Biomimicry for Optimization, Control, and Automation

Kevin M. Passino

Dept. Electrical Engineering
The Ohio State University
2015 Neil Avenue
Columbus, OH 43210-1272
(614) 292-5716, passino@ee.eng.ohio-state.edu
http://eewww.eng.ohio-state.edu/~passino

THE OHIO STATE UNIVERSITY

Copyright

FORAGING

Focus:

- Bacterial foraging (modeling, simulation, optimization)
- Stability analysis of swarms, swarms approach to cooperative control for robots
- Games, cooperative and competitive foraging
- Intelligent social foraging (planning, attention, learning, communications, multiple agents/cooperation/competition)

Bacterial Foraging as Optimization

Foraging Theory

Elements of Foraging Theory

➙ Foraging theory is based on the assumption that animals search for and obtain nutrients in a way that maximizes their energy intake E per unit time T spent foraging.

$$\frac{E}{T}$$

➙ Foraging is an optimization process created by evolution.

- Gives time for reproduction, etc.

- Foraging is very different for different species.

➙ "Environment" sets pattern of predators/prey (nutrients).

➙ Forager physiological characteristics affect success (also evolve).

➙ Nutrients are distributed in "patches"

➙ Typical foraging decision strategies?

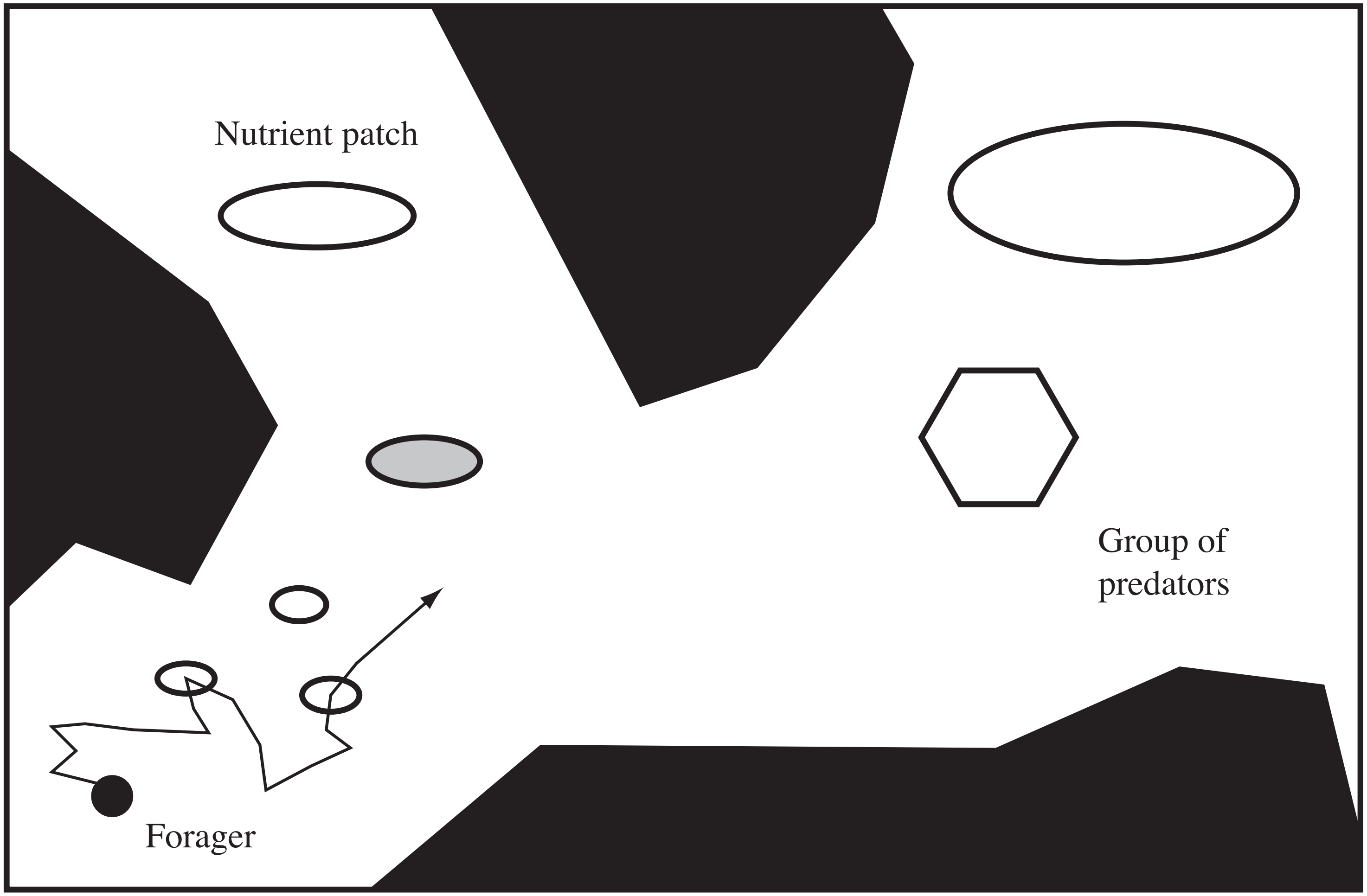

Figure 196: Foraging landscape and scenario.

- Decision-making in foraging is a control strategy for organism guidance.
- "Optimal foraging theory"? Relevance? Optimal guidance strategies!

Social/Intelligent Foraging

- Advantages to "social" foraging? Need communications:
  - Wider area search
  - Access to "information center" for helping with survival.
  - Increased capabilities to cope with larger prey.
  - Protection from predators.
- Think of a group (swarm) as a *single* living creature—emergent intelligence.
- Examples: Pack of wolves hunting, flock of birds, swarm of bees, colony of ants (algorithms here), or school of fish.

➙ “Intelligent foraging” uses rules, planning, attention, learning.

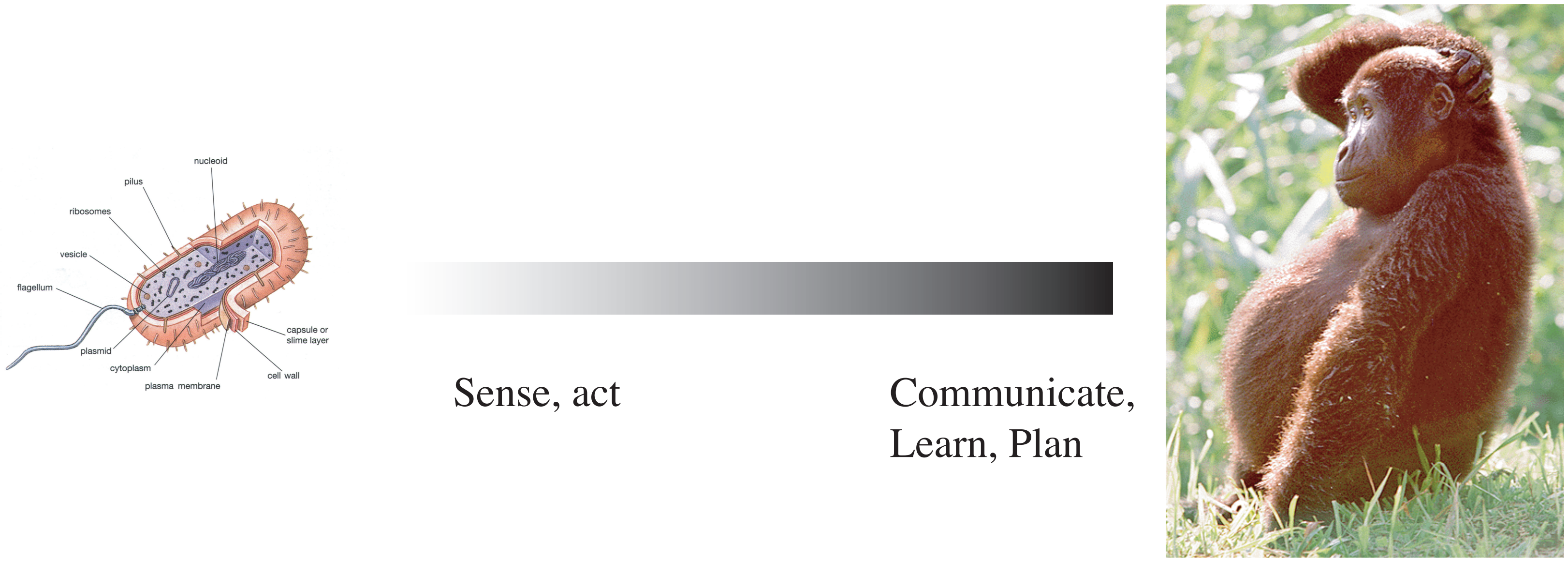

Figure 197: Cognitive spectrum for foraging.

➙ Entire spectrum interesting from an engineering perspective.

- Let’s start at the bottom...

Chemotactic (Foraging) Behavior of ***E. coli***

- *E. coli*: Diameter: $1\mu m$, Length: $2\mu m$

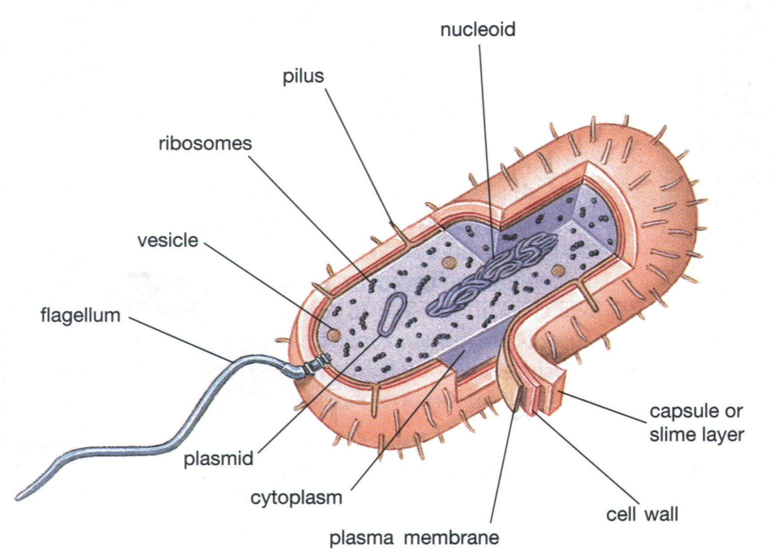

Figure 198: *E. coli* bacterium (from [2]).

- Can reproduce (split) in 20 min.

Motility and Chemotaxis

- Motility via reversible rigid 100 – 200 rps spinning flagella each driven by a biological "motor"

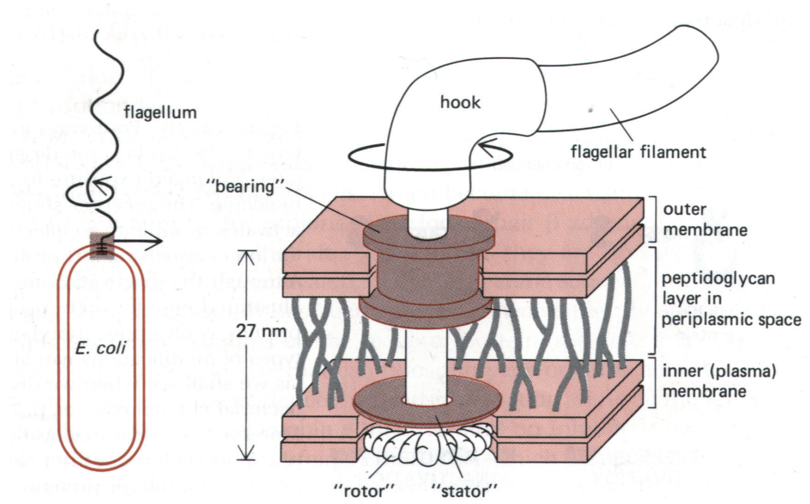

Figure 199: *E. coli* biological "motor" (from [1]).

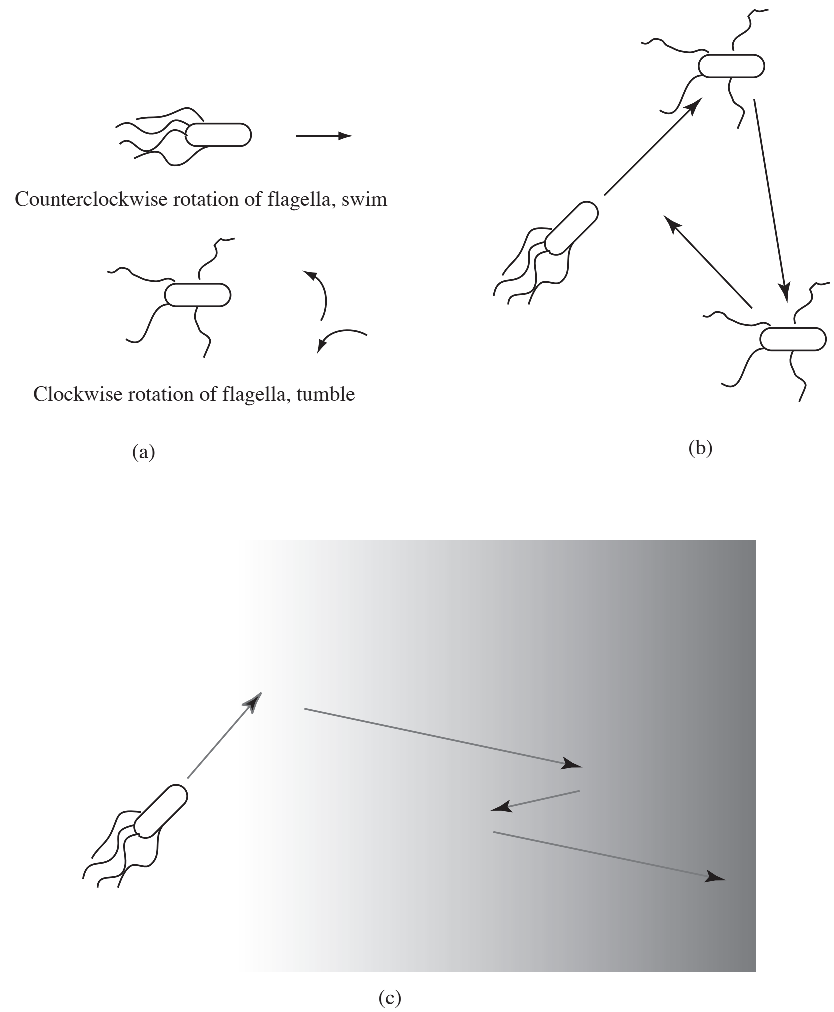

Figure 200: Chemotactic behavior.

Decision Making in Foraging

1. If in neutral medium alternate tumbles and runs
   $\Rightarrow$ Search

2. If swimming up nutrient gradient (or out of noxious substances) swim longer (climb up nutrient gradient or down noxious gradient)
   $\Rightarrow$ Seek increasingly favorable environments

3. If swimming down nutrient gradient (or up noxious substance gradient), then search
   $\Rightarrow$ Avoid unfavorable environments

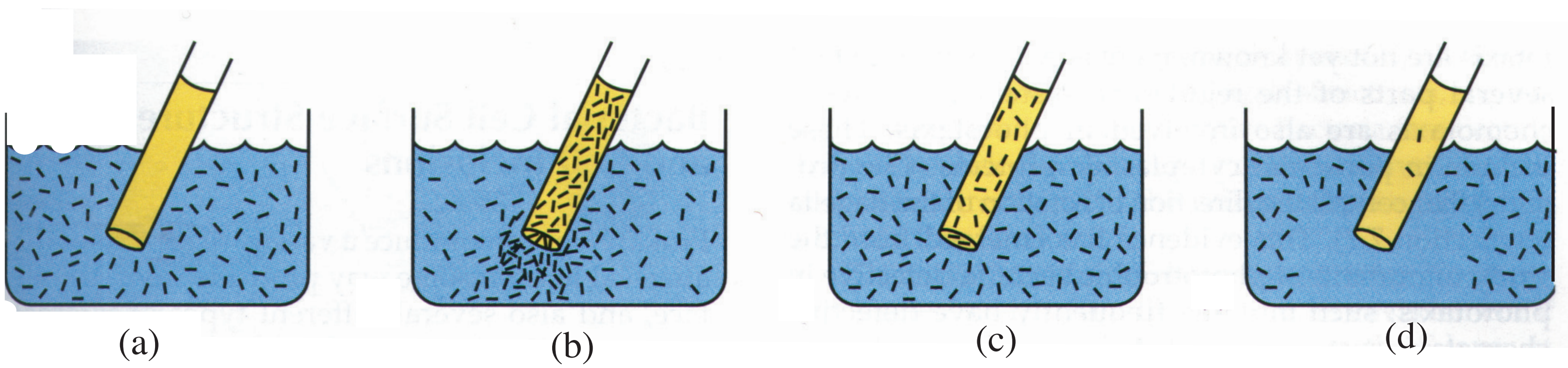

Figure 201: Capillary experiment (from [14]).

