16.410/413
Principles of Autonomy and Decision Making

Lecture 25: Differential Games

Sertac Karaman

Massachusetts Institute of Technology

December. 8, 2010

Outline

- Game theory and sequential games (recap previous lecture)
- Dynamical (control) systems and optimal control
- Dynamic Game Theory
- Numerical Methods
- A special case: Pursuit-evasion.

Game theory (Recap)

Zero-sum Games

Gains/losses of each player is balanced by the gains/losses of the all the other players.

Cooperative vs. non-cooperative.

Cooperative if groups of players may enforce binding agreements.

Nash equilibrium

No player can gain more by unilaterally changing strategy.

An example

Remember the prisoner's dilemma:

	Player B cooperates	**Player B defects**
Player A cooperates	(-1,-1)	(-10, 0)
Player A defects	(0,-10)	(-5,-5)

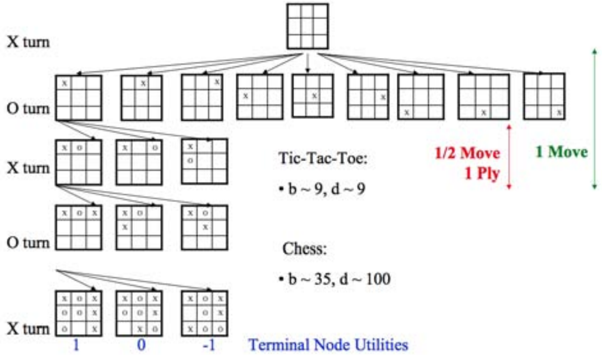
- Non-zero sum.
- Cooperation could have been enforced; otherwise may or may not arise.

Game Theory (Recap)

Zero-sum Two-player Sequential games

Key characteristics

- Two players
- Zero-sum reward
- Sequential moves (from a finite set)
- Perfect information
- Terminates in a finite number of steps



- We have used alpha-beta pruning to solve such games.
- Today, we will study **non-cooperative dynamic games**.

Dynamic games

- Dynamic games:

> Actions available to each agent depends on its current state which evolves according to a certain dynamical system.
> *Sets of states/actions is usually a continuum.*

- In many cases, the agents involved in the game are subject to dynamics.

Some (major/relevant) application areas:

- Dogfight
- Aircraft landing subject to wind (or other) disturbance
- Air traffic control
- Economics & Management Science

History of Dynamic Games

- Introduction of dynamic games is attributed to Rufus Isaacs (1951).

  Book: R. Isaacs, Differential Games: A mathematical theory with applications to warfare and pursuit, control and optimization, 1965.

- Later the theory was developed by many contributors including A. Merz and J. Breakwell.

- More recent contributions by T. Basar and coworkers.

  Book: Basar and Olsder, Dynamic Noncooperative Game Theory, 1982.

Dynamic Games Literature

- Dynamic games has a very rich literature.

Images of book covers removed due to copyright restrictions:

Isaacs, Rufus. *Differential Games: A Mathematical Theory with Applications to Warfare, Pursuit, Control and Optimization.* Dover, 1999. ISBN: 9780486406824.

Basar, Tamer, and Geert Jan Olsder. *Dynamic Noncooperative Game Theory*. 2nd ed. SIAM, 1999. ISBN: 9780898714296.

Dockner, Engelbert, Steffen Jorgensen, Ngo Van Long, and Gerhard Sorger. *Differential Games in Economics and Management Science.* Cambridge University Press, 2001. ISBN: 9780521637329.

Dynamical systems

- Two definitions of time:

Discrete Time

$t \in \mathbb{N}$: time takes values in $\{0, 1, 2, \ldots\}$.

- Can be thought of as "steps".
- Good models of computers and digital systems.

Continuous Time

$t \in \mathbb{R}_{\geq 0}$: time takes values in $[0, \infty)$.

- Models of systems arising from (large-scale) physical phenomena
- Examples: airplanes, cars, room temperature, planets moving around the sun

- **(Autonomous) Discrete-time Dynamical Systems** described by *difference equations*:

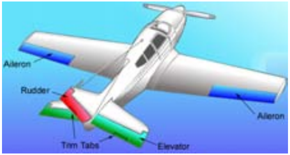
$$x[t+1] = f(x[t])$$

- **(Autonomous) Continuous-time Dynamical Systems** described by *differential equations*:

$$\dot{x} = \frac{dx}{dt}(t) = f(x(t)), \quad x(t) \in X \text{ (state space)}$$

Dynamical Control Systems

- Almost all engineering systems have a certain set of inputs.



