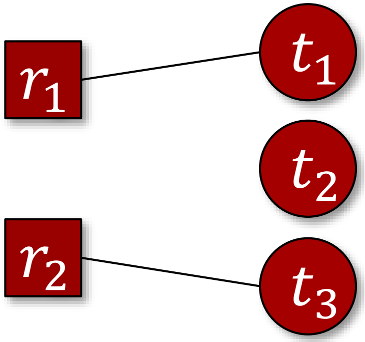
<u>0.2</u>

	t_1	t_2	t_3
r_1	1	0	0
r_2	0	1	0



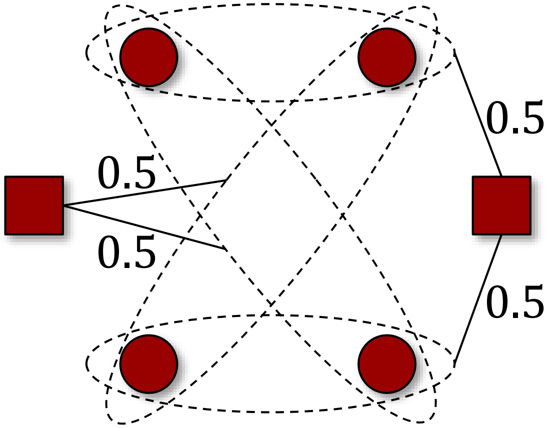
<u>0.5</u>

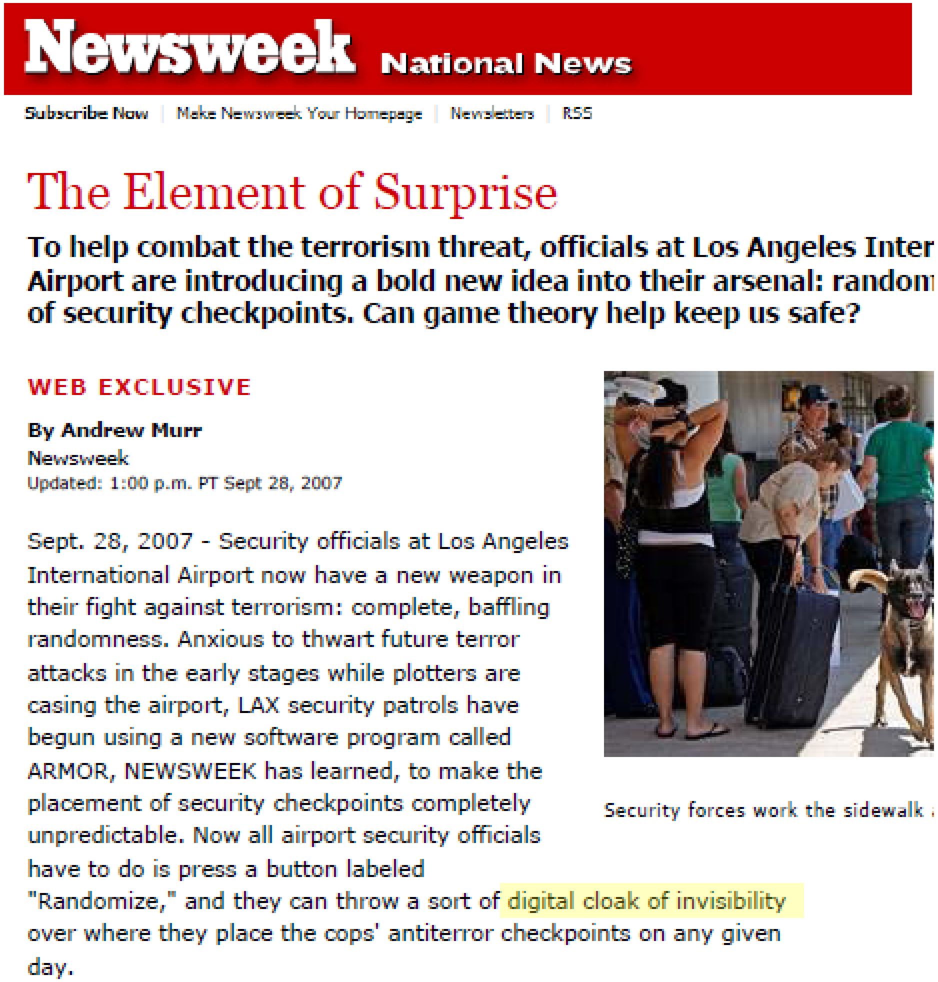
	t_1	t_2	t_3
r_1	1	0	0
r_2	0	0	1



Generalizing

- What about schedules of size 2?
- Air Marshals domain has such schedules: outgoing+incoming flight
- Previous apporoach fails
- Theorem [Korzhyk et al. 2010]: problem is NP-hard