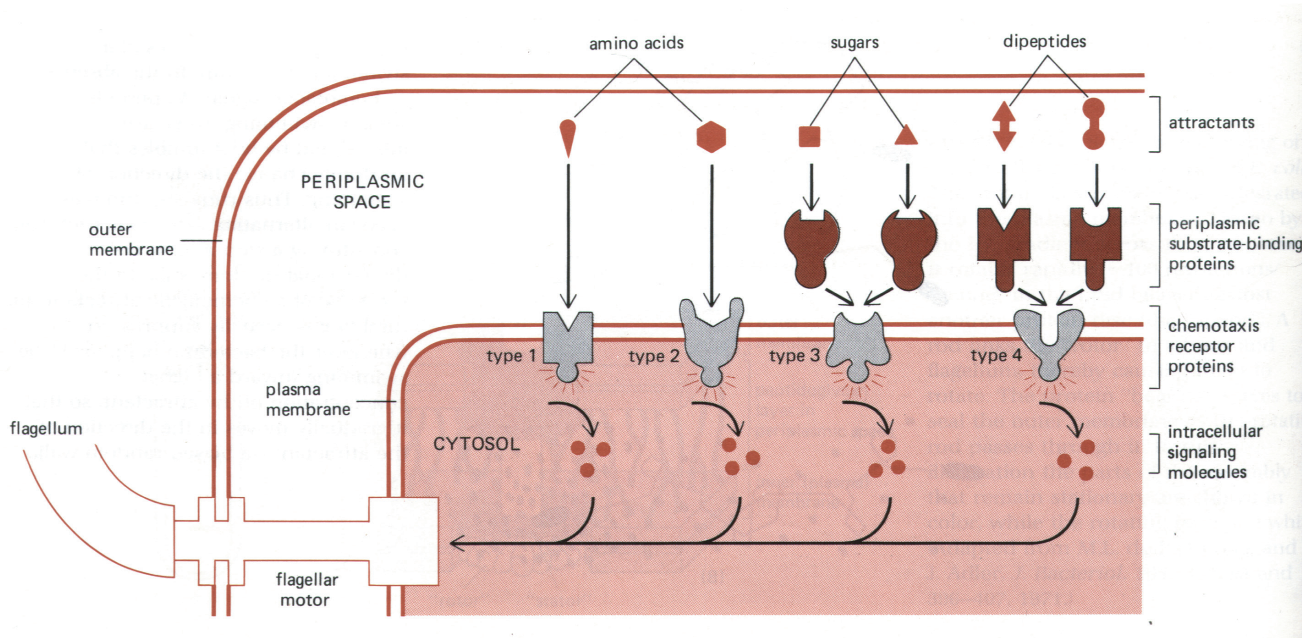

Figure 202: Sensing and control in *E. coli* (from [1]).

- The sensors are very sensitive, and overall there is a "high gain."
- Averages sensed concentrations and computes an approximation to a *time* derivative.

➙ Probably the best understood sensory and decision-making system in biology (understood/simulated at molecular level).

Elimination/Dispersal and Evolution

- Bacteria often killed and dispersed (can be viewed as part of their motility)
  - Mutations in *E. coli* affect, e.g., reproductive efficiency at different temperatures, and occur at a rate of about 10^{-7} per gene, per generation.
  - *E. coli* occasionally engage in a type of "sex" called "conjugation" (Figure 203)

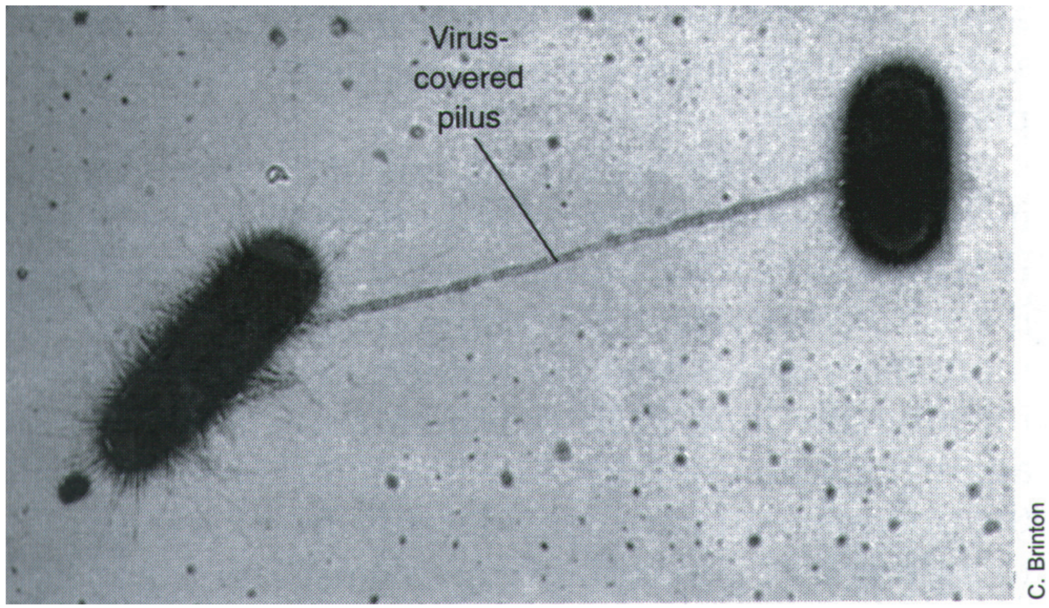

Figure 203: Conjugation in *E. coli* (from [14]).

Other Taxes

1. Change cell shape and number of flagella based on medium!
2. Oxygen (aerotaxis), light (phototaxis), temperature (thermotaxis), magnetic flux lines (magnetotaxis)

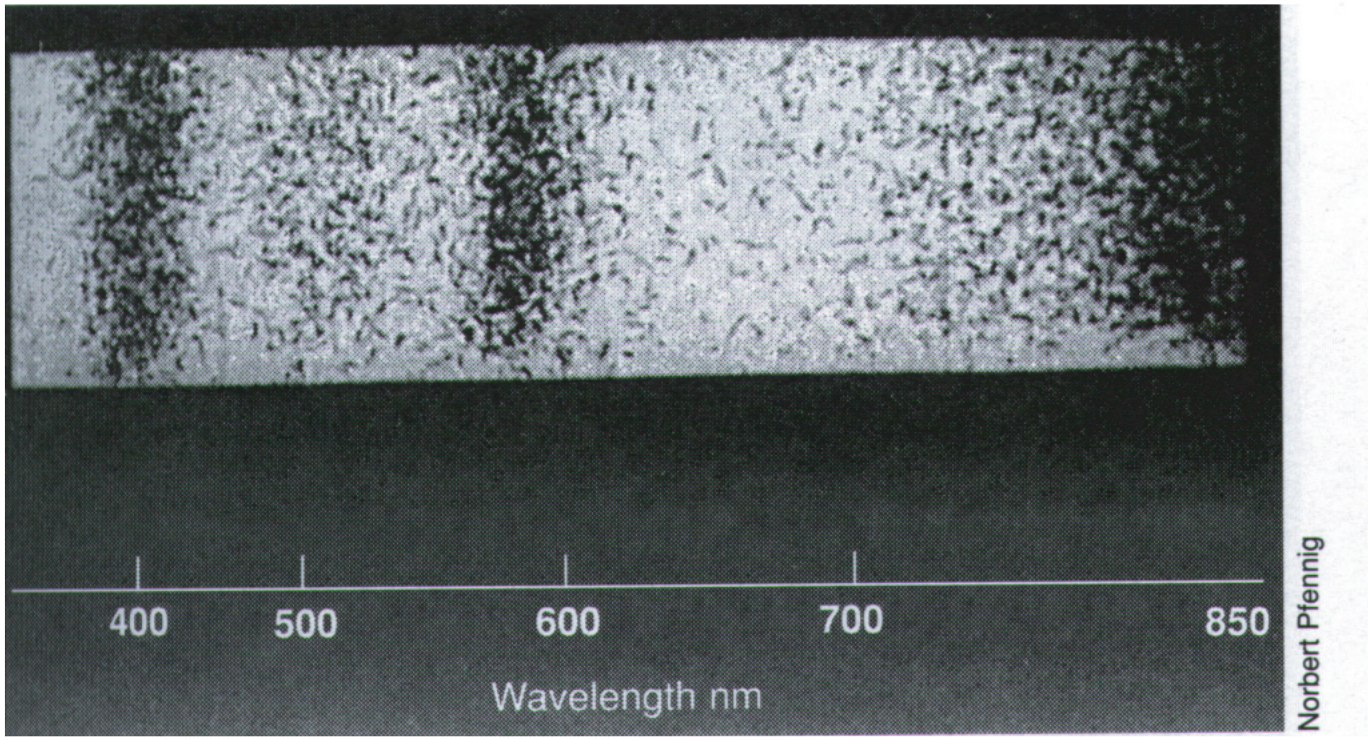

Figure 204: Phototaxis behavior of the phototropic bacterium *Thiospirillum jenense* (from [14]).

Swarms

➙ *E. coli* and *S. typhimurium* can form intricate stable spatio-temporal patterns in certain semi-solid nutrient media

- Radially eat their way through the medium.
- Cell-to-cell attractant signals.
- The bacteria protect each other.

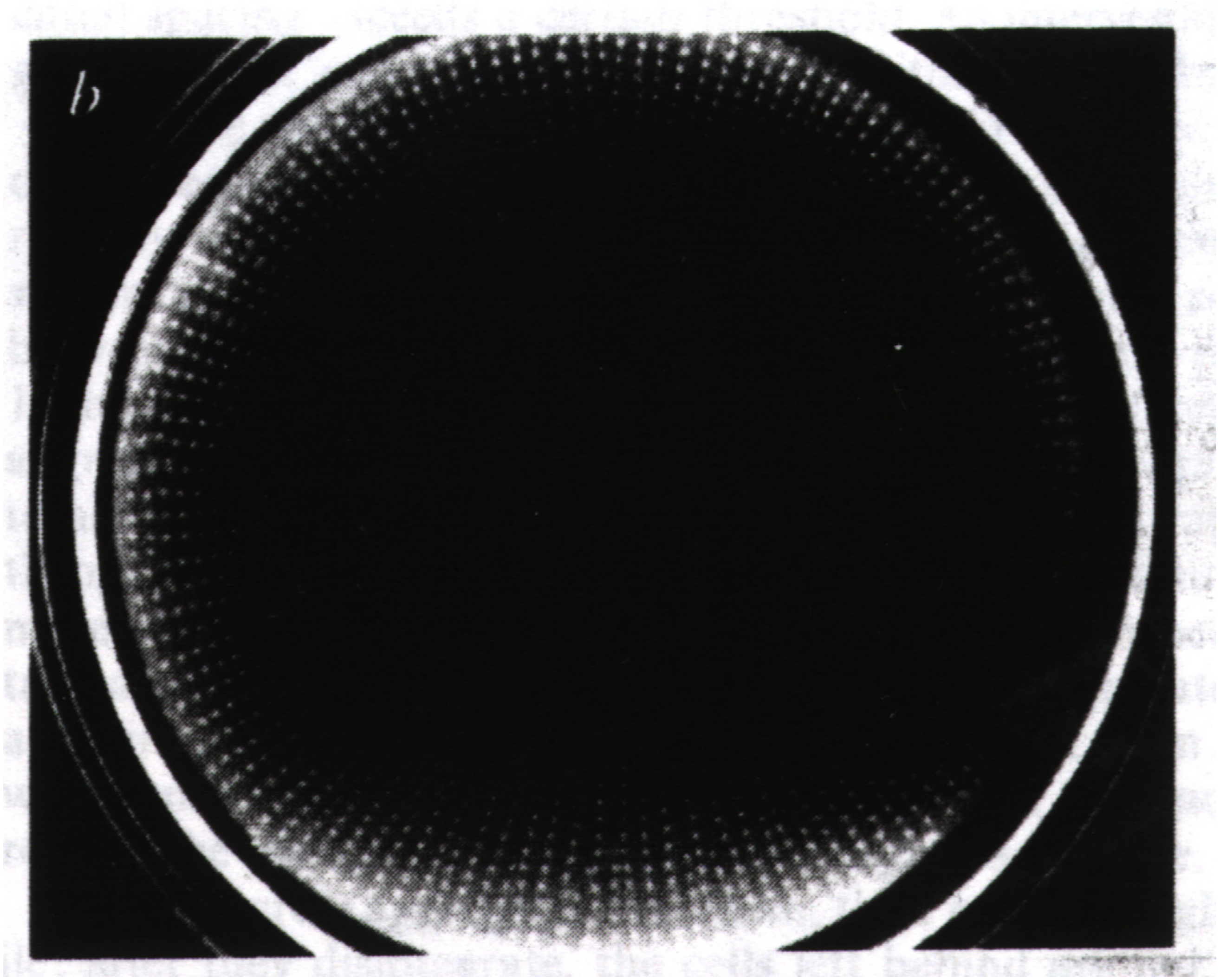

Figure 205: Swarm pattern of *E. coli* (from [3]).

Bacterial Swarm Foraging for Optimization

- Find the minimum of

$$J(\theta),\ \theta \in \Re^p$$

when we do not have $\nabla J(\theta)$.

➙ Suppose θ is the position of a bacterium, and $J(\theta)$ represents an attractant-repellant profile so:

1. $J > 0 \Rightarrow$ noxious
2. $J = 0 \Rightarrow$ neutral
3. $J < 0 \Rightarrow$ food

➙ Let

$$P(j,k,\ell) = \left\{\theta^i(j,k,\ell) | i = 1,2,\ldots,S\right\}$$

be the set of all S bacterial positions at the j^{th} chemotactic step, k^{th} reproduction step, and ℓ^{th} elimination-dispersal event.

- Let $J(i, j, k, \ell)$ denote the cost at the location of the i^{th} bacterium $\theta^i(j, k, \ell) \in \Re^p$.
- Let N_c be the length of the lifetime of the bacteria as measured by the number of chemotactic steps.

➙ To represent a tumble, a unit length random direction, say $\phi(j)$, is generated; then we let

$$\theta^i(j+1, k, \ell) = \theta^i(j, k, \ell) + C(i)\phi(j)$$

so $C(i) > 0$ is the size of the step taken in the random direction specified by the tumble.

➙ If at $\theta^i(j+1, k, \ell)$ the cost $J(i, j+1, k, \ell)$ is better (lower) than at $\theta^i(j, k, \ell)$, then another chemotactic step of size $C(i)$ in this same direction will be taken, and repeat that up to a maximum number of steps, N_s.

➙ Cell-to-cell signaling via an attractant:

1. Attractants are essentially "food" for other cells (chemotactically attracted to it)
2. Use $J^i_{cc}(\theta)$, $i = 1, 2, \ldots, S$, to represent locally secreted food.

- Repel? Via local consumption, and cells are not food for each other. Again, use $J^i_{cc}(\theta)$.
- Example: Consider the $S = 2$ case...

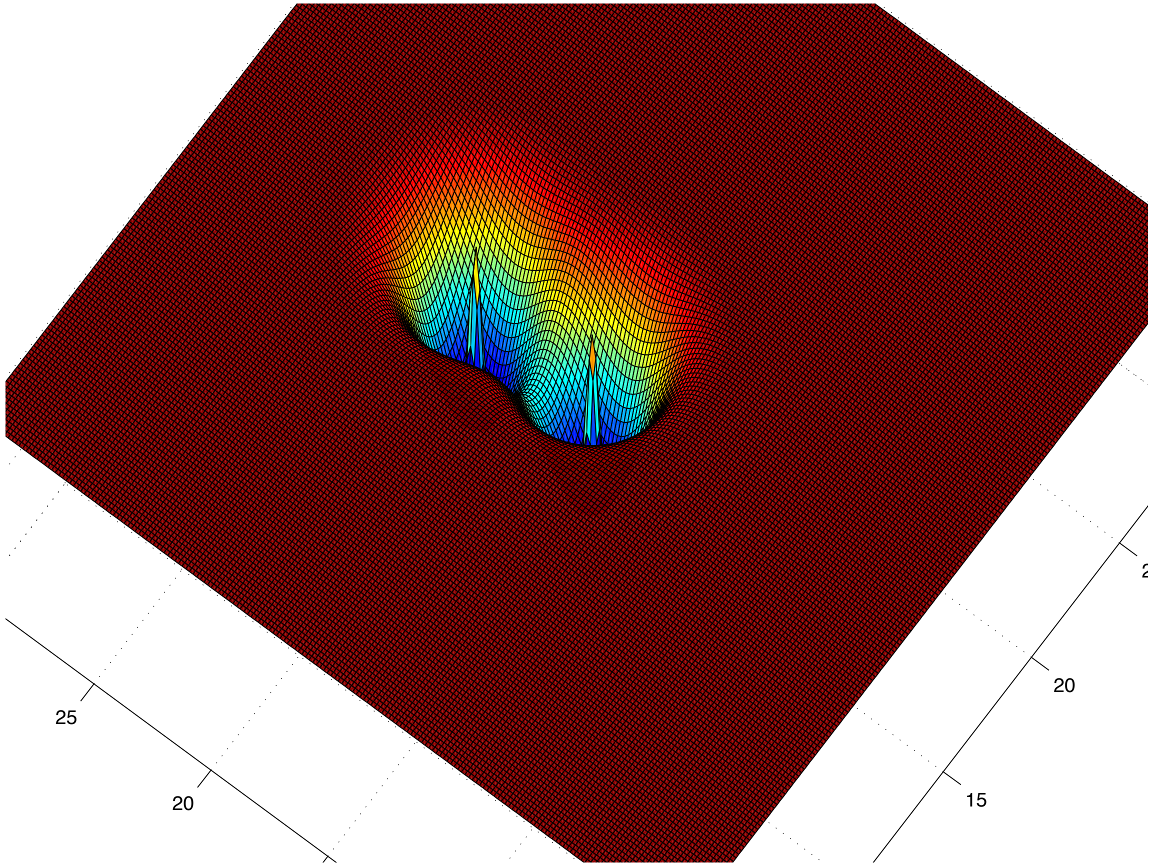

Figure 206: Example cell-to-cell attractant model, $S = 2$.

- For swarming consider minimization of

$$J(i,j,k,\ell) + J_{cc}(\theta)$$

so cells try to find nutrients, avoid noxious substances, and try to move towards other cells, but not too close to them.

- The $J_{cc}(\theta)$ function dynamically deforms the search landscape to represent the desire to swarm.

