November 6, 2002

beeve@cs.rpi.edu
Algorithmic Robotics Laboratory
Rensselaer Polytechnic Institute
Kris Beevers

Pursuit-Evasion Problem
Game Theory, Robotics and the
Dynamic Game Theory: the order in which decisions are made is important [1]

- Non-cooperative
- Cooperative: formation of coalitions

Two main branches [2]

Psychologists would say: the theory of social situations [2]

In a nutshell: multiperson decision-making [1]

What is Game Theory?
Conflict situation: players value possible outcomes differently

- payoffs (both good and bad) depend on the values of the player
- what players know about moves of other players, the environment,
- moves, players

A game:

Each person involved pursues his or her own (partly conflicting) interests

Non-cooperative Game Theory [1]
Strategic form: map from strategies to payoffs

- Play on their behalf
- Set of instructions that a player could give to a friend or program (to

Strategy: Fundamental notion in noncooperative Game theory

Strategies [2]
and confess

Game theory predicts each player will thus follow their own self-interests.

- No matter what other player does, it's best for me to confess.
- If other player confesses, also best to confess (1 instead of 0).
- If other player doesn’t confess, best for me to confess (9 instead of 5).

**But reasoning as follows:**

- Total payoff highest when neither confesses (5, 5).

<table>
<thead>
<tr>
<th></th>
<th>Player 2</th>
<th>Player 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>5, 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0, 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1, 1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Strategic Form: Prisoners' Dilemma
Advantages

- Cooperation between two or more players may lead to mutual

Nonzero-sum Games: sum of cost functions nonconstant

- Constant-sum (like Prisoners’ Dilemma): transform to zero-sum

- Usually two players

Zero-sum Game: sum of cost functions of the players is zero

Zero vs. Nonzero-sum Games [1]
If both escalate, both are worse off.

If one guesses the other will give in, he will escalate.

- BUT: both won't necessarily give in.

By giving in, both can maximize their minimal payoff.

<table>
<thead>
<tr>
<th></th>
<th>( Y )</th>
<th>D</th>
<th>( Y )</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Y )</td>
<td>0</td>
<td>1-10</td>
<td>( Y )</td>
<td>D</td>
</tr>
<tr>
<td>D</td>
<td>1-10</td>
<td>0</td>
<td>D</td>
<td>1-10</td>
</tr>
</tbody>
</table>

Payoff matrix for Johnny:

Players: Johnny, Oscar. Both have option to escalate a brawl or give in.

Example Game [3]: "Chicken" (nonzero-sum).

Another example game [3]: "Chicken" (nonzero-sum).

\textbf{Nash Equilibrium}
n neither has anything to gain by deviating from equilibrium *

\[
\text{equilibrium}
\]

- if both escalate with probability \( \frac{0}{1} \), they are in Nash equilibrium and escalate \( O \) with probability \( \frac{0}{1} \) \( \text{should escalate} \)

- if escalate with probability \( O \) \( \frac{0}{1} \) \( \text{should yield} \)

\[
\hat{v} - x + \hat{v}x - 10x = f \quad \text{(expected payoff for \( f \))}
\]

\( x, \hat{v} : \text{probabilities that \( f \), escalate, resp.)} \quad \text{(Nash equilibrium (cont.))}


- J. Hespanha, M. Prandini, S. Sastry

- multiple-robot motion planning (coordinatization)
- motion planning under environment uncertainties
- motion planning under uncertainty in sensing and control

- A Game-Theoretic Framework for Robot Motion Planning (PhD thesis)

- S. LaValle

Game Theory: Applications in Robotics
Pursuit-Evasion Games

- dispersion
- aggregation
- flocking
- following

Collective behaviors

- navigation
- foraging/search-and-rescue
- obstacle avoidance

'Degenerate\' cases (iniminate opponents)

Several obvious applications in robotics/distributed robotics [4]
Pursuit-Evasion and Game Theory [4]

- Problem first posed by Isaacs in 1950s [5]
- Considered extensively in aerial combat context (e.g. missiles)
- Different from previous games
  - continuous unfolding of moves, continuous variation in strategies
- Game theory can handle pursuit-evasion
  - optimal pursuit strategy depends on evasion strategy adopted by other player and vice-versa—just what game theory is good at
  - continuous nature modeled by differential equations
  - approach: pursuers minimize time to capture, evaders maximize time to capture
\( S \) \( \in \) \((t)^0 x, (t)^d x, (t)^e x\) = \((t)^s x\) ●

\( \mathcal{L} \) \( \ni \) \((1 + t)^0 x = (t)^0 x\) (fixed) ●

Obstacle positions (fixed) ●

Positions at time \( t \): (pursuer \( \# \), evader \( \# \)) ●

Some cells may contain obstacles; configuration of obstacles not perfectly known ●

\( \{ \cdots , \#_1 , \#_2 \} = \mathcal{L} \) All events take place at a set of equally-spaced times ●

\( \{ \#_1 , \cdots , \#_n \} = \mathcal{X} \) Pursuit region: finite collection of cells ●

\( \{ \#_1 , \cdots \} = \mathcal{X} \) Single evader (player D) ●

Notation
Pursuers' evaders reach chosen adjacent cells with probability $p^e$.

