

## MIT ES.S20 Lecture 10: Differential Games, Physics, and Control

### Lecture Outline

1. Pursuit Games
2. Optimal Play for the HC
3. Population and Replication Dynamics
4. Statistics Mechanics and Tournaments
5. Resource Flow and Multiagent Systems
6. Game Patterns
7. Variance and Sampling

### Pursuit Games

Demonstration: Car on a grid. One player cannot go backward or turn left.

- Classical Problem: the Homocidal Chauffeur

The control variables are  $\phi(t)$  and  $\psi(t)$

$$\begin{aligned} dx/dt &= \text{Rotation of pedestrian about } (R/\phi, 0) + \text{motion of person in ortho. direction} \\ &= -w_1 * y * \phi / R + w_2 \sin \psi \end{aligned}$$

$$\begin{aligned} dy/dt &= \text{Rotation of pedestrian about } (R/\phi, 0) \\ &\quad + \text{motion of person in parallel direction} \\ &\quad - \text{motion of car in parallel direction} \\ &= w_1 * x * \phi / R + w_2 \cos \psi - w_1 \end{aligned}$$

- Classical Problem: the Rocket Pursuit Game
- Classical Problem: Princess and Monster
- Algorithmic Problem: Graph Clearing

It's fairly easy to see that the Princess and Monster problem can be generalized to an arbitrary number of pursuers and an arbitrarily large graph. In-class problem: try to play the game on paths, stars, etc. How many pursuers do we need? How long does it take?

In general, this problem is NP-Complete and we can reduce from the *partition problem* (Can we partition a string of integers into equally large sets? This is a special case of subset sums)

- Formulating Problems: Games of kind and degree

$$\begin{aligned}\text{Payoff} &= \text{Terminal Payoff} + \text{Integral Payoff} \\ &= H(s) + \int G(x, \phi, \psi) dt\end{aligned}$$

Purely terminal payoff gives a *game of kind* and purely integral payoff gives a *game of degree*

## Optimal Play for the Homocidal Chauffeur

- Optimizing the Control Functions

First, we can define a *value* of the differential game so that  $V$  is the payoff with optimal play at a given point in space. Also, define our EOM in terms of some  $f(x, \phi, \psi) = dx/dt$ .

After some differential time segment, we must have that the integral payoff balanced the change in  $V$  if  $V$  is to make sense as a value. This gives the first PDE:

$$\sum_i V_i f_i(x, \phi, \psi) + G(x, \phi, \psi) = 0$$

If we differentiate this equation by some direction  $x_k$ , we obtain by elementary calculus that

$$dV_k/dt = -G_k - \sum_i V_i f_{ik}$$

Where the time-derivative in  $V_k$  comes from the fact that  $dV_i/dV_k = dV_k/dV_i$  and  $f_i$  looks like  $dx_i/dt$ . This gives another set of differential equations that we can solve for the value with.

To solve the actual game, we can reverse time to  $dT = -dt$  and use the set of equations

$$\begin{aligned}dx_k/dT &= -f_k \\ dV_k/dT &= \sum_i V_i f_{ik} + G_k\end{aligned}$$

These can be integrated in reverse to derive the value at all points. These look like the *Hamilton-Jacobi Equations*: we've just derived a potential so that the control paths minimize it. There is an entire field organized around this sort of problem, *Optimal Control Theory*.

- The Homocidal Chauffeur

Draw the Barrier: the surface of starting points for which the outcome is neutral (payoff zero). This is a convenient thing to do because it is suggestive that we

can describe the regions where capture will occur. We can draw barrier diagrams for the Homocidal Chauffeur and use them to characterize it as a game.

This also is suggestive of the approach of Level Sets, surfaces on which we can win in  $N$  turns.

There's a lot of mathe to get there, but capture will occur iff (for speed ratio  $g = w_2 / w_1$ ),  $1 / R > \sqrt{1-g^2} + \arcsin(g) - 1$

## Population and Replication Dynamics

Demonstration: Choose strategy A or B repeatedly. For whichever one is picked most often, I will "kill" one of the players randomly. The goal is to get killed the least number of times.