- Take N_{re} reproduction steps.

- For reproduction, healthiest bacteria (ones that have lowest accumulated cost over their lifetime) split, and then kill other unhealthy half of population.
- Let N_{ed} be the number of elimination-dispersal events (for each one, each bacterium is subjected to elimination-dispersal with probability p_{ed}).
- Biologically valid model? Capturing gross characteristics of chemotactic hill-climbing and swarming.

Example: Function Optimization

- Find minimum of function in Figure 207 ($[15, 5]^\top$ is the global minimum point, $[20, 15]^\top$ is a local minimum).
- Standard ideas from optimization theory can be used to set the algorithm parameters.

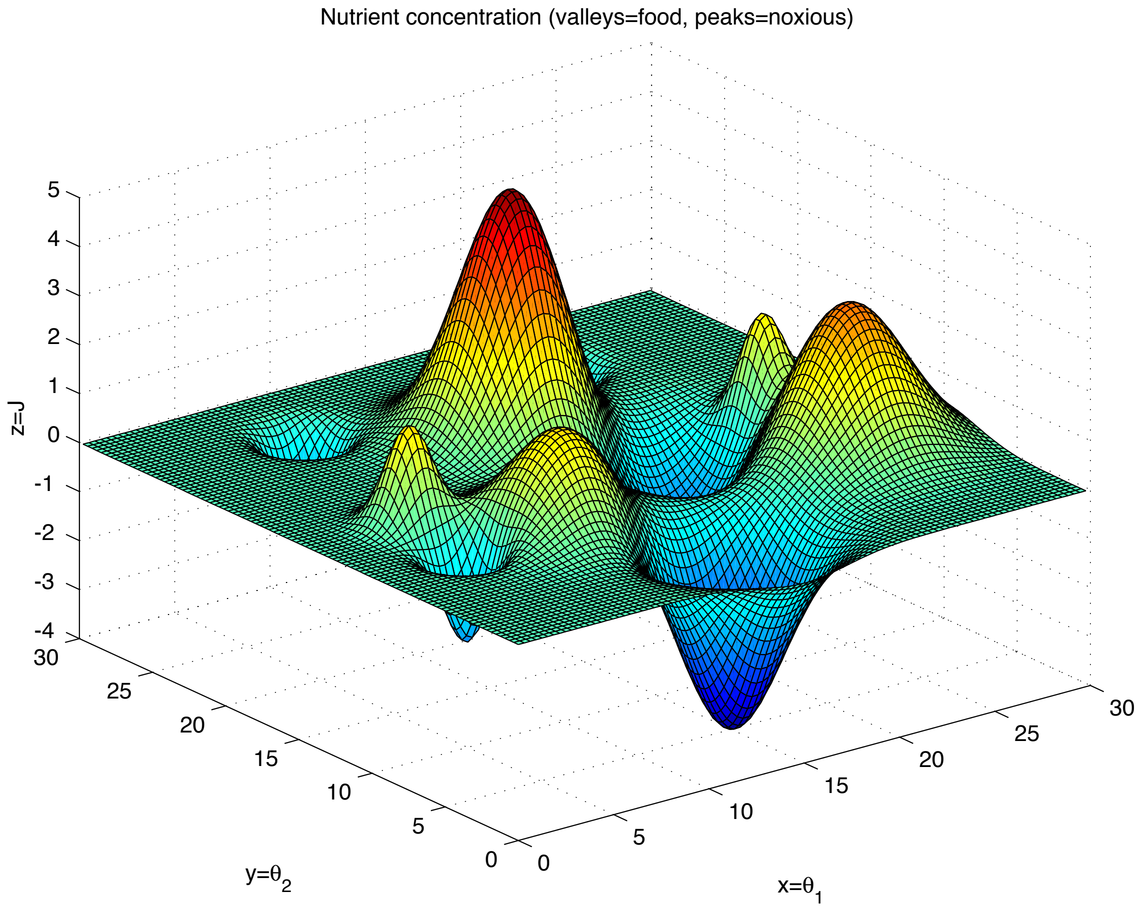

Figure 207: Function with multiple extremum points.

➙ No swarming:

- $S = 50$, $N_c = 100$, $C(i) = 0.1$, $i = 1, 2, \ldots, S$, $N_s = 4$ (a biologically-motivated choice)
- $N_{re} = 4$, $N_{ed} = 2$, $p_{ed} = 0.25$,
- Random initial bacteria distribution.

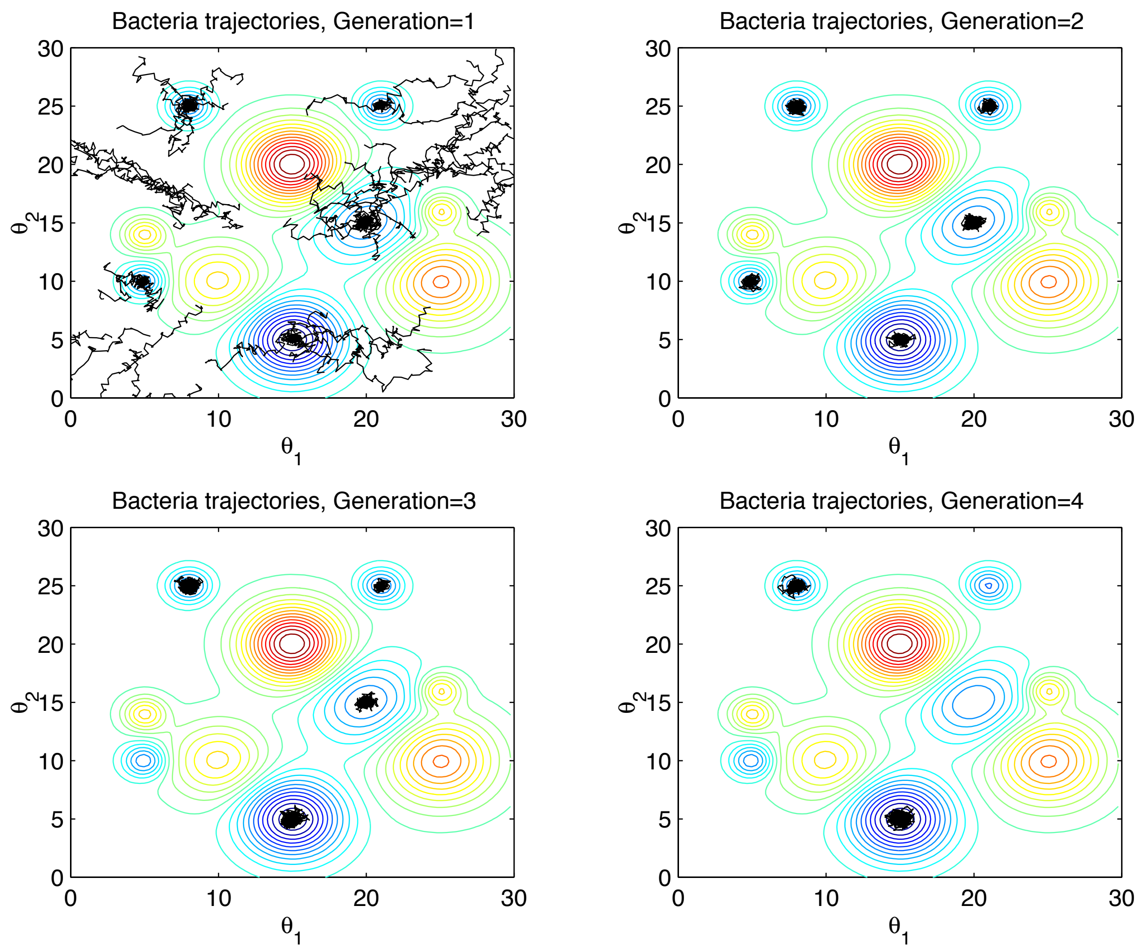

Figure 208: Bacterial motion trajectories, generations 1-4.

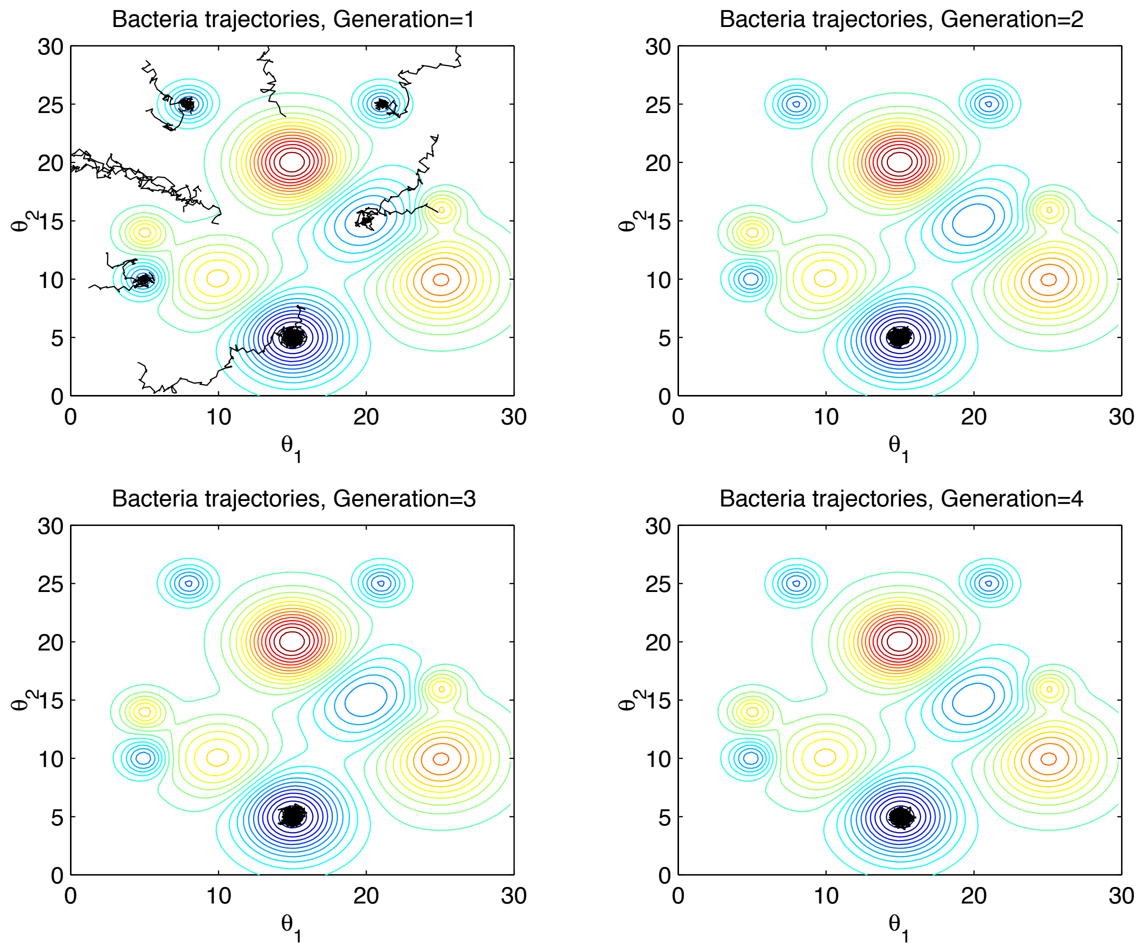

Figure 209: Bacterial motion trajectories, generations 1-4, after an elimination-dispersal event.

➜ Swarm effects:

- Emulate Figure 205 by considering optimization over Figure 210.
- Initially, place all cells at the peak $[15, 15]^\top$.

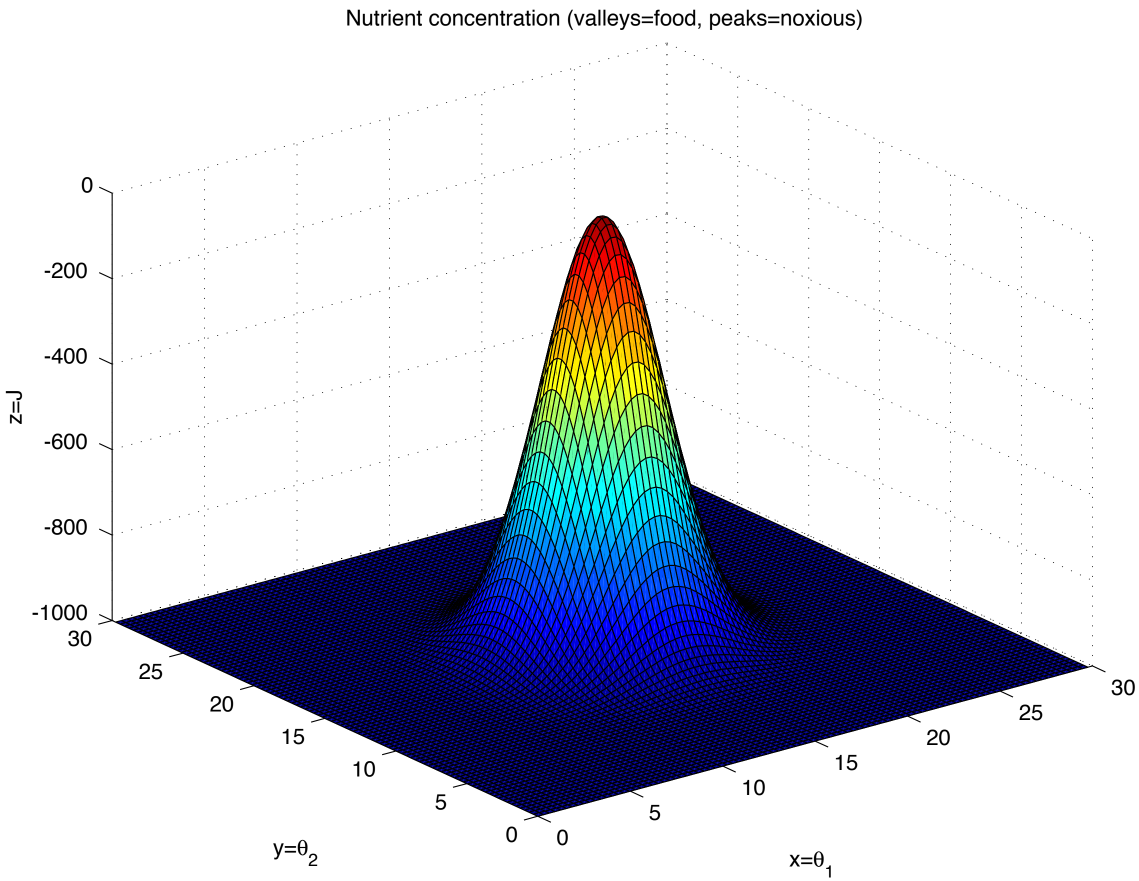

Figure 210: A nutrient surface for testing swarming.

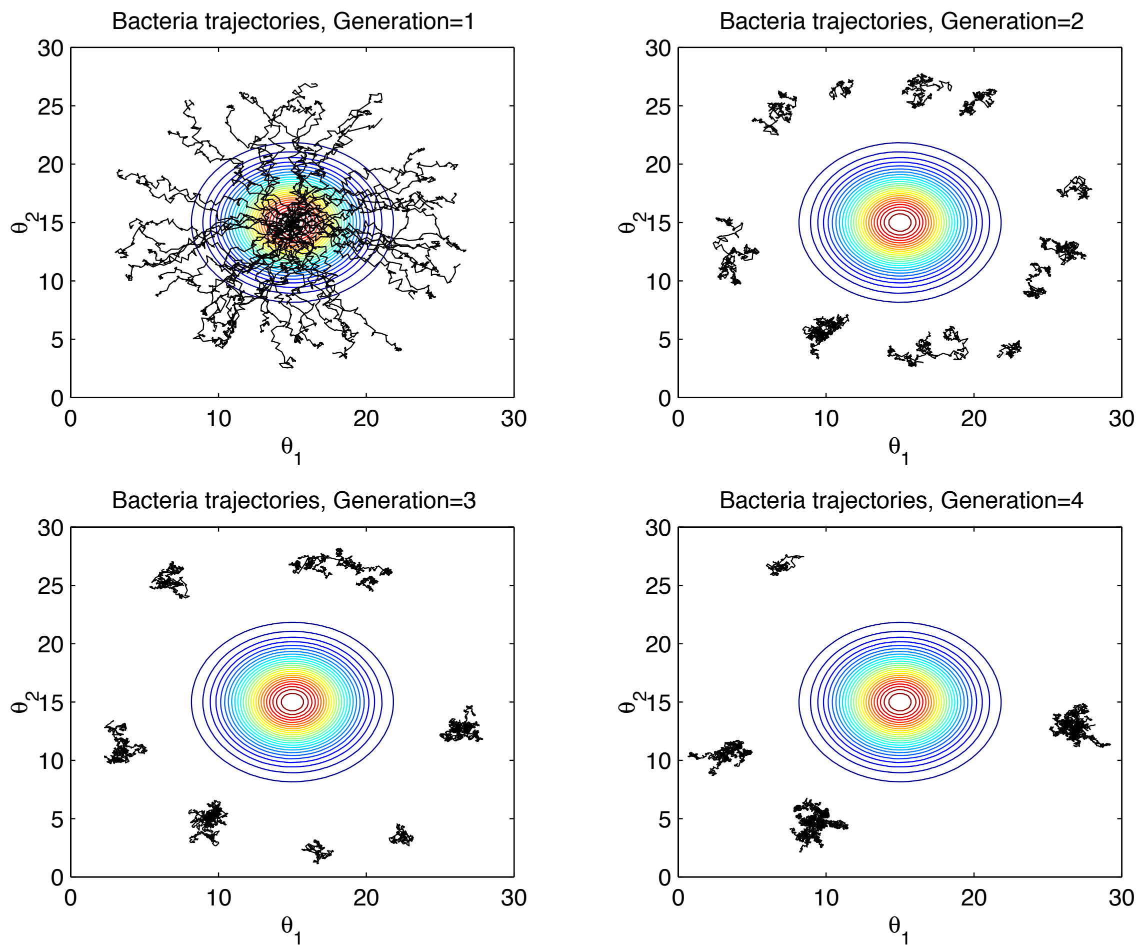

Figure 211: Swarm behavior of *E. coli* on a test function.