Newsweek National News

Subscribe Now | Make Newsweek Your Homepage | Newsletters | RSS

The Element of Surprise

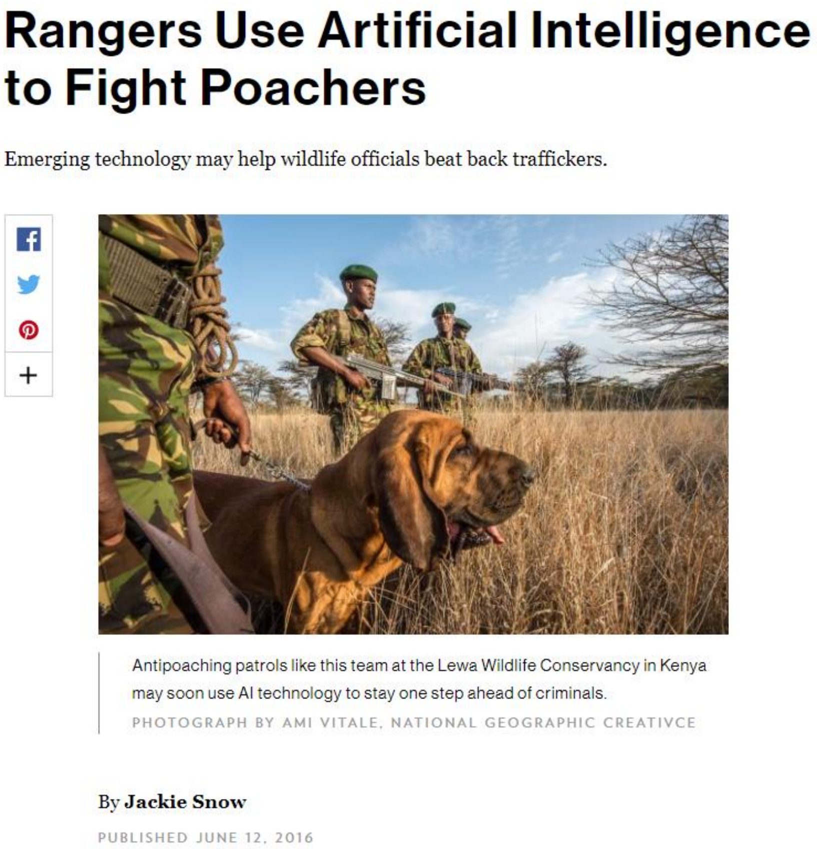
To help combat the terrorism threat, officials at Los Angeles Inter Airport are introducing a bold new idea into their arsenal: randon of security checkpoints. Can game theory help keep us safe?

WEB EXCLUSIVE

By Andrew Murr
Newsweek
Updated: 1:00 p.m. PT Sept 28, 2007

Security forces work the sidewalk .

Sept. 28, 2007 - Security officials at Los Angeles International Airport now have a new weapon in their fight against terrorism: complete, baffling randomness. Anxious to thwart future terror attacks in the early stages while plotters are casing the airport, LAX security patrols have begun using a new software program called ARMOR, NEWSWEEK has learned, to make the placement of security checkpoints completely unpredictable. Now all airport security officials have to do is press a button labeled "Randomize," and they can throw a sort of digital cloak of invisibility over where they place the cops' antiterror checkpoints on any given day.



Rangers Use Artificial Intelligence to Fight Poachers

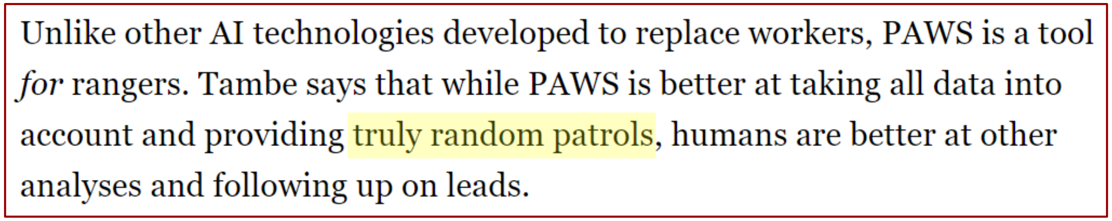
Emerging technology may help wildlife officials beat back traffickers.