> The behavior of the system is determined by its current state and the inputs.

- **Discrete-time dynamical control systems**

> *Difference equation:* $x[t+1] = f(x[t], u[t])$.

- **Continuous-time dynamical control systems**

> *Differential equation:* $\dot{x}(t) = f(x(t), u(t)), x(t) \in X, u(t) \in U$.

- From now on we will only discuss continuous-time systems, although the discussion can easily be extended to discrete-time systems

Dynamical Control Systems: Examples

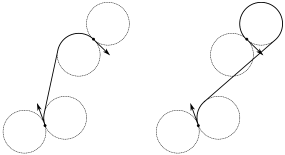
- **Single integrator**

$$\dot{x} = u, \quad |u| \leq 1.$$

  - Can be extended to multiple dimensions easily.

- **Dubins' car**
  - States: x, y, θ; Input: $u \in [-1, 1]$.

$$\begin{aligned} \dot{x} &= v\cos(\theta) \\ \dot{y} &= v\sin(\theta) \\ \dot{\theta} &= u \end{aligned}$$



  - The car can not turn on a dime, i.e., has a minimum turning radius.

Optimal Control

- Often times in engineering, we would like design to maximize a certain performance (equivalently minimize a cost function)
- Let $g(x, u) : X \times U \to \mathbb{R}_+$ associate state-input pairs with a cost "density". Define:
$$L(u) = \int_{t=0}^{T} g(x(t), u(t)) dt,$$
where $\dot{x}(t) = f(x(t), u(t))$ for all $t \in [0, T]$ (T might be infinity).
- **Optimal control problem** is to find $u(t)$ such that $L(u)$ is minimized.
- Optimal control is widely studied. Generally, solution methods are based on dynamic programming and the *principle of optimality*.
- Analytical techniques apply when, e.g., linear dynamics (f linear) and quadratic cost (g quadratic).

Differential games

Dynamical systems with many independently-controlled inputs

Player 1 controls $u^1(t) \in U_1$, Player 2 controls $u^2 \in U_2$:

$$\dot{x} = f(x, u^1(t), u^2(t))$$

- The state evolves according to both players decisions.

- **Payoff function**:
  - For each player $i \in \{1, 2\}$, define $g^i : X \times U_1 \times U_2 \to \mathbb{R}_+$

$$L_i(u^1, u^2) = \int_{t=0}^{T} g^i(x(t), u^1(t), u^2(t))$$

- Each player wants maximize her own payoff *(knowing that the other player is doing the same)*.
- Another type of **dynamic game** is the **difference game** which is defined by *difference equations* instead of *differential equations*.
- This formulation can be extended to multiple players easily.

Example: Pursuit-evasion

- Consider an **airplane**

$$\dot{x}^1(t) = f^1(x^1(t), u^1(t))$$

and a **missile** chasing the airplane

$$\dot{x}^2(t) = f^1(x^2(t), u^2(t))$$

- That is,

$$\dot{x} = \begin{bmatrix} \dot{x}^1(t) \\ \dot{x}^2(t) \end{bmatrix} = f(x(t), u^1(t), u^2(t)) = \begin{bmatrix} f^1(x^1(t), u^1(t)) \\ f^2(x^2(t), u^2(t)) \end{bmatrix}$$

- Define

$$T(x) = \min\{t \mid x^1(t) = x^2(t)\}, \quad T(x) = \infty \text{ if } x^1(t) \neq x^2(t) \text{ for all } t.$$

- Let us define **utilities** as

$$L^1(u^1, u^2) = T(x) \quad L^2(u^1, u^2) = -T(x)$$

Types of differential games: **Information patterns**

Information pattern

- $\eta^i(t)$: the information available to player i at time t.

Open-loop information pattern

$$\eta^i(t) = \{x_0\}, \quad t \in [0, T].$$

- Each player observes the initial condition of others and picks an open-loop control: $u^i(t) : [0, \infty) \to U^i$.
- During the evolution of the system, the players can not change their controls.

Closed-loop information pattern

$$\eta^i(t) = \{x(t'), 0 \leq t' \leq t\}, \quad t \in [0, T]$$

- Each player picks a closed loop control (that depends on the trajectory of the system, i.e., the other player's control inputs): $\Gamma^i(t, x) : [0, \infty) \times X \to U^i$.
- That is, player's can adjust their controls depending on the state of the system.