Social Bacterial Foraging: M. xanthus

- Exotic social motile behavior *Myxobacteria* ("slime bacteria")

  Gliding, Slime Trails, and Swarm Foraging Behavior

- Lives in soil and on leaves on forest floors, is a "gliding" bacterium.
  - Single bacterium is isolated in an appropriate nutrient-rich environment moves forward and backward and it seems to make no progress.
- As it moves it lays down a slime trail, others tend to follow.
- Also engage in social foraging (via mutations "social motility" and "adventurous motility")

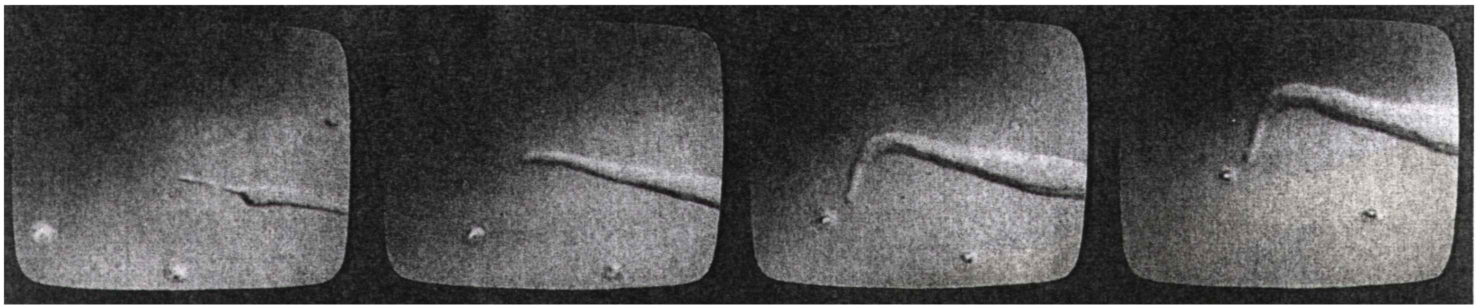

Figure 212: Swarm chemotactic behavior of *M. xanthus* (figure taken from [16], but the images were made by M. Dworkin of the Univ. of Minnesota).

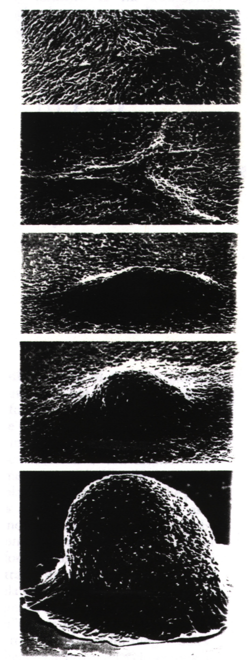

Figure 213: Formation of a mound of *M. xanthus*. In the top frame aggregation is just starting via cell alignment, then via aggregation the mound grows as shown in subsequent frames (figure taken from [13]).

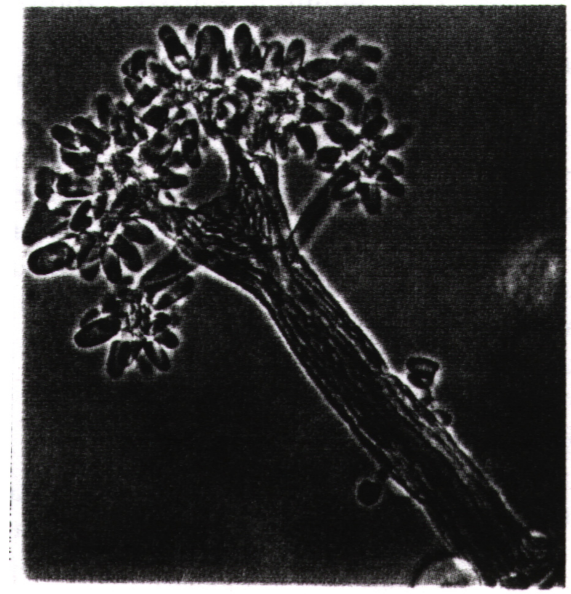

Figure 214: Fruiting body of *Chondromyces crocatus*. If you watch a movie of the formation of this fruiting body it resembles a growing plant (figure taken from [13], but image is from H. Reichenbach, Inst. for Biotechnical Research, Germany).

Other Types of Swarming/Complex Bacterial Patterns

- Luminous bacteria symbiotic with squid
- Soil-dwelling *streptomycete* colonies can grow a branching network of long fiber-like cells that can penetrate and degrade vegetation
- *Proteus mirabilis* swimmer/swarmer capabilities

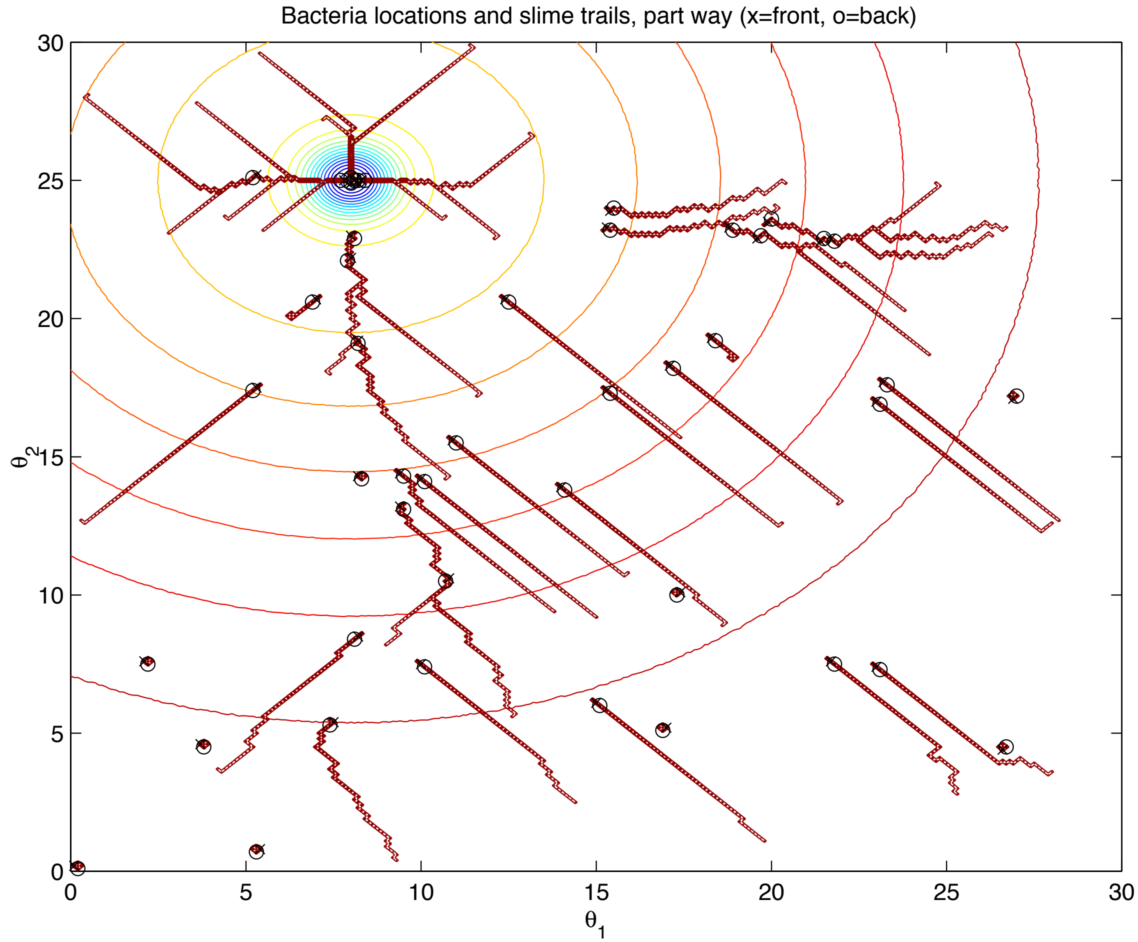

Figure 215: Cellular automaton: *M. xanthus* bacterial swarm foraging algorithm results, after 50 chemotactic steps.

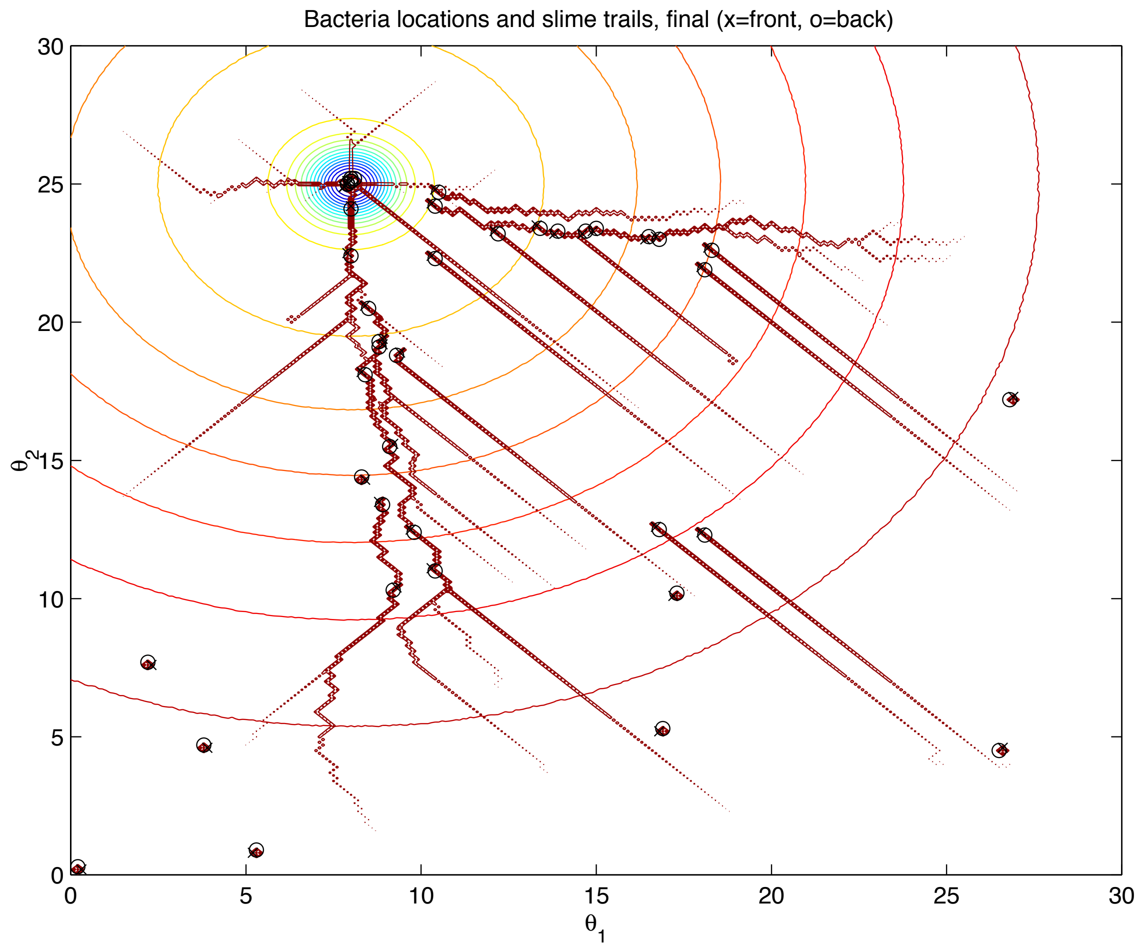

Figure 216: *M. xanthus* bacterial swarm foraging algorithm results, after 100 chemotactic steps.

Stability Analysis of Social Foraging Swarms

- Examples: Bees, bacteria, wildebeests, birds, robots, etc.

Swarm and Environment Models

Agent Dynamics and Communications

→ N "agents," point mass dynamics,

$$\begin{aligned} \dot{x}^i &= v^i \\ \dot{v}^i &= \frac{1}{M_i} u^i \end{aligned} \tag{93}$$

where $x^i \in \Re^n$ is the position, $v^i \in \Re^n$ is the velocity, M_i is the mass, and $u^i \in \Re^n$ is the (force) control input for the i^{th} agent.

- Nonlinear and stochastic models can be used.
- Consider $n = 3$
- Could use "communication topology" (G, A) where

$G = \{1, 2, \ldots, N\}$ is a set of nodes (the agents) and

$$A = \{(i, j) : i, j \in G, i \neq j\}$$

- Fixed? Time-varying? Link characteristics?

Agent to Agent Attraction and Repulsion

➙ Agent seeks to be in a position that is "comfortable" relative to its neighbors

➙ Attraction/repulsion terms: Can have linear, nonlinear, local/global, static/dynamic.

➙ Balance between attraction and repulsion (a basic concept in swarm dynamics that is sometimes referred to as an "equilibrium" even though it may not be one in the stability-theoretic sense)

- Include in u^i for each agent:

➙ Attraction: Linear attraction

$$-k_a \left(x^i - x^j\right)$$

where $k_a > 0$ is a scalar that represents the strength of attraction.

– Repulsion:

 * Seek a "comfortable distance": A term in u^i with

$$\left[-k\left(||x^i - x^j|| - d\right)\right]\left(x^i - x^j\right)$$

 * Repel when close: A term in u^i

$$k_r \exp\left(\frac{-\frac{1}{2}\|x^i - x^j\|^2}{r_s^2}\right)\left(x^i - x^j\right) \tag{94}$$

 * Hard repulsion for collision avoidance: Term like

$$\left[\max\left\{\left(\frac{a}{b||x^i - x^j|| - w} - \epsilon\right), 0\right\}\right]\left(x^i - x^j\right) \tag{95}$$

Environment and Foraging

- Agents move over a "resource profile" (e.g., nutrient profile) $J(x)$, where $x \in \Re^n$.
- Agents move in the direction of the negative gradient of $J(x)$

$$-\nabla J(x) = -\frac{\partial J}{\partial x}$$

  in order to move away from bad areas and into good areas of the environment (e.g., to avoid noxious substances and find nutrients).
- Examples:
  - Plane: In this case we have $J(x) = J_p(x)$ where

$$J_p(x) = R^\top x + r_o$$

    where $R \in \Re^n$ and r_o is a scalar. Here, $\nabla J_p(x) = R$.

– Quadratic: In this case we have $J(x) = J_q(x)$ where

$$J_q(x) = \frac{r_m}{2}\|x - R_c\|^2 + r_o$$

where r_m and r_o are scalars and $R_c \in \Re^n$. Here, $\nabla J_q(x) = r_m\,(x - R_c)$.

- ➙ Agents can sense profile, with noise
- ➙ Multi-objective agents
- ➙ Can sense points or regions of environment
- ➙ Consumption, time-varying profile

Stability Analysis of Swarm Cohesion Properties

Sensing, Noise, and Error Dynamics

- Let

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x^i$$

be the center of the swarm and

$$\bar{v} = \frac{1}{N}\sum_{i=1}^{N} v^i$$

be the average velocity (vector) which we view as the velocity of the group of agents.

- Each agent can sense the distance from itself to $\bar{x}$, and the difference between its own velocity and $\bar{v}$.
- Each agent knows its own velocity, but not its own position.
- Sensors for this in nature? Robotics?

➙ Agent objective: Move so as to end up at or near $\bar{x}$ and have its velocity equal to $\bar{v}$

- $\bar{x}$ and $\bar{v}$ are generally time-varying

➙ Error system:

$$e_p^i = x^i - \bar{x}$$

and

$$e_v^i = v^i - \bar{v}$$

- Notice

$$\begin{aligned} \dot{e}_p^i &= e_v^i \\ \dot{e}_v^i &= \frac{1}{M_i} u^i - \frac{1}{N} \sum_{j=1}^{N} \frac{1}{M_j} u^j \end{aligned} \tag{96}$$

➙ Challenge: Specify the u^i so that we get good cohesion properties and successful social foraging.

- Assume that each agent can sense its position and velocity relative to $\bar{x}$ and $\bar{v}$, but with some bounded errors.

➙ Let $d_p^i(t) \in \Re^n$, $d_v^i(t) \in \Re^n$ be these errors for agent i, respectively.

- Assume that $d_p^i(t)$ and $d_v^i(t)$ are sufficiently smooth and are independent of the state of the system.
- Each agent senses the gradient of J_p, but with some sufficiently smooth error $d_f^i(t) \in \Re^n$ (sensor noise or noise on the profile).
- Will use term "noise" loosely
- Assume:

$$
\begin{aligned}
\|d_p^i\| &\leq D_p \\
\|d_v^i\| &\leq D_v \\
\|d_f^i\| &\leq D_f
\end{aligned}
$$

where $D_p > 0$, $D_v > 0$, and $D_f > 0$ are known constants.

- Thus, each agent can sense

$$\hat{e}_p^i = e_p^i - d_p^i$$

$$\hat{e}_v^i = e_v^i - d_v^i$$

and

$$\nabla J_p\left(x^i\right) - d_f^i$$

at the location x^i where the agent is located.