Antipoaching patrols like this team at the Lewa Wildlife Conservancy in Kenya may soon use AI technology to stay one step ahead of criminals.

PHOTOGRAPH BY AMI VITALE, NATIONAL GEOGRAPHIC CREATIVCE

By **Jackie Snow**

PUBLISHED JUNE 12, 2016



Unlike other AI technologies developed to replace workers, PAWS is a tool *for* rangers. Tambe says that while PAWS is better at taking all data into account and providing truly random patrols, humans are better at other analyses and following up on leads.

Limitations

- The defender knows the utility function of the attacker
  - ○ Solution: machine learning
- The attacker perfectly observes the defender's randomized strategy
  - ○ MDPs, although this may not be a major concern
- The attacker is perfectly rational, i.e., best responds to the defender's strategy
  - ○ Solution: bounded rationality models

Learning to play

- Suppose the defender does not know the attacker's utility function
- The defender can interact with the attacker by playing a strategy and observing the attacker's best response
- Theorem [Haghtalab et al. 2014]: For any $\epsilon, \delta > 0$, with probability $1 - \delta$, it is possible to learn a defender strategy that is optimal up to ϵ with a number of best-response queries that is polynomial in n and logarithmic in $1/\epsilon$, $1/\delta$

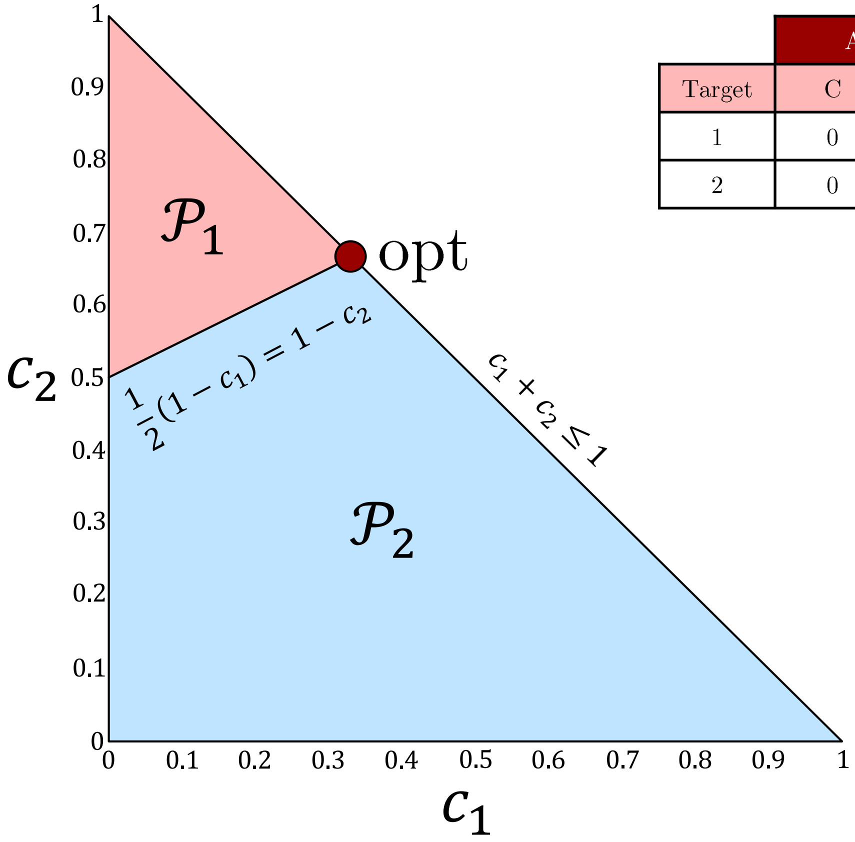
Learning to play

- We can reformulate the LP as

$$\begin{aligned} &\textbf{max}\ u_d(t,\boldsymbol{c}) \\ &\textbf{s.t.}\ t \text{ is attacked under } \boldsymbol{c} \\ &\qquad \boldsymbol{c} \text{ is implementable} \end{aligned}$$