Types of differential games: **Payoff structures**

Zero-sum games

Payoffs of the players sum up to zero (or, equivalently a constant), i.e.,

$$L^1(u^1, u^2) + L^2(u^1, u^2) = 0.$$

This can be extended to multiple players easily.

- Examples of zero-sum games:
  - Pursuit-evasion, dog fight (?).
- Generally, management science examples are non-zero sum.
  - Markets (determine market clearing prices), Choosing divident rates (to keep shareholders happy), Supply chain management (game against demand rates).

Nash equilibrium concept

No player can improve payoff by unilaterally changing her strategy.

(u^{1*}, u^{2*}) is a Nash equilibrium point if

$$L^1(u^{1*}, u^{2*}) \geq L^1(u^1, u^{2*}), \qquad \text{for all } u^1.$$

$$L^2(u^{1*}, u^{2*}) \geq L^2(u^{1*}, u^2), \qquad \text{for all } u^2.$$

Nash equilibrium concept can be extended to multiple players easily.

- Most markets end up in a Nash equilibrium.
  - No company can improve payoff (aggregate gains) by unilaterally changing strategy (production rates).

Types of differential games: **Equilibria concepts**

Saddle-point Equilibrium

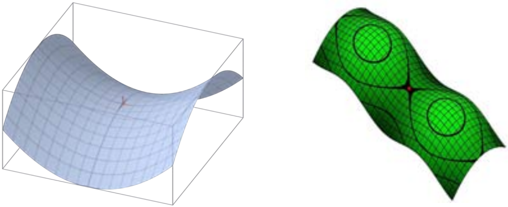
Saddle-point Equilibria Concept

- Saddle-point equilibrium arises in zero-sum differential games.
- Assume there is a single payoff function: $J(u^1, u^2)$.
- Player 1 wants to maximize, Player 2 wants to minimize.

(u^{1*}, u^{2*}) is a saddle point equilibrium point if

$$J(u^1, u^{2*}) \leq J(u^{1*}, u^{2*}) \leq J(u^{1*}, u^2).$$

Note that this can **not** *be extended to multiple players.*



Images are in the public domain. Source: Wikipedia.

Stackelberg equilibrium

One player is the leader announces her strategy first, the followers play accordingly.

- From the leader's point of view:

$$\max_{u^1} \min_{u^2} J(u^1, u^2).$$

- Most markets works according to this rules.
  - Coca Cola sets the price, all others follow.

Open-loop vs. Closed-loop

Lady in the lake

A lady is swimming in a circular-shaped lake. Right when she is in the middle a man comes nearby with the intention of catching her when she comes out.

- The man can not swim.
- The lady can swim slower than the man can run
- The lady can run faster than the man.

Man wins if he captures the lady; lady wins if she escapes.

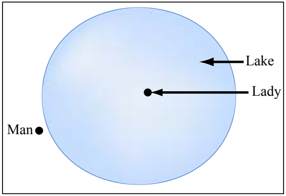


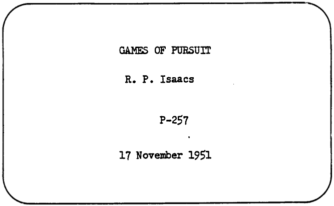
Image by MIT OpenCourseWare.

- In this case, open-loop strategies do not make sense (at least for the man).
- What is a closed-loop strategy for the lady to wins?

Homicidal Chauffeur

A homicidal driver wants to kill a pedestrian. The pedestrian is slow but much more agile.

- Driver is modeled by a Dubins' car.
- The pedestrian is a single integrator with bounded velocity.



GAMES OF PURSUIT

R. P. Isaacs

P-257

17 November 1951

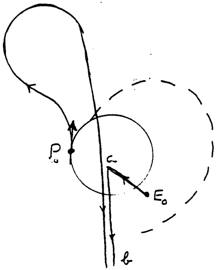


Fig. 1

RAND Corporation, Games of Pursuit, P-257, 1951. Reprinted with permission.

Computational Methods

- **Direct methods**:
  - Formulate a mathematical program and solve.
  - How shall we handle the min-max type of objective function?
  - Bilevel programming is one promising approach.

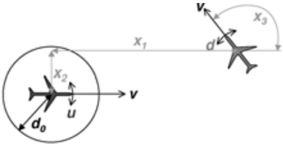
Computational Methods

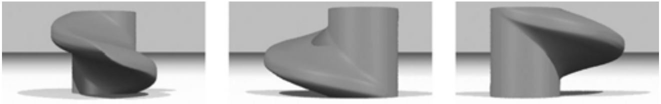
- **Indirect methods**:
  - Using necessary and sufficient conditions, write down a partial differential equation (PDE) that the solution must satisfy.
  - Solve this PDE using level sets, multiple-shooting, collocation, etc.