➙ Suppose that in order to steer itself each agent uses

$$\begin{aligned} u^i &= -M_i k_a \hat{e}_p^i - M_i k_a \hat{e}_v^i - M_i k_v v^i \\ &+ M_i k_r \sum_{j=1, j\neq i}^{N} \exp\left(\frac{-\frac{1}{2}\|\hat{e}_p^i - \hat{e}_p^j\|^2}{r_s^2}\right)\left(\hat{e}_p^i - \hat{e}_p^j\right) \\ &- M_i k_f \left(\nabla J_p\left(x^i\right) - d_f^i\right) \end{aligned} \tag{97}$$

- Each agent knows its own mass M_i and velocity v^i.
- $k_v > 0$ is the gain for a "velocity damping term."
- $k_a > 0$ is the "attraction gain"

- k_r is a "repulsion gain" which sets how much the agents want to be away from each other.

- Notice that

$$\hat{e}_p^i - \hat{e}_p^j = \left(\left(x^i - \bar{x}\right) - d_p^i\right) - \left(\left(x^j - \bar{x}\right) - d_p^j\right) = \left(x^i - x^j\right) - \left(d_p^i - d_p^j\right)$$

- If $D_p = D_v = 0$, there is no sensing error on attraction and repulsion, thus, $\hat{e}_p^i = e_p^i$, $\hat{e}_v^i = e_v^i$, and $e_p^i - e_p^j = x^i - x^j$.

Social Foraging in Noise: Groups Can Increase Foraging Effectiveness

➙ Substitute this choice for u^i into the error dynamics in Equation (96) and study their stability properties.

➙ First, however, we will study how the *group* can follow the resource profile in the presence of noise.

- Consider $\dot{e}_v^i = \dot{v}^i - \dot{\bar{v}}$.

- First, note that

$$\begin{aligned}
\dot{\bar{v}} &= \frac{1}{N}\sum_{i=1}^{N}\frac{1}{M_i}u^i \\
&= -\frac{k_a}{N}\sum_{i=1}^{N}\left(\hat{e}_p^i + \hat{e}_v^i\right) - \frac{k_v}{N}\sum_{i=1}^{N}v^i \\
&+ \frac{k_r}{N}\sum_{i=1}^{N}\sum_{j=1, j\neq i}^{N}\exp\left(\frac{-\frac{1}{2}\|\hat{e}_p^i - \hat{e}_p^j\|^2}{r_s^2}\right)\left(\hat{e}_p^i - \hat{e}_p^j\right) \\
&- \frac{k_f}{N}\sum_{i=1}^{N}\left(R - d_f^i\right) \qquad (98)
\end{aligned}$$

$$\frac{1}{N}\sum_{i=1}^{N}\hat{e}_p^i = \frac{1}{N}\sum_{i=1}^{N}\left(\left(x^i - \bar{x}\right) - d_p^i\right) = \bar{x} - \frac{1}{N}N\bar{x} - \frac{1}{N}\sum_{i=1}^{N}d_p^i = -\frac{1}{N}\sum_{i=1}^{N}d_p^i$$

- Also, the term due to repulsion in Equation (98) is zero as we show next.

- Note that

$$\sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} \exp\left(\frac{-\frac{1}{2}\|\hat{e}_p^i - \hat{e}_p^j\|^2}{r_s^2}\right)(\hat{e}_p^i - \hat{e}_p^j) =$$

$$\left[\sum_{i=1}^{N} \hat{e}_p^i \sum_{j=1, j \neq i}^{N} \exp\left(\frac{-\frac{1}{2}\|\hat{e}_p^i - \hat{e}_p^j\|^2}{r_s^2}\right)]\right]$$

$$- \quad \left[\sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} \exp\left(\frac{-\frac{1}{2}\|\hat{e}_p^i - \hat{e}_p^j\|^2}{r_s^2}\right) \hat{e}_p^j\right] \qquad (99)$$

- The last term in Equation (99)

$$\sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} \exp\left(\frac{-\frac{1}{2}\|\hat{e}_p^i - \hat{e}_p^j\|^2}{r_s^2}\right) \hat{e}_p^j =$$

$$\sum_{j=1}^{N} \sum_{i=1, i \neq j}^{N} \exp\left(\frac{-\frac{1}{2} \| \hat{e}_p^i - \hat{e}_p^j \|^2}{r_s^2} \right) \hat{e}_p^j$$

and since

$$\exp\left(\frac{-\frac{1}{2} \| \hat{e}_p^i - \hat{e}_p^j \|^2}{r_s^2} \right) = \exp\left(\frac{-\frac{1}{2} \| \hat{e}_p^j - \hat{e}_p^i \|^2}{r_s^2} \right)$$

we have

$$\sum_{j=1}^{N} \sum_{i=1, i \neq j}^{N} \exp\left(\frac{-\frac{1}{2} \| \hat{e}_p^i - \hat{e}_p^j \|^2}{r_s^2} \right) \hat{e}_p^j =$$

$$\sum_{j=1}^{N} \hat{e}_p^j \sum_{i=1, i \neq j}^{N} \exp\left(\frac{-\frac{1}{2} \| \hat{e}_p^j - \hat{e}_p^i \|^2}{r_s^2} \right)$$

but this last value is the same as the first term in Equation (99).

- So overall its value is zero.

- This gives us

$$\dot{\bar{v}} = \frac{k_a}{N}\sum_{i=1}^{N} d_p^i + \frac{k_a}{N}\sum_{i=1}^{N} d_v^i + \frac{k_f}{N}\sum_{i=1}^{N} d_f^i - k_v\bar{v} - k_f R$$

- Letting $\bar{d}_p(t) = \frac{1}{N}\sum_{i=1}^{N} d_p^i(t)$ and similarly for $\bar{d}_v(t)$ and $\bar{d}_f(t)$ we get

$$\dot{\bar{v}} = -k_v\bar{v} + \underbrace{k_a\bar{d}_p + k_a\bar{d}_v + k_f\bar{d}_f - k_f R}_{z(t)} \qquad (100)$$

➙ This is an exponentially stable system with a time-varying but bounded input $z(t)$ so we know that $\bar{v}(t)$ is bounded.

- To see this, choose a Lyapunov function

$$V_{\bar{v}} = \frac{1}{2}\bar{v}^\top\bar{v}$$

defined on $D = \{\bar{v} \in \Re^n \mid \|\bar{v}\| < r_v\}$ for some $r_v > 0$, and we have

$$\dot{V}_{\bar{v}} = \bar{v}^\top \dot{\bar{v}} = -k_v \bar{v}^\top \bar{v} + z(t)^\top \bar{v}$$

with

$$\left\| \frac{\partial V_{\bar{v}}}{\partial \bar{v}} \right\| = \|\bar{v}\|$$

- Note that $\|z(t)\| \leq \|k_a \bar{d}_p\| + \|k_a \bar{d}_v\| + \|k_f \bar{d}_f\| + \|k_f R\| \leq \delta$, where $\delta = k_a D_p + k_a D_v + k_f D_f + k_f \|R\|$. If $\delta < k_v \theta r_v$ for all $t \geq 0$ for some positive constant $\theta < 1$, and all $\bar{v} \in D$, then it can be proven that for all $\|\bar{v}(0)\| < r_v$, and some finite T we have

$$\|\bar{v}(t)\| \leq \exp\left[-(1-\theta)k_v t\right] \|\bar{v}(0)\|, \ \forall\ 0 \leq t < T$$

and

$$\|\bar{v}(t)\| \leq \frac{\delta}{k_v \theta}, \ \forall\ t \geq T$$

- Since this holds globally we can take $r_v \to \infty$ so these inequalities hold for all $\bar{v}(0)$.

- If δ and θ are fixed, with increasing k_v we get that $\|\bar{v}(t)\|$ decreases faster for $0 \leq t < T$ and smaller bound on $\|\bar{v}(t)\|$ for $t \geq T$.

➙ If δ gets larger with k_v and θ fixed, $\|\bar{v}(t)\|$ has larger bound for $t \geq T$; hence if the magnitude of the noise increases this increases δ and hence there can be larger magnitude changes in the ultimate average velocity of the swarm (e.g., the average velocity could oscillate).

➙ Note that if in Equation (100) $z(t) \approx 0$ (e.g., due to noise that destroys the directionality of the resource profile R), then the above bound may be reduced but the swarm could be going in the wrong direction.

➙ Average sensing errors of the group is what changes the direction of the group's movement relative to the direction of the gradient of $J_p(x)$.

- For big swmars it can be that $\bar{d}_p \approx \bar{d}_v \approx \bar{d}_f \approx 0$ to give a zero average sensing error and the group will perfectly follow the proper direction for foraging
- Also, groups can often climb noisy gradients better than an individual - Grunbaum.

Cohesive Social Foraging in Noise

- Consider the $\dot{v}^i$ term of $\dot{e}_v^i = \dot{v}^i - \dot{\bar{v}}$ in Equation (96).
  - Note that

$$
\begin{aligned}
\dot{v}^i = \frac{1}{M_i} u^i &= -k_a \hat{e}_p^i - k_a \hat{e}_v^i - k_v v^i \\
&+ k_r \sum_{j=1, j \neq i}^{N} \exp\left(\frac{-\frac{1}{2} \|\hat{e}_p^i - \hat{e}_p^j\|^2}{r_s^2} \right) \left(\hat{e}_p^i - \hat{e}_p^j\right) \\
&- k_f \left(\nabla J_p \left(x^i\right) - d_f^i \right) \\
&= -k_a e_p^i + k_a d_p^i - k_a e_v^i + k_a d_v^i - k_v v^i
\end{aligned}
$$

$$
\begin{aligned}
+ \quad & k_r \sum_{j=1, j\neq i}^{N} \exp\left(\frac{-\frac{1}{2}\|\hat{e}_p^i - \hat{e}_p^j\|^2}{r_s^2} \right) \left(\hat{e}_p^i - \hat{e}_p^j\right) \\
- \quad & k_f \left(R - d_f^i\right)
\end{aligned}
$$

- Hence, we have

$$
\begin{aligned}
& \dot{e}_v^i = \dot{v}^i - \dot{\bar{v}} = \\
& -k_a e_p^i - k_a e_v^i - k_v e_v^i + k_a \left(d_p^i - \bar{d}_p\right) + k_a \left(d_v^i - \bar{d}_v\right) \\
& +k_r \sum_{j=1, j\neq i}^{N} \exp\left(\frac{-\frac{1}{2}\| \left(x^i - x^j\right) - \left(d_p^i - d_p^j\right) \|^2}{r_s^2} \right) \left(\left(x^i - x^j\right)\right. \\
& \left. - \left(d_p^i - d_p^j\right)\right) + k_f \left(d_f^i - \bar{d}_f\right)
\end{aligned}
$$

➙ To study the stability of the error dynamics, and hence swarm cohesiveness, define

$$
E^i = [{e_p^i}^\top, \ {e_v^i}^\top]^\top
$$

and $E = \left[{E^1}^\top, {E^2}^\top, \ldots, {E^N}^\top\right]^\top$, and choose a Lyapunov function

$$V(E) = \sum_{i=1}^{N} V_i\left(E^i\right)$$

where

$$V_i\left(E^i\right) = {E^i}^\top P E^i$$

with $P = P^\top$ and $P > 0$ (a positive definite matrix).

- We know that

$$\lambda_{min}(P){E^i}^\top E^i \leq {E^i}^\top P E^i \leq \lambda_{max}(P){E^i}^\top E^i$$

- Notice that with I an $n \times n$ identity matrix, we have

$$\dot{E}^i = \underbrace{\begin{bmatrix} 0 & I \\ -k_a I & -\left(k_a + k_v\right) I \end{bmatrix}}_{A} E^i + \underbrace{\begin{bmatrix} 0 \\ I \end{bmatrix}}_{B} g^i(E)$$

where

$$
\begin{aligned}
g^i(E) &= k_a\left(d_p^i - \bar{d}_p\right) + k_a\left(d_v^i - \bar{d}_v\right) \\
&+ k_r \sum_{j=1, j\neq i}^{N} \exp\left(\frac{-\frac{1}{2}\|\hat{e}_p^i - \hat{e}_p^j\|^2}{r_s^2}\right)\left(\hat{e}_p^i - \hat{e}_p^j\right) \\
&+ k_f\left(d_f^i - \bar{d}_f\right) \qquad (101)
\end{aligned}
$$

- Note that any matrix

$$
\begin{bmatrix} 0 & I \\ -k_1 I & -k_2 I \end{bmatrix}
$$

with $k_1 > 0$ and $k_2 > 0$ has eigenvalues given by the roots of $\left(s^2 + k_2 s + k_1\right)^n$, which are in the strict left half plane.

- Since $k_a > 0$ and $k_v > 0$, the matrix A above is Hurwitz (i.e., has eigenvalues all in the strict left half plane).

- We have

$$\dot{V}_i = {E^i}^\top P \dot{E}^i + {\dot{E}^i}^\top P E^i = {E^i}^\top \underbrace{\left(PA + A^\top P\right)}_{-Q} E^i + 2{E^i}^\top P B g^i(E) \tag{102}$$

- Note that if Q, defined in this manner, is such that $Q = Q^\top$ and $Q > 0$, then the unique solution P of $PA + A^\top P = -Q$ has $P = P^\top$ and $P > 0$ as needed.

- Also, since $\|B\| = 1$, ${E^i}^\top Q E^i \geq \lambda_{min}(Q) {E^i}^\top E^i$, and $\|P\| = \lambda_{max}(P)$ with $P = P^\top > 0$, we have

$$\begin{aligned} \dot{V}_i &\leq -\lambda_{min}(Q) \left\|E^i\right\|^2 + 2\left\|E^i\right\| \lambda_{max}(P) \|g^i(E)\| \\ &= -\lambda_{min}(Q) \left(\left\|E^i\right\| - \frac{2\lambda_{max}(P)}{\lambda_{min}(Q)} \|g^i(E)\| \right) \left\|E^i\right\| \end{aligned} \tag{103}$$

- Suppose for a moment that for each $i = 1, 2, \ldots, N$, $\|g^i(E)\| < \beta$ for some known β.

- Then, if

$$\|E^i\| > \frac{2\lambda_{max}(P)}{\lambda_{min}(Q)}\|g^i(E)\| \tag{104}$$

we have that $\dot{V}_i < 0$.

- Hence, the set

$$\Omega_b = \left\{E : \ \|E^i\| \leq 2\frac{\lambda_{max}(P)}{\lambda_{min}(Q)}\|g^i(E)\|, \ i = 1, 2, \ldots, N\right\} \tag{105}$$

is attractive and compact.

➙ Also we know that within a finite amount of time, $E^i \to \Omega_b$.

➙ This means that we can guarantee that if the swarm is not cohesive, it will seek to be cohesive, but this can only be guaranteed if it is a certain distance from cohesiveness as indicated by Equation (104).

➙ It remains to show that for each i, $\|g^i(E)\| < \beta$ for some β.

- Note that

$$
\begin{aligned}
\|g^i(E)\| &\le k_a\|d_p^i - \bar{d}_p\| + k_a\|d_v^i - \bar{d}_v\| + k_f\|d_f^i - \bar{d}_f\| \\
&+ k_r \sum_{j=1, j\neq i}^{N} \exp\left(\frac{-\frac{1}{2}\|\psi\|^2}{r_s^2}\right)\|\psi\|
\end{aligned} \tag{106}
$$

where $\psi = \hat{e}_p^i - \hat{e}_p^j = (x^i - x^j) - (d_p^i - d_p^j)$.

- Notice that $\frac{1}{N}\sum_{j=1}^{N}\|d_p^j\| \le D_p$ since $\|d_p^j\| \le D_p$.
- Also

$$
d_p^i - \frac{1}{N}\sum_{j=1}^{N} d_p^j \le \|d_p^i\| + \frac{1}{N}\|\sum_{j=1}^{N} d_p^j\| \le \|d_p^i\| + \frac{1}{N}\sum_{j=1}^{N}\|d_p^j\|
$$

$\|d_p^i - \bar{d}_p\| \le 2D_p$, $\|d_v^i - \bar{d}_v\| \le 2D_v$, and $\|d_f^i - \bar{d}_f\| \le 2D_f$.

➙ For the last term in Equation (106), note that as $\|x^i - x^j\|$ becomes large for all i and j, the agents are all far from each other and the repulsion term goes to zero.

➙ Also, the term due to the repulsion is bounded with a unique maximum point.

- To find this point note that

$$\frac{\partial}{\partial\|\psi\|}\left(\|\psi\|\exp\left(\frac{-\frac{1}{2}\|\psi\|^2}{r_s^2}\right)\right) =$$

$$\exp\left(\frac{-\frac{1}{2}\|\psi\|^2}{r_s^2}\right) - \frac{\|\psi\|^2}{r_s^2}\exp\left(\frac{-\frac{1}{2}\|\psi\|^2}{r_s^2}\right)$$

- The maximum point occurs at a point such that

$$1 - \frac{\|\psi\|^2}{r_s^2} = 0$$

or when $\|\psi\| = r_s$.

- Hence, we have

$$
\begin{aligned}
\|g^i(E)\| &\leq 2k_a\left(D_p + D_v\right) + 2k_f D_f + k_r \sum_{j=1, j\neq i}^{N} \exp\left(-\frac{1}{2}\right) r_s \\
&= 2k_a\left(D_p + D_v\right) + 2k_f D_f + k_r r_s (N-1) \exp\left(-\frac{1}{2}\right) = \beta
\end{aligned}
$$

- If you substitute this value for β into Equation (105) you get the set Ω_b that ultimately all the trajectories will end up in.