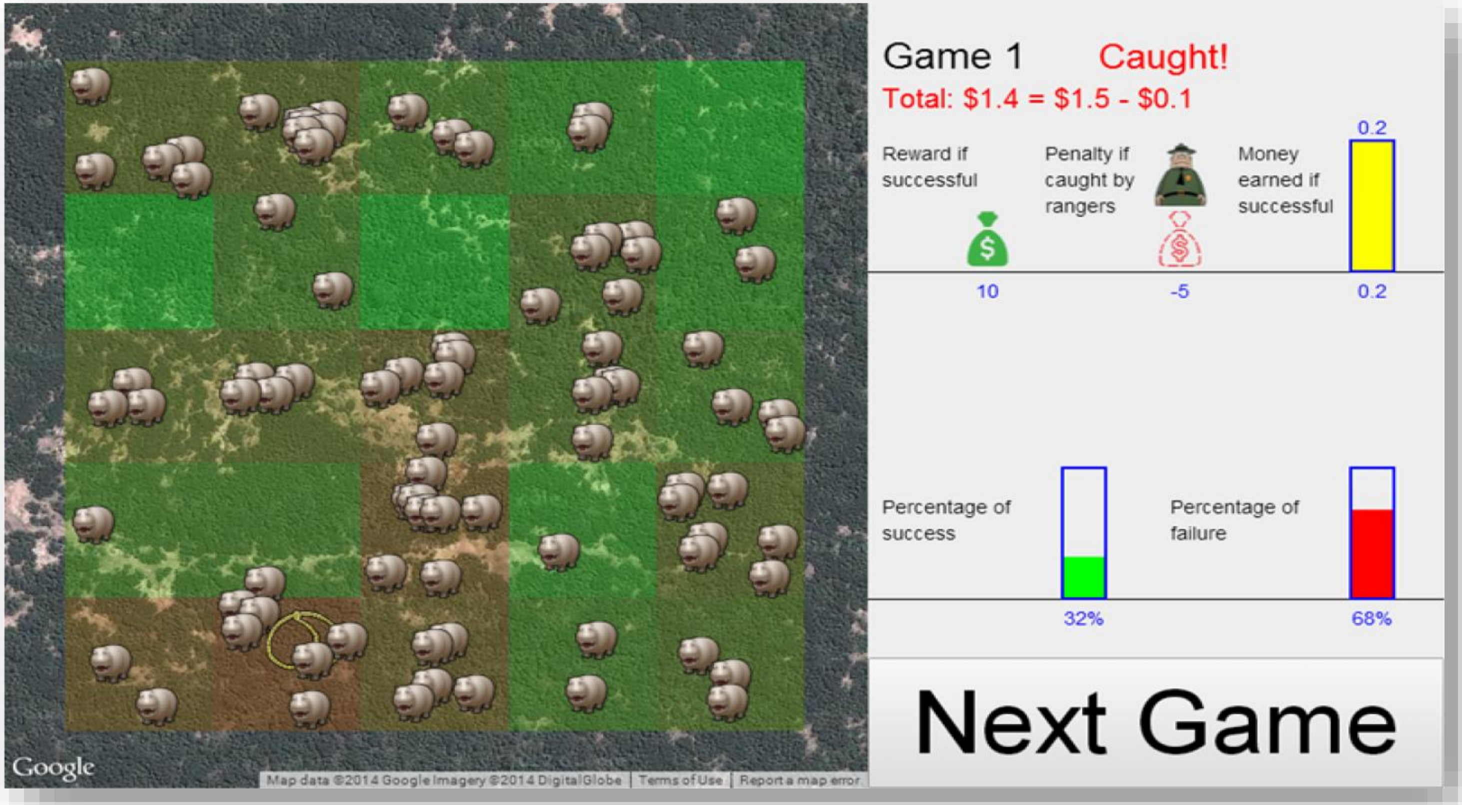
- It holds that:
  - The objective function is linear in $\boldsymbol{c}$
  - The feasible region $\mathcal{P}_t$ is convex
  - There is a membership oracle for the feasible region
- Result now follows using an algorithm for optimizing a linear function in a convex region using membership queries [Kalai and Vempala 2006]

Learning to play



	Attacker		Defender	
Target	C	U	C	U
1	0	1/2	0	-1
2	0	1	0	-1

Testing bounded rationality



Game 1 Caught!

Total: $1.4 = $1.5 - $0.1

Reward if successful	Penalty if caught by rangers	Money earned if successful
10	-5	0.2

Percentage of success	Percentage of failure
32%	68%

0.2
Next Game
Google
Map data ©2014 Google Imagery ©2014 DigitalGlobe | Terms of Use | Report a map error

[Kar et al. 2015]

Summary

- Terminology:
  - Stackelberg game
  - Strong Stackelberg equilibrium
  - Security game
- Algorithms:
  - Multiple LPs algorithm
  - Polynomial time algorithm for singleton schedules