Computational Techniques for the Verification of Hybrid Systems

CLAIRE J. TOMLIN, IAN MITCHELL, ALEXANDRE M. BAYEN, AND MEEKO OISHI

PROCEEDINGS OF THE IEEE, VOL. 91, NO. 7, JULY 2003





Source: Figures 4 and 6. Tomlin, Claire, Ian Mitchell, Alexandre Bayen, and Meeko Oishi. "Computational Techniques for the Verification of Hybrid Systems." *Proceedings of the IEEE* 91, no. 7 (2003): 986-1001.

Incremental Sampling-based Methods

- Consider a two-player zero-sum pursuit-evasion game

$$\dot{x}_e = f_e(x_e(t), u_e(t)), \qquad \dot{x}_p = f_p(x_p(t), u_p(t)),$$

  where e is the evader and p is the pursuer.

$$\frac{d}{dt}x(t) = \frac{d}{dt}\begin{bmatrix} x_{\rm e}(t) \\ x_{\rm p}(t) \end{bmatrix} = f(x(t), u(t)) = \begin{bmatrix} f_{\rm e}\left(x_{\rm e}(t), u_{\rm e}(t)\right) \\ f_{\rm p}\left(x_{\rm p}(t), u_{\rm p}(t)\right) \end{bmatrix}, \quad \text{for all } t \in \mathbb{R}_{\geq 0},$$

- Define
  (i) $X_{\rm goal}$: **goal region**,
  (ii) $X_{\rm obs,i}$: **obstacle region** for both players $i \in \{e, p\}$,
  (iii) $X_{\rm capt}$: **capture set**.

- Define **terminal time** of the game

$$T = \min\{t \in \mathbb{R}_{\geq 0} : x(t) \in X_{\rm goal} \cup X_{\rm capt}\}$$

- Define the **payoff function**

$$L(u_e, u_p) = \begin{cases} T, & \text{if } x(T) \in X_{\rm goal}; \\ \infty, & \text{otherwise.} \end{cases}$$

  Evader tries to minimize, pursuer tries to maximize.

Incremental Sampling-based Methods

Problem description

- **Open-loop information structure**:
  - The players pick open-loop controls and let the dynamical system evolve.
- **Stackelberg equilibrium**:
  - The evader picks her strategy first, the pursuer observes the evader and picks his strategy accordingly.
- We can think of this as an unbalanced *information structure*:
  - Evader's information structure: open-loop
  - Pursuer's information structure: closed-loop
- Also, assume that the pursuer is in a stable equilibrium.
- **A motivating example**: Aircraft avoiding missiles.

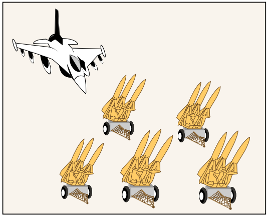


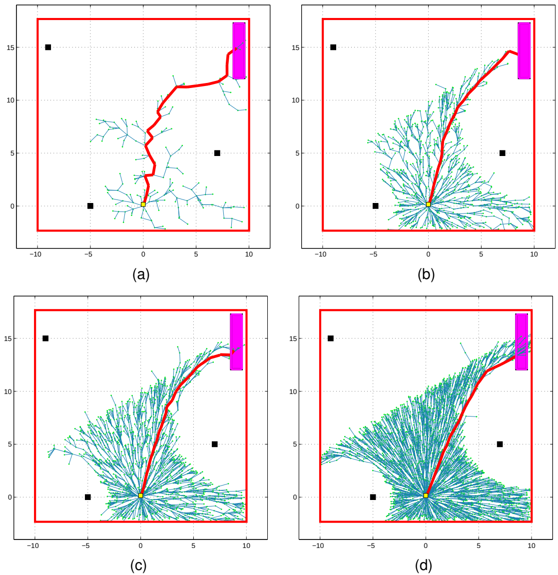
Image by MIT OpenCourseWare.

- Missiles detected by the satellite, but not directly observed by the airplane.
- The airplane must find a safe way through the field.