Cohesive Social Foraging with No Noise: Optimization Perspective

- When there is no noise, tighter bounds and stronger results can be obtained.
- First, we can eliminate the effect of P via $\lambda_{max}(P)$ on the bound for the no-noise case. As

- sume there is no sensor noise so $D_p = D_v = D_f = 0$. Choose

$$
\begin{aligned}
u^i &= -M_i k_a e_p^i - M_i k_a e_v^i - M_i k_v v^i \\
&+ M_i k_r \left(B^\top P^{-1} B\right) \sum_{j=1, j\neq i}^{N} \exp\left(\frac{-\frac{1}{2}\|e_p^i - e_p^j\|^2}{r_s^2}\right)\left(e_p^i - e_p^j\right) \\
&- M_i k_f R
\end{aligned} \tag{107}
$$

where $P = P^\top$, $P > 0$ was defined earlier, so P^{-1} exists.

- Also

$$
\begin{aligned}
\dot{V}_i &\leq -\lambda_{min}(Q)\left\|E^i\right\|^2 \\
&+ 2{E^i}^\top PB\left(k_r B^\top P^{-1} B \sum_{j=1, j\neq i}^{N} \exp\left(\frac{-\frac{1}{2}\|e_p^i - e_p^j\|^2}{r_s^2}\right)\left(e_p^i - e_p^j\right)\right) \\
&= -\lambda_{min}(Q)\left\|E^i\right\|^2
\end{aligned}
$$

$$
\begin{aligned}
+ \quad & 2k_r {E^i}^\top B \sum_{j=1, j\neq i}^{N} \exp\left(\frac{-\frac{1}{2}\|e_p^i - e_p^j\|^2}{r_s^2} \right) (e_p^i - e_p^j) \\
\leq \quad & -\lambda_{min}(Q) \left\| E^i \right\|^2 + 2k_r \left\| E^i \right\| (N-1) \exp\left(-\frac{1}{2}\right) r_s
\end{aligned}
$$

- So $\dot{V}_i < 0$ if $\left\| E^i \right\| > \frac{2k_r(N-1)r_s}{\lambda_{min}(Q)} \exp\left(-\frac{1}{2}\right)$. Let

$$
\Omega_b' = \left\{ E : \ \left\| E^i \right\| \leq \frac{2k_r r_s (N-1)}{\lambda_{min}(Q)} \exp\left(-\frac{1}{2}\right), \ i = 1, 2, \ldots, N \right\}
$$

➙ Next, note that in the set Ω_b, we have bounded e_p^i and e_v^i but we are not guaranteed that $e_v^i \to 0$ for any i.

➙ Achieving $e_v^i \to 0$ for all i would be a desirable property since this represents that $v^i = \bar{v}$ for all i so that the group will all move cohesively in the same direction.

➙ Consider Ω_b', and consider a Lyapunov function

$V^o(E) = \sum_{i=1}^N V_i^o\left(E^i\right)$ with

$$V_i^o\left(E^i\right) = \frac{1}{2}k_a {e_p^i}^\top e_p^i + \frac{1}{2}k_a {e_v^i}^\top e_v^i + k_r r_s^2 \sum_{j=1, j\neq i}^{N} \exp\left(\frac{-\frac{1}{2}\|e_p^i - e_p^j\|^2}{r_s^2}\right)$$

- Note that this Lyapunov function satisfies $V_i^o\left(E^i\right) \geq 0$.

➙ You should view the objective of the agents as being that of *minimizing* this Lyapunov function; they try to minimize the distance to the center of the swarm, match the average velocity of the group, and minimize the repulsion effect (to do that the agents move away from each other).

- We have

$$\begin{aligned} \nabla_{e_p^i} V_i^o &= k_a e_p^i - k_r \sum_{j=1, j\neq i}^{N} \exp\left(\frac{-\frac{1}{2}\|e_p^i - e_p^j\|^2}{r_s^2}\right)\left(e_p^i - e_p^j\right) \\ \nabla_{e_v^i} V_i^o &= e_v^i \end{aligned}$$

so

$$
\begin{aligned}
\dot{V}_i^o &= \left(\nabla J\left(E^i\right)\right)^\top \dot{E}_i \\
&= k_a e_p^{i\top} e_v^i - k_r \sum_{j=1, j\neq i}^{N} \exp\left(\frac{-\frac{1}{2}\|e_p^i - e_p^j\|^2}{r_s^2}\right)\left(e_p^i - e_p^j\right)^\top e_v^i \\
&\quad + e_v^{i\top}\left(-k_a e_p^i - k_a e_v^i - k_v e_v^i + k_r \sum_{j=1, j\neq i}^{N} \exp\left(\frac{-\frac{1}{2}\|e_p^i - e_p^j\|^2}{r_s^2}\right)\left(e_p^i - e_p^j\right)\right) \\
&= -\left(k_a + k_v\right) e_v^{i\top} e_v^i
\end{aligned}
$$

- Hence,

$$\dot{V}^o = -\left(k_a + k_v\right)\sum_{i=1}^{N} \|e_v^i\|^2 \leq 0$$

on $E \in \Omega$ for a compact set Ω.

- Choose Ω so it is positively invariant, which is clearly possible, and so $\Omega_e \in \Omega$ where

$$\Omega_e = \{E : \ \dot{V}^o(E) = 0\} = \{E : \ e_v^i = 0, \ i = 1, 2, \ldots, N\}$$

➙ From LaSalle's Invariance Principle we know that if $E(0) \in \Omega$ then $E(t)$ will converge to the largest invariant subset of Ω_e.

- Hence
$$e_v^i(t) \rightarrow 0$$
as $t \rightarrow \infty$.

- When $R = 0$ (no resource profile effect), $\bar{v}(t) \rightarrow 0$ and hence $v^i(t) \rightarrow 0$ as $t \rightarrow \infty$ for all i (i.e., ultimately no oscillations in the average velocity).

- If $R \neq 0$, then $\dot{\bar{v}} = -k_v \bar{v} - k_f R$ and $\bar{v}(t) \rightarrow -\frac{k_f}{k_v} R$ as $t \rightarrow \infty$, and thus, $v^i(t) \rightarrow -\frac{k_f}{k_v} R$ for all i as $t \rightarrow \infty$, i.e, the group follows the profile.

➙ These results help to highlight the effects of the noise.

➙ The noise makes it so that the swarm may not follow the profile as well (but makes following it possible when it may not be possible for a single individual), and it destroys tight cohesion characterized by getting $e_v^i(t) \to 0$.

Cohesion Characteristics and Swarm Dynamics

- Suppose that u^i is given by Equation (97).

Effects of Parameters on Swarm Size

- The size of Ω_b in Equation (105), which we denote by $|\Omega_b|$, is directly a function of several known parameters.

- Consider the following cases:
  - *No sensing errors:* If there are no sensing errors, i.e., $D_p = D_v = D_f = 0$, and if $Q = k_a I$ we obtain

$$\Omega_b = \left\{ E : \left\| E^i \right\| \leq \frac{2k_r r_s(N-1)}{k_a} \lambda_{max}(P) \exp\left(-\frac{1}{2}\right), \; i = 1, 2, \ldots, N \right\}$$

    If N, k_r, and r_s are fixed, then if k_a increases from zero we get $\frac{\lambda_{max}(P)}{k_a} \to 1$ from above and we get a decrease in $|\Omega_b|$, but only up to a certain point.

– Sensing errors: There are several characteristics of interest:

  * Noise cancellations: In the special situation when $d_p^i = d_p^j$, $d_v^i = d_v^j$, and $d_f^i = d_f^j$ for all i and j, then $d_p^i - \bar{d}_p = d_v^i - \bar{d}_v = d_f^i - \bar{d}_f$ for all i and it is as if there were no error and $|\Omega_b|$ is smaller.
  * Repel effects: For fixed values of N, k_a, and k_r if we increase r_s each agent has a larger region from which it will repel its neighbors so $|\Omega_b|$ is larger. For fixed k_r, k_a, and r_s if we let $N \to \infty$, then $|\Omega_b| \to \infty$ as we expect due to the repulsion (the bound is conservative since it depends on the special case of all agents being aligned on a line so there are $N-1$ inter-agent distances that sum to make the bound large).
  * Attraction can amplify noise: Let $D_s = D_p + D_v$ and J quantify the size of the set Ω_b. Next, we study the special case of choosing $Q = k_a I$. Fix all values of the parameters

except k_a and D_s. A plot of J versus k_a and D_s is shown in Figure 217, where the locus of points are those values of k_a that minimize J for each given value of D_s.

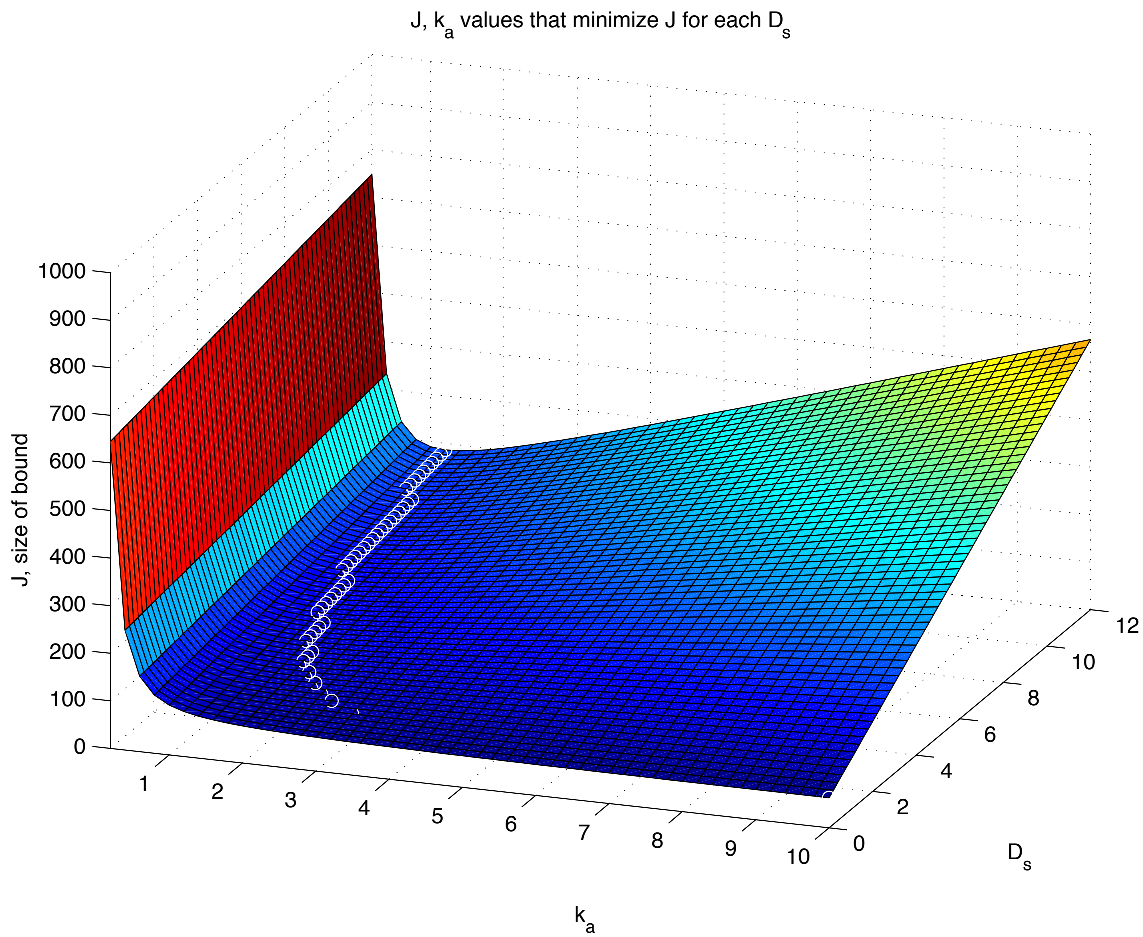

Figure 217: Values of k_a that minimize J for given values of noise magnitude D_s.

This plot shows that if there is a set magnitude of the noise, then to get the best cohesiveness (smallest Ω_b) k_a should not be too small (or it would not hold the group together), but also not too large since then the noise is also amplified by the attraction gain and poor cohesion results. Could you interpret the plot as a type of fitness function, and then make any conclusions about the evolution of the agent parameters?

– Swarm size N: In some situations, when N is very large, $\bar{d}_p \approx \bar{d}_v \approx \bar{d}_f \approx 0$ and there is no biasing of sensing errors so that on average they are zero and this reduces the above bound on $\|g^i(E)\|$.

Swarm Dynamics: Individual and Group

➙ Assume no noise. We use linear attraction, velocity damping, the Gaussian form for the repulsion term, and the resource profile with the shape of a plane.

➙ $N = 50$, $k_a = 1$, $k_r = 10$, $r_s^2 = 0.1$, $k_v = k_f = 0.1$, $R = [1, 2, 3]^\top$, and $r_o = 0$.

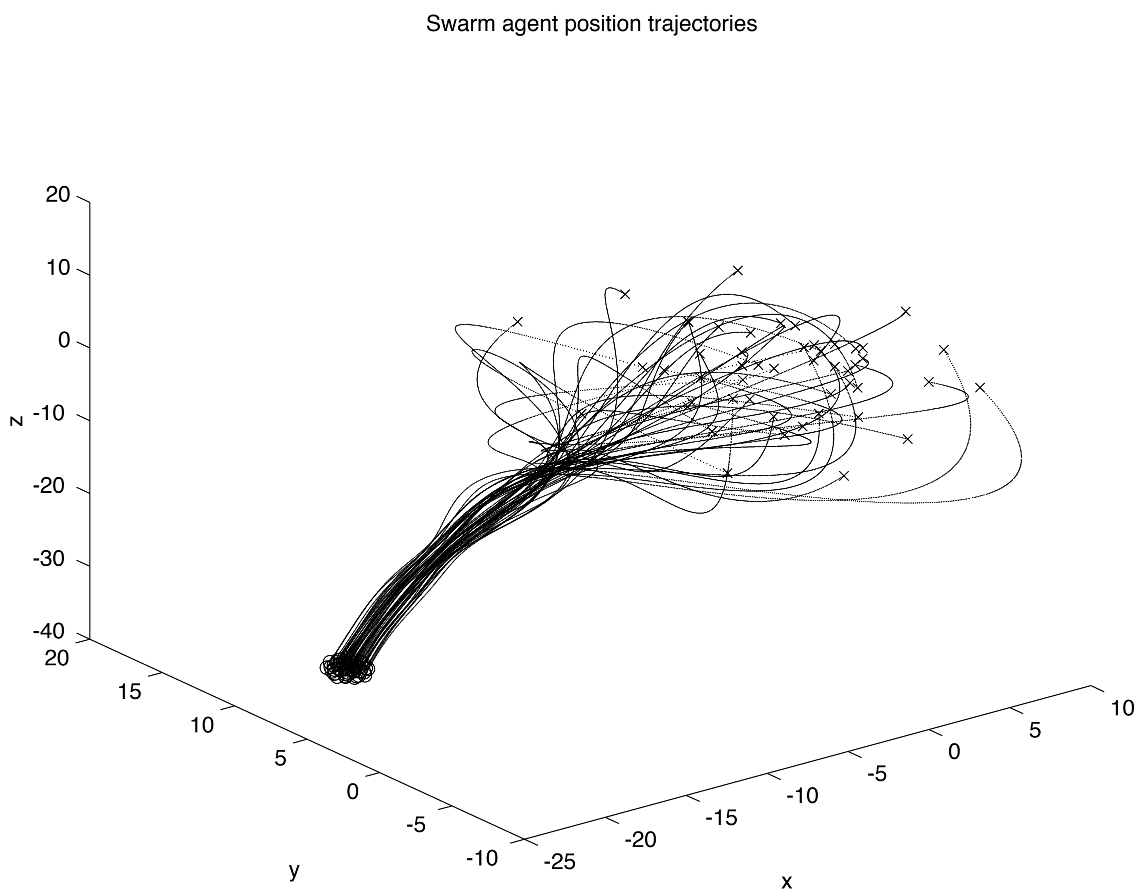

Figure 218: Agent trajectories in a swarm.

- The group achieves a certain level of aggregation relatively quickly, then the *group* moves to follow the foraging profile.

Design Example: Cooperative Robot Swarms

Robot Swarm Formulation

- ➙ Robot swarm = group of robots that move in a cohesive fashion in order to perform some task.
- ➙ Here, the task is simply to get the group to a certain location in a factory so that they can perform some activity together there.
  - Obstacle avoidance by group is a problem
  - Use same problem as in planning for obstacle avoidance.
    - ➙ "Obstacle function," "goal function" combined to get cost function
    - – Think of as a nutrient profile or resource profile (avoidance, goal-seeking)
    - – Use swarm agent model above with $M_i = 1$.
    - – Each robot perfectly knows the swarm center and swarm

average velocity and its own velocity (also consider noise case)

Performance in Obstacle Avoidance and Noise Effects

- ➙ $N = 20$, $k_p = 1$, $k_v = 0.1$, $k_f = 0.1$, $k_r = 10$, and $r_s^2 = 1$. We pick some random initial locations and velocities, but within some fixed ranges.
  - Pick $w_1 = 120$ and $w_2 = 0.1$ after some tuning.
- ➙ Use an Euler approximation and simulate the swarm as a discrete-time system.
- ➙ Do stability results of last section hold? No.