$\chi \supset (x)\forall x$ Set of cells reachable in one time step by an agent at $x$.

Outcome of the evader's uncertain action will produce the desired transition probability: probability that next state will be $s'$ given $s$ at time $t$.

$\mathcal{A} \in (t)p', \mathcal{U} \in (t)n$ Next desired positions for pursuers, evader: $\mathcal{A}$ resp. the sets of actions available to $\mathcal{U}$, $\mathcal{D}$ resp. Every time instant $t$ each player can choose control actions from $\mathcal{A}, \mathcal{U}$.
Observations
for a probability that depends on \( n \) and \( \theta \).

Probability measures vary with policies, so we denote them as \( p_\pi(n, \theta) \) (e.g. • Probability distribution (a, policy)

\( \pi \), each player selects action for time \( t \) according to some probability

\( \pi, \theta \): stochastic policies

\( \pi \), \( \theta \): stochastic policies
Problem Formulation
observations

So, pursuers try to maximize estimate of evader’s cost based on

information (since the evader knows it)

\( \pi \) pursuit similar to \( n \pi d \) but takes into account \( n \pi d \) scalar in distribution of corresponding to action \( n \pi d \) over distribution of actions \( p ' \), \( n ' \), \( s ' \), \( s \) given

\( \pi \) is a transition probability function (\&). Given

\[
[ \pi = \pi \pi (Z | b , d ) \pi \pi ]^{1 - q_1 - q_1 - q_1} \pi = (b , d ) \pi \pi
\]

\[
( (Z = Z | s = (q)s )^{1 - q_1 - q_1 - q_1} d ) (p , n ' , s , s ) d \bigcup (Z \pi \pi n d ) \bigcup = (Z \pi \pi b , d ) \pi \pi
\]

Cost functions:

One-Step Nash Equilibrium Solution
• Natural tendency for the game to be played at Nash equilibrium

• Well-known solution: Nash equilibrium

moves, what exactly does "optimize a cost" mean?

Since each player's incurred cost depends on the other player's choice of

\[ \int_{U}^{D} J_{t}(x) \text{ represent cost functions optimized at time } t \text{ by } U \text{ and } D \]

Nash Equilibrium Solution (cont.)
Pair $(\vec{b}, \vec{d})$ is called a one-step Nash equilibrium.

\[ b_A (Z \vec{b}, \vec{d}) \geq (Z \vec{b}, \vec{d}) \]
\[ d_A (\vec{b}, \vec{d}) \leq (\vec{b}, \vec{d}) \]

Players choose actions equal to $(Z)_{\vec{g}} (X)_{\vec{r}}$ satisfying $\vec{b}, \vec{d}$. Nash Equilibrium Solution (cont.)
is forced to choose \( \hat{b} \) \( \hat{d} \). Essentially, can do this because \( \hat{d} \) is rational evader.

\( \text{Nash Equilibrium Solution (cont.)} \)
Nash Equilibrium

Solution (cont.)
(measurement data)

- knows pursuers' locations perfectly (because it has access to their
  - can perfectly sense adjacent cells $\mathcal{A}(x)$ for obstacles

**Example (Simulation)**

- Evader

  \[ (x)^{uf} \text{ (for } \mathcal{A}) \]

  \[ (x)^{df} \text{ (for } \mathcal{A}) \]

  false positives \& false negatives

- perfect sensing for cell pursuer is currently in

  \[ \text{senses for evaders} \]

  - can perfectly sense adjacent cells $\mathcal{A}(x)$ for obstacles

  - can perfectly determine position $x$

- Pursuers
Frames taken every four time steps

\[ I_0 = u_f = d_f \]

slow evader \( (\text{dark circle}) \)

[light stars] \( I = d_d \)

3 fast pursuers

400 cells

Example (cont.)
Example (cont.)
Game theory [3]

- modern Game theory offers (among other approaches) evolutionary
- who's to say other players aren't irrational

- Rationality assumption

- equilibrium; 2-3 robots, up to an hour of computation
- LaValle/Hutchinson [7]: coordination problem solved with Nash

Applications

Some Problems With Game Theory and Robotics


3. J. Hofbauer and K. Sigmund. Evolutionary Games and Population


References


Games: A One-Step Nash Approach. Submitted to the 39th Conf. on

Corporation, 1951.
Games of pursuit. Technical Report P-257, Rand

5. Isaacs, R.