Incremental Sampling-based Methods

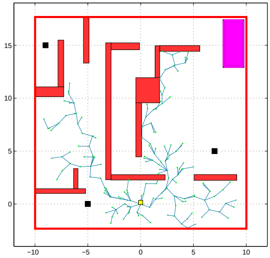
The Algorithm

- We will use incremental sampling-based motion planning methods. In particular, the RRT* algorithm.
- Let `EvaderTree` and `PursuerTree` denote the tree of feasible trajectories maintained by the evader and the pursuer, respectively.
- `GrowEvaderTree` adds one vertex to `EvaderTree` and returns this vertex.
- **Algorithm**:
  1. Initialize `EvaderTree` and `PursuerTree`.
  2. **For** $i := 1$ to n **do**
  3. $\quad x_{\mathrm{new},e} \leftarrow$ `GrowEvaderTree`.
  4. $\quad$ **If** $\{x_p \in \texttt{PursuerTree} \mid \|x_{\mathrm{new},e} - x_p\| f(i), \; (x_{\mathrm{new},e}, x_p) \in X_{\mathrm{capt}}\} \neq \emptyset$ **then**.
  5. $\qquad$ delete $x_{\mathrm{new},e}$.
  6. $\quad$ **EndIf**
  7. $\quad x_{\mathrm{new},p} \leftarrow$ `GrowPursuerTree`.
  8. $\quad$ delete $\{x_e \in \texttt{EvaderTree} \mid \|x_e - x_{\mathrm{new},p}\| \leq f(i), \; (x_{\mathrm{new}}, x_p) \in X_{\mathrm{capt}}\}$
  9. **EndFor**
- For computational efficiency pick $f(i) \approx \frac{\log n}{n}$.

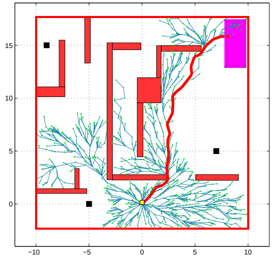


(a)
(b)
(c)
(d)

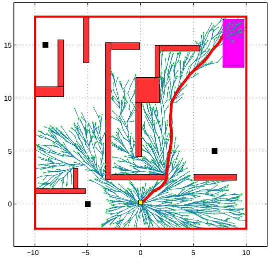
Incremental Sampling-based Methods



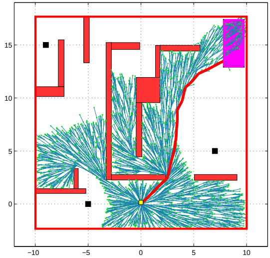
(e)



(f)

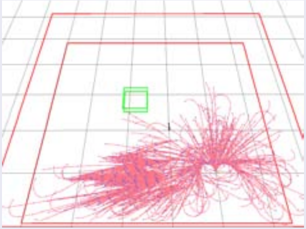


(g)

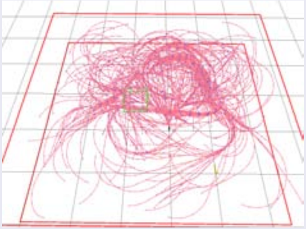


(h)

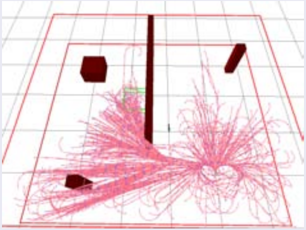
Incremental Sampling-based Methods



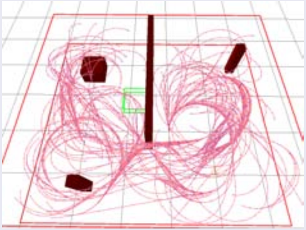
Evader



Pursuer



Evader



Pursuer

Incremental Sampling-based Methods

Probabilistic Soundness

The probability that the solution returned by the algorithm is sound converges to one as the number of samples approaches infinity.

Probabilistic Completeness

The probability that the algorithm returns a solution, if one exists, converges to one as the number of samples approaches infinity.

- The algorithm is incremental and sampling-based:
- An approximate solution is computed quickly and improved if the time allows.
- The approach is amenable to real-time computation,
- Also, computationally effective extensions to high dimensional state-spaces,
- May be valuable in online settings.

Conclusions

In this lecture, we have studied **dynamic games**:

- **Description of time**: Discrete-time, Continuous-time.
- **Information patterns**: Open-loop, Closed-loop (feedback).
- **Payoff structures**: Zero-sum, Nonzero-sum games.
- **Equilibrium concepts**: Nash, Saddle-point, and Stackelberg.
- **Simple examples**: Lady in the lake, Homicidal chauffeur.
- **Numerical solutions**: Direct methods, Indirect methods.
- **Incremental sampling-based algorithms**

16.410 / 16.413 Principles of Autonomy and Decision Making

Fall 2010