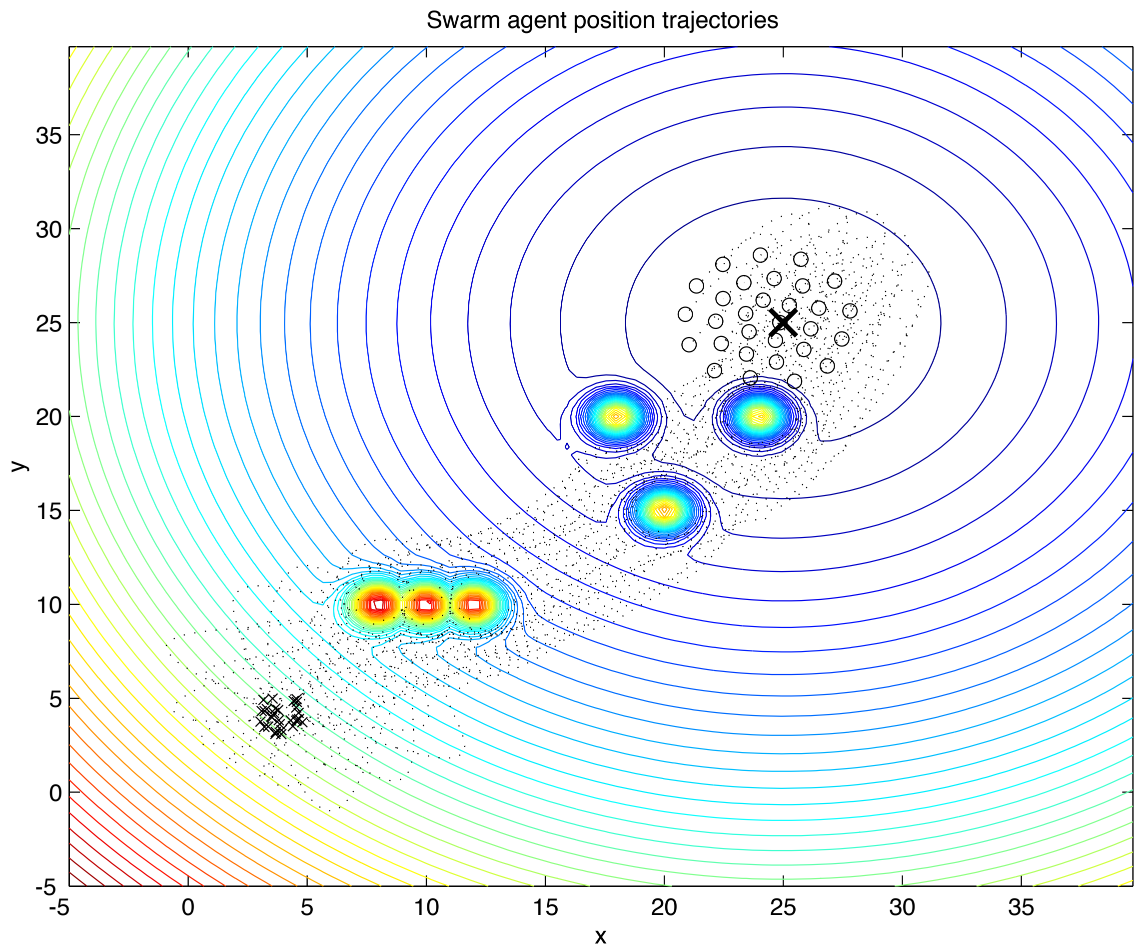

Figure 219: Robot swarm, robot position trajectories showing how they avoid collisions with obstacles and reach the goal position (no noise).

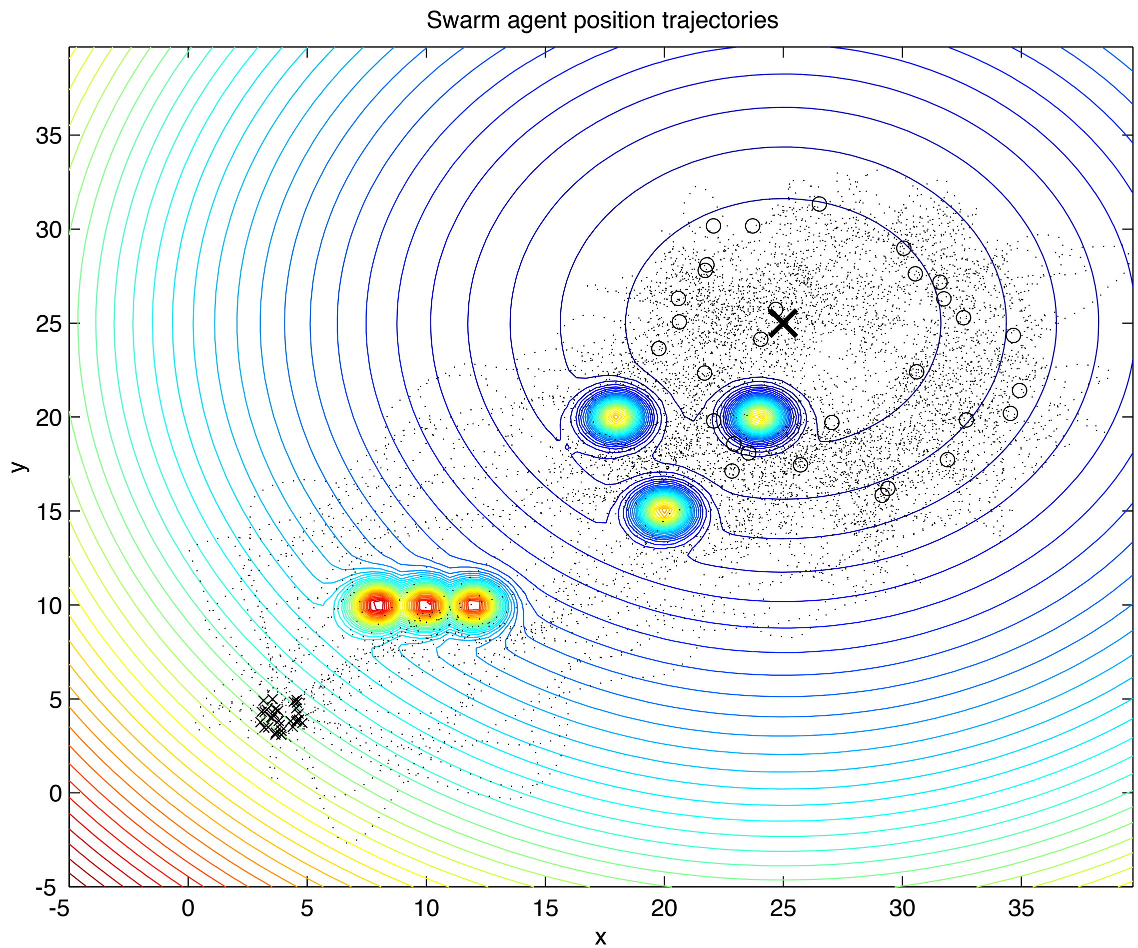

Figure 220: Robot swarm, robot position trajectories, noise case.

Additional Robot Swarm Design Challenges

- Complex mazes, dead ends, circular loops, mobile obstacles, uncertainty
- Does the group help or hurt ability to overcome these issues (e.g., adding a network can help, but could add even more uncertainty)?

Competitive and Intelligent Foraging

- Foraging invovles both search and competition for food (limited resources, fighting, predator/prey)
- "Foraging games" - games where animals compete for resources or cooperate to obtain resources.
- Aspects:
  - Resources limited
  - Environment places constraints (including aspects of predators)
  - Physiology of animal places constraints
- Cooperative foraging is social foraging
- Intelligent foraging: planning, learning, attention
- Evolution designs both cooperative and competitive foraging strategies

Introduction to Game Theory

Strategies and Information for Decisions

Players, Rules, and Payoffs

- Two "players" ("decision-makers"), P_1 and P_2
- θ^1, θ^2 are "decision variables" ofP_1 and P_2
- J_1 and J_2 are cost functions of P_1 and P_2 (i.e., gain or loss for given decisions).
- P_1 (P_2) wants to minimize (maximize) J_1.
- If

$$J_1(\theta^1, \theta^2) + J_2(\theta^1, \theta^2) = 0$$

  for all θ^1 and θ^2, we have a "zero sum game"
- Initially suppose that ("actions")

$$\theta^1 \in \{1, 2, \ldots, D_1\}$$

and

$$\theta^2 \in \{1, 2, \ldots, D_2\}$$

- For simplicity, $\theta^1 = i$ and $\theta^2 = j$ for P_1 and P_2
- Denote costs by $J_1(i, j)$ and $J_2(i, j)$.
- Two player case: $J_1(i, j)$ and $J_2(i, j)$ are $D_1 \times D_2$ matrices, J_1^{ij} and J_2^{ij}
- Suppose that the game is only played once (i.e., P_1 and P_2 only make decisions once).
- The players make decisions with full information about payoffs, they make their decisions simultaneously, and are "rational."
- A pair of decision strategies for a two-player finite game is denoted by (i, j).
- The "outcome" of the game is J_1^{ij}.
- An "optimal" strategy is (i^*, j^*).

Security Strategies, Saddle Point Strategies, and Decision Information

- Let $D_1 = 5$ and $D_2 = 3$.
- Suppose

$$J_1^{ij} = \begin{bmatrix} -3 & 4 & 4 \\ 0 & -5 & 2 \\ -2 & 1 & -4 \\ 2 & 3 & -4 \\ 2 & -2 & -5 \end{bmatrix} \tag{108}$$

- Rows correspond to P_1 decisions and columns to P_2 decisions.

➙ In a "security strategy" a player makes decisions to secure losses against what ever the other player might do (i.e., it minimizes its maximum possible loss).

- P_1 picks row i^* such that any value in any column of this row is no bigger than the largest value of any other row $i \neq i^*$ (i.e., pick the row that minimizes the maximum size column value).

- For Equation (108) the list of maximum values for each row is

$$\begin{array}{c} 4 \\ 2 \\ 1 \\ 3 \\ 2 \end{array}$$

so the security strategy is for P_1 to pick $i^* = 3$.

- The "loss ceiling" (i.e., the most it can lose) is 1, which is less than the other possible losses, and this is called the "security level" of P_1.

- P_2 can adopt a security strategy by choosing the column j^*

whose row values are smaller than the smallest value found for another column $j \neq j^*$.

- In this case, the list of minimum values is

$$-3 \quad -5 \quad -5$$

so the security strategy for P_2 is $j^* = 1$ and P_2 secures gains at the "gain floor" (its security level) of -3 (i.e., he pays no more than 3).

- The security level of P_1 is never below the security level of P_2.
- The "outcome" of the game,

$$J_1^{i^* j^*} = J_1^{31} = -2$$

will lie between the two security levels.

- Special case: the security levels of the two players are the same, the security strategies of the two players are in "equilibrium" with each other since they are "optimal" with respect to each other; in this case they are called "saddle point strategies."

➙ A pair of strategies is said to be in a "saddle point equilibrium" if unilateral deviations by one player from its strategy will not benefit that player.

- In general, for a $D_1 \times D_2$ matrix game if (i^*, j^*) is the pair of chosen strategies, and if

$$J_1^{i^* j} \leq J_1^{i^* j^*} \leq J_1^{i j^*}$$

for all $i \in \{1, 2, \ldots, D_1\}$ and $j \in \{1, 2, \ldots, D_2\}$, then (i^*, j^*) constitutes a saddle point equilibrium.

- The value of $J_1^{i^* j^*}$ is called the "saddle point value."

➙ Order of play and knowledge changes outcomes

- If use sequence of information get a "dynamic" rather than "static" game

- ➙ Above are “pure” strategies (there are also “mixed” ones)
- ➙ All concepts extend to multiple players (“nature” could be one)
- ➙ There are “noncooperative games” (above, below) and “cooperative games” (below) and ones with elements of both.

Extensive Forms and Decision Trees

- The matrix form of a game is called its “normal form.”
- An “extensive form” of a game involves creating a type of labeled (multi-stage) “decision tree” to represent the game, like:

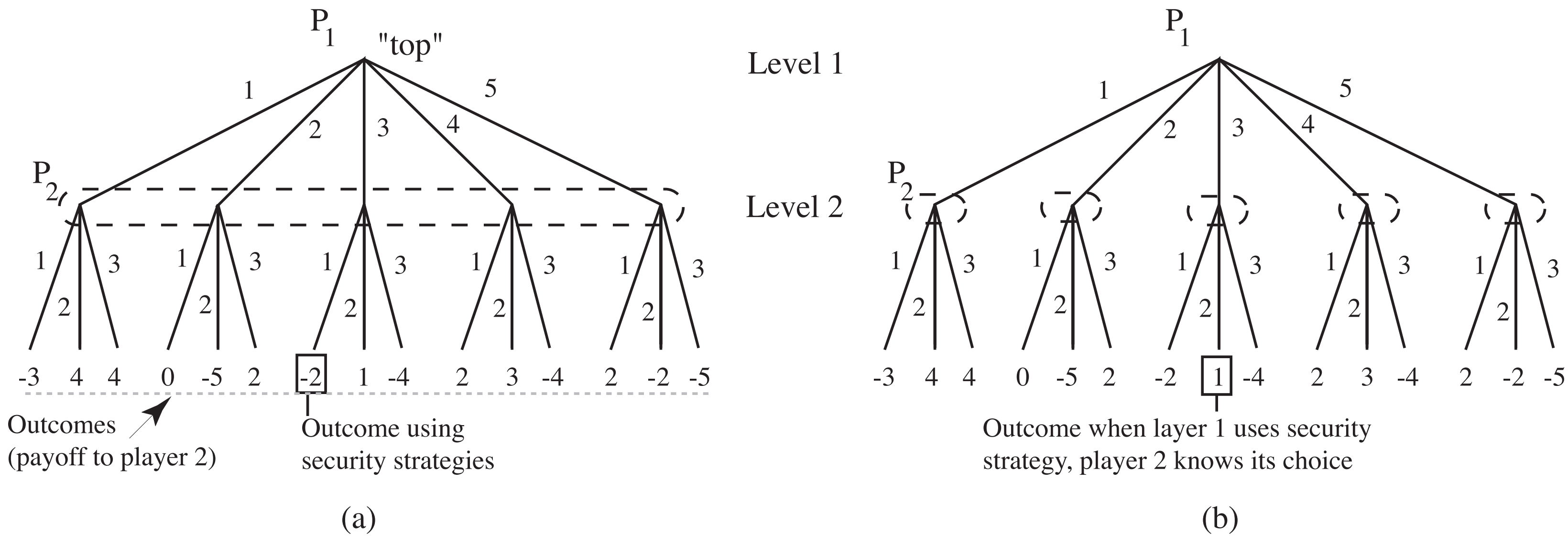

Figure 221: Example of an extensive representation, in this case in (a) for the matrix game in Equation (108) and in (b) for the same payoff matrix but when P_1 chooses first and then P_2 knows its decision before choosing.

➙ Decisions versus strategies: Decision is the *specific* value chosen while a strategy is a rule (policy, function) for picking actions based on some set of information

Competitive Games: Nash, Minimax, and Stackelberg Strategies

Nash Equilibrium Strategies

- Non zero sum case: that

$$J_1(i,j) + J_2(i,j) \neq 0$$

  - Payoff matrices J_1^{ij} and J_2^{ij} denote losses of P_1 and P_2
  - Assume players are rational
  - Unless otherwise stated we assume that there is no cooperation and decisions are made independently.

- The basic problem for each player is that the outcome resulting from their decision also depends on what the other player decides.

  - A pair of strategies is said to be in a "saddle point equilibrium" if unilateral deviations by one player from its strategy will not benefit that player.

- Many types of "equilibrium" strategies

➙ A strategy pair (i^*, j^*) is a noncooperative (Nash) equilibrium solution to a "bimatrix" game (J_1^{ij}, J_2^{ij}) if the inequalities

$$J_1^{i^*j^*} \leq J_1^{ij^*} \tag{109}$$

and

$$J_2^{i^*j^*} \leq J_2^{i^*j} \tag{110}$$

are both satisfied for all $i \in \{1, 2, \ldots, D_1\}$ and all $j \in \{1, 2, \ldots, D_2\}$.

- The pair $(J_1^{i^*j^*}, J_2^{i^*j^*})$ is the noncooperative (Nash) equilibrium outcome of the game.

➙ There can be no Nash solutions, one Nash solution, or many Nash solutions.

- If $J_1^{ij} = -J_2^{ij}$ for all i and j then we have a zero sum game, and the Nash solution is a saddle point equilibrium for the game.

➙ Example: $D_1 = 5$ and $D_2 = 3$.

- Suppose

$$J_1^{ij} = \begin{bmatrix} -1 & 5 & -3 \\ -2 & 5 & 1 \\ 4 & 3 & -2 \\ -5 & -1 & 5 \\ 3 & 0 & 2 \end{bmatrix}, \; J_2^{ij} = \begin{bmatrix} -1 & 2 & -3 \\ 2 & -3 & 1 \\ -2 & 3 & 1 \\ -1 & 1 & -1 \\ 4 & -4 & 1 \end{bmatrix} \qquad (111)$$

- To solve for the Nash equilibria, consider candidate (i^*, j^*) pairs in turn and test if they satisfy both the inequalities in Equations (109) and (110).

- To do this, consider $(1, 1)$ and see if J_1^{11} is less than all other row elements of column one to test Equation (109).

- Since it is not, $(1, 1)$ cannot be a Nash solution.
- $(1, 2)$ is not a Nash solution.
- If you test $(1, 3)$ you will see that $J_1^{13} \leq J_1^{i3}$ for all i so it is a candidate, so test the inequality in Equation (110) and you will find that $J_2^{13} \leq J_2^{1j}$ for all j; hence $(i^*, j^*) = (1, 3)$ is indeed a Nash equilibrium.
- The Nash equilibrium outcome is $(-3, -3)$ so that both players *gain* 3.
- Show that $(4, 1)$ is also a Nash solution, but that all others are not.

➙ A Nash equilibrium solution is special since if the players adopt it, then they have no reason after playing the game to regret their decisions.

- There can be more than one Nash equilibrium so the question of which one to use arises.
- It is not possible to totally order the Nash strategies according to the values of their outcomes because they are defined by *pairs* of numbers.
- One Nash strategy is "better" than another if *both* outcomes are better than the other.