- Evolution

Assume that we have a classical dual-strategy payoff matrix model over strategies A and B so that we have the performance (here, fitness):

$$\begin{aligned} f_A &= a x_A + b x_B \\ f_B &= c x_A + d x_B \end{aligned}$$

where  $x_A$  is the proportion of the population using strategy A. The evolutionary dynamics approach is that we can turn this into a system of differential equations:

$$\begin{aligned} dx_A/dt &= x_A (f_A - \langle f \rangle) = x_A (1 - x_A) (f_A - f_B) \\ dx_B/dt &= x_B (f_B - \langle f \rangle) = x_B (1 - x_B) (f_B - f_A) \end{aligned}$$

This, of course, assumes independence and does not *a priori* represent limitations on resources (although I think this can be generally embedded).

- Moran Processes

We can use something that looks like a discrete stochastic process for modeling this system in a finite population: let  $i$  be the number of individuals using A and  $j$  be the number using B.

$$\begin{aligned} P(i \rightarrow i+1) &= a (i-1) + b j \\ P(i \rightarrow i-1) &= c i + d (j-1) \\ P(i \rightarrow i) &= 1 - P(i \rightarrow i+1) - P(i \rightarrow i-1) \end{aligned}$$

Note that similar systems can be formulated on continuous stochastic processes, continuum populations, graphs, etc. This is the simplest dynamic possible.

- Shape of the Process

Define an absorbing state as a state in which everyone goes to A or B. The literature gives the probability that the system is absorbed by B from B as

$$r_{AB} = (1 + \sum_{j=1..i-1} \sum_{k=1..j} (g_k / f_k)) / (1 + \sum_{j=1..N-1} \sum_{k=1..j} (g_k / f_k))$$

This already gives the nontrivial result that the ratio of the probabilities of absorption is  $r_{AB} / r_{BA} = \prod_k (f_k / g_k)$

Example:  $f_A = r$ ,  $f_B = 1$  gives

$$\begin{aligned} p_{AB} &= (1-1/r) / (1-(1/r)^N) \\ p_{BA} &= (1-r) / (1-r^N) \end{aligned}$$

- Selection Dynamics

Define  $\zeta$  as  $d-b$  and  $\xi$  as  $a-c$  being the initial disadvantage of each (the disadvantage of moving to the A or B, respectively). If we define the difference in fitness

$$\begin{aligned} h_i &= \xi' i - \zeta' j \\ \text{for } \xi' &= \xi - (a-d)/N \\ \zeta' &= \zeta + (a-d)/N \end{aligned}$$

The N-dependence means that the characteristics change at the thresholds

$$\begin{aligned} N_1 &= (a-d) / \xi \\ N_2 &= (a-d) / \zeta \end{aligned}$$

$$\begin{aligned} \xi' > 0, \zeta' > 0 &\Rightarrow \text{Bistable} \\ \xi' > 0, \zeta' < 0 &\Rightarrow \text{A-Dominant} \\ \xi' < 0, \zeta' > 0 &\Rightarrow \text{B-Dominant} \\ \xi' < 0, \zeta' < 0 &\Rightarrow \text{Polymorphic} \end{aligned}$$

Conclusion: Spiteful strategy is best with few players

- Evolutionarily Stable Strategy

A strategy is Evolutionarily Stable iff a strategy cannot be invaded by a rare strategy. This is representable for a state of mostly A and some number of B less than  $N_1$ ! This is a special case of a Nash Equilibrium.

- Additional Problems

What happens when there are spatially-dependent resources? (Location Games)

What happens when players can signal one another?

What if the influence of a player's choice is local?

Relevant areas: Statistical Mechanics, Chaos, Dynamical Systems.

## **Statistical Mechanics and Tournaments**

### **Resource Flow and Multiagent Systems**

The bimatrix formalism is not a very good representation of many real-world system because it only deals with linear, first order dynamics. I'd like to also introduce a model that I forgot in my previous lecture: functional flow models.

### **Game Patterns**

- Bounding Surfaces, Visual metaphor
- Tradeoffs between performance domains
- Stochastic Processes