➙ Problem: If one player picks one Nash strategy and another picks another Nash strategy then they could both do worse (how to resolve? use cooperation?)

Infinite Games and Reaction Curves

- Consider an infinite number of strategy (decision) choices by one or both of the players.

➙ Example: Actions of P_1 can be

$$\theta^1 \in [-4, 4]$$

and for P_2

$$\theta^2 \in [-5, 5]$$

- Suppose

$$J_1(\theta^1, \theta^2) = -\exp\left(-\frac{(\theta^1 - 2)^2}{8} - \frac{(\theta^2 - 4)^2}{2}\right)$$

and

$$J_2(\theta^1, \theta^2) = -\exp\left(-\frac{(\theta^1 - 1)^2}{1} - \frac{(\theta^2 + 1)^2}{6}\right)$$

each of which has one global minimum, and also for which if you fix θ^1 (θ^2) there is a unique minimum point in the other value.

➙ Define the "reaction curve" of P_1 to be

$$R_1(\theta^2) = \arg\min_{\theta_1} J_1(\theta^1, \theta^2)$$

where we are using the assumption of uniqueness of the minimum point so that there is only one point at which the minimum is achieved.

➙ The reaction curve $R_1(\theta^2)$ defines how P_1 should react for every possible action of P_2 in order to minimize its losses.

- Similarly, for P_2

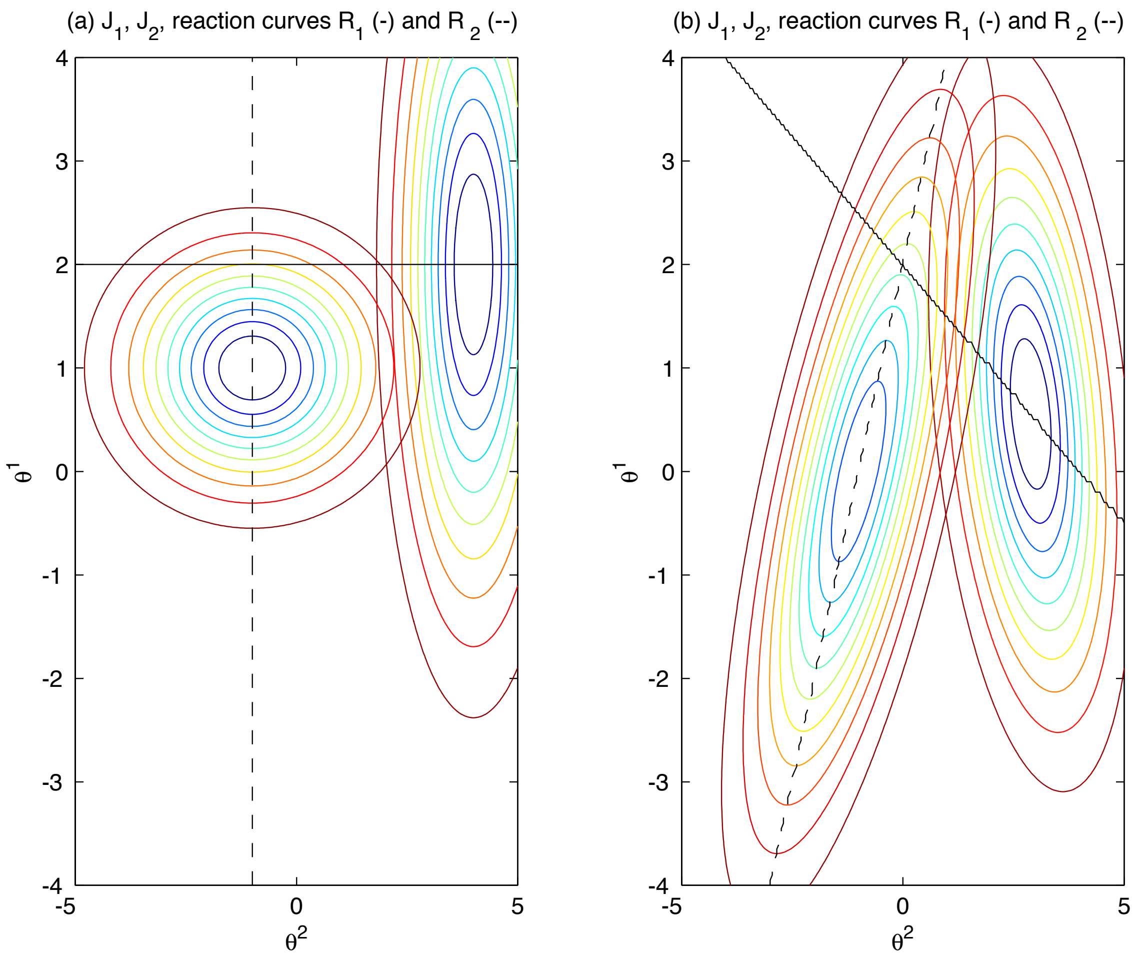

Figure 222: (a) Contour plots of J_1 and J_2 and reaction curves R_1 (solid) and R_2 (dashed); (b) same but for different cost functions J_1 and J_2.

- Any intersection point of the two curves in Figure 222(a) is a Nash equilibrium, so $(2, -1)$ is the unique Nash equilibrium.
- Curve shapes can be complex, sets not lines, may not intersect.

Stable Nash Equilibria

- Suppose for the game in Figure 222(b) we have P_1 play first, then P_2, followed by P_1, and so on
- Suppose that we number the moves with an index k.
- Suppose that each player knows the other's last move, and takes an action that minimizes its losses given this past move.
- For an arbitrarily chosen first decision by P_1, the "trajectory" in decision-space is shown in Figure 223.
- Since the reaction curves were already computed for this case it is easy to construct the trajectory since once P_1 makes a

decision $\theta^1(k)$ at iteration k (k odd), we have

$$\theta^2(k+1) = R_2(\theta^1(k))$$

and once P_2 makes a decision $\theta^2(k)$ at iteration k (k even), we have

$$\theta^1(k+1) = R_1(\theta^2(k))$$

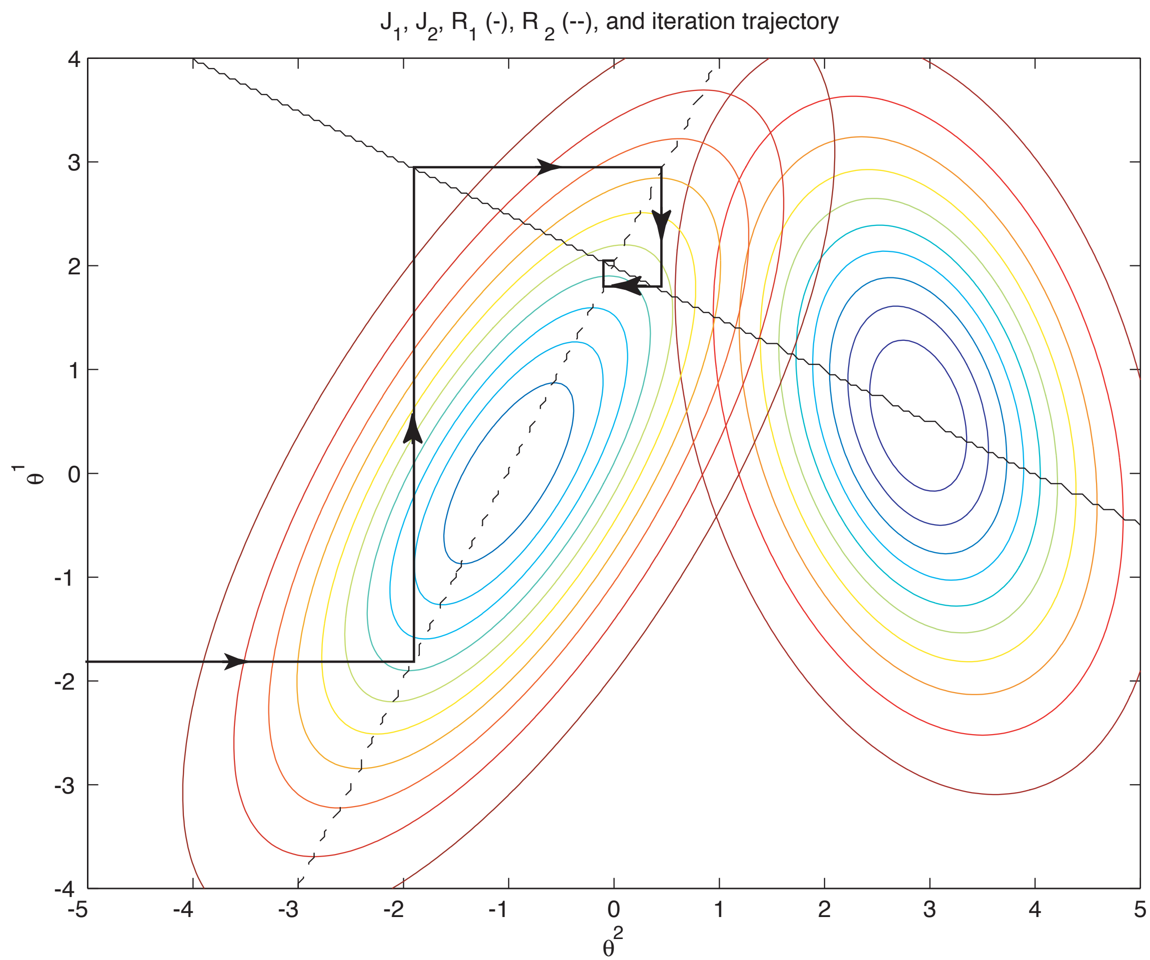

Figure 223: Contour plots of J_1 and J_2, reaction curves R_1 (solid) and R_2 (dashed), and iteration trajectory (arrows indicate direction of time).

- ➙ If for every initial choice by P_1 (i.e., $\theta^1(1)$), the trajectory in the decision space moves to the point $(2,0)$ then the point $(2,0)$ is said to be a "stable Nash equilibrium."

Minimax Strategies

- ➙ To find the minimax strategy you simply find the security strategy for P_1 based on J_1^{ij} and the security strategy for P_2 based on J_2^{ij} and then taken together these directly provide a strategy pair that is the "minimax" strategy for the bimatrix game.

  - P_1 (P_2) does not need knowledge of J_2^{ij} (J_1^{ij}) to compute its strategy.
  - The minimax concept essentially ignores whether the opponent is rational or not.
  - These facts can be important in practical applications.

- Can lead to very conservative strategies in some applications.
- There are also "Stackelberg strategies" ("leader-follower")

Cooperation and Pareto-Optimal Strategies

- Above, assumed games to be noncooperative so we assumed that there was a type of adversarial relationship between the two players.
- There are, however, games where the two players may be able to share information to try to do better; that is, they may cooperate.

Multiobjective Optimization and Pareto Optimality

➙ A *multiobjective optimization problem* is in the form of

$$\text{minimize: } \{J_1(\theta), \ldots, J_N(\theta)\}$$
$$\text{subject to: } \theta = [(\theta^1)^\top, \ldots, (\theta^N)^\top]^\top \in \Theta$$

➙ *Simulataneously* minimize a *set* of cost functions (called a "cost function vector") by changing θ.

- A general optimization problem.

- Here, we assume that $\theta^i = [\theta_1^i, \ldots, \theta_n^i]^\top$, $i = 1, 2, \ldots, N$, so that decisions are $n \times 1$ vectors, rather than just scalars, and there are N costs to minimize (e.g., in a two-player game we will have $N = 2$).

➙ A decision vector $\theta^* \in \Theta$ is *Pareto optimal* if there does not exist any other $\theta \in \Theta$ such that

$$J_i(\theta) \leq J_i(\theta^*)$$

for all $i = 1, 2, \ldots, N$, and at the same time

$$J_j(\theta) < J_j(\theta^*)$$

for at least one index j.

- A cost function vector is called Pareto optimal if the corresponding decision vector is Pareto optimal.

➙ A cost function vector is Pareto optimal if you cannot improve one cost value without degrading others.

➙ There can be many Pareto optimal solutions.

- In multiobjective optimization there is a need to specify *preferences* to be able to pick which Pareto optimal solution specifies an acceptable solution (e.g., one that balances the wins and losses of two players).
- Is there a "decision maker" who will specify these preferences in some manner?
- Decision maker may provide a "value function" to say what is best.

➙ Example: θ^1 and θ^2 are scalars, so that θ is a 2×1 vector.

- Quadratic costs so they are convex.
- Suppose

$$J_1(\theta) = J_1(\theta^1, \theta^2) = \left(\theta^1 - 2\right)^2 + \left(\theta^2 - 3\right)^2$$

$$J_2(\theta) = J_2(\theta^1, \theta^2) = \left(\theta^1 + 2\right)^2 + \left(\theta^2 + 2\right)^2$$

- Contour plots in Figure 224.

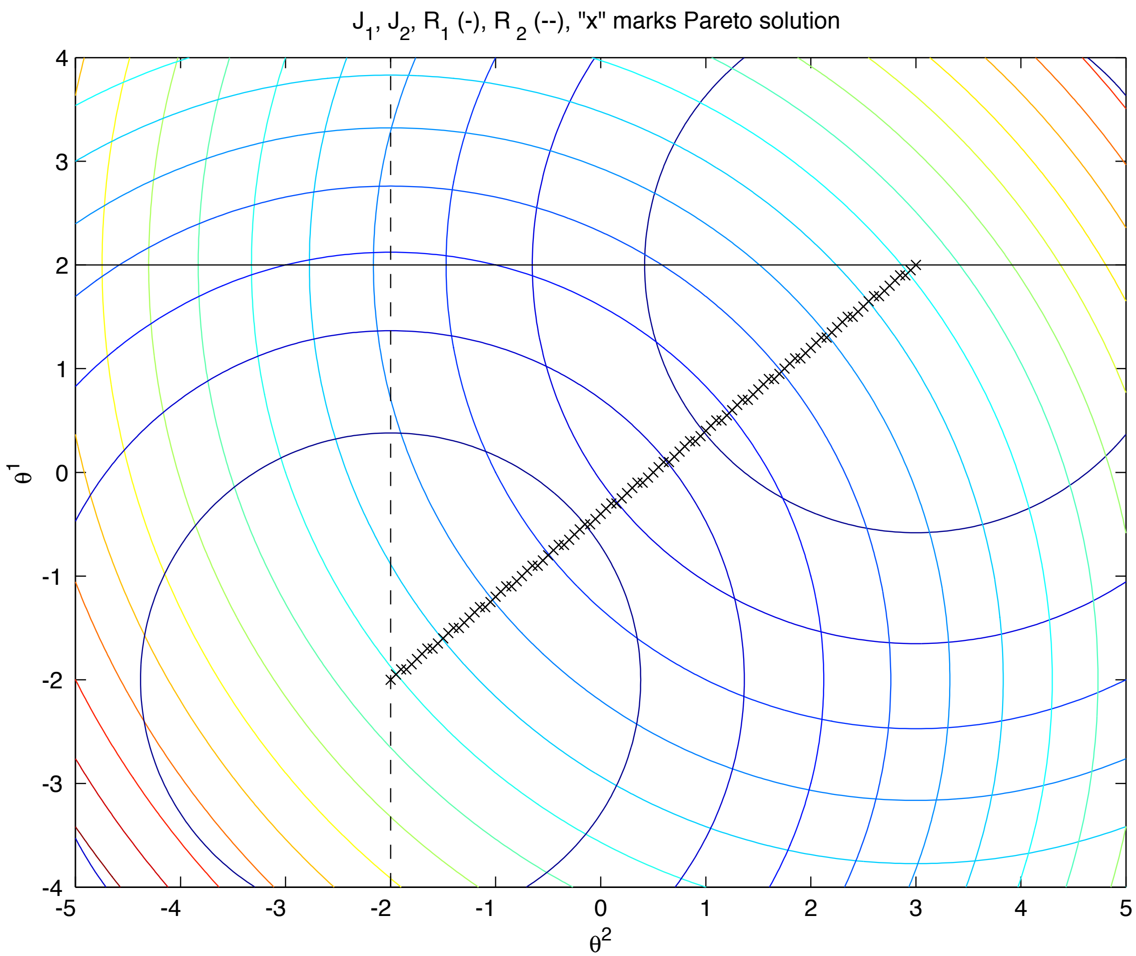

Figure 224: Example family of Pareto-optimal points for two quadratic cost functions ("$\times$" marks Pareto solutions).

➙ If a line on this contour plot is tangent to a contour of *both* costs, then the point of tangency for both costs is a Pareto optimal solution.

- Why? Suppose gradient at such a point θ^* is in the opposite direction for each cost
- Imagine that you are at some Pareto optimal solution θ^* in Figure 224.
- The direction of the negative gradient is the direction to move from θ^* in order to get a steepest amount of decrease in the value of one cost function
- The key observation is that if you perturb θ^* along the gradient in direction of the minimum point for J_1 (J_2) the cost for J_1 (J_2) goes down, but the cost for J_2 (J_1) goes up.
- So, we cannot reduce one cost without increasing the other, which is the very definition of Pareto optimality.

- The set of all Pareto optimal solutions is sometimes called the "family" of Pareto optimal solutions.

➙ If we define a "Pareto cost" (value function) to be

$$J_p(\theta) = pJ_1(\theta) + (1-p)J_2(\theta)$$

for $p \in [0,1]$ (this is called the "scalarization" approach to constructing the Pareto cost), then the family of Pareto points is the set of (unique) global minima for $J_p(\theta)$ as p varies from zero to one, which is just the equation for the line between the two minimum points in Figure 224.

Pareto Optimal Solutions for Games

- A standard way to define Pareto-optimality is

$$J_p^{ij} = pJ_1^{ij} + (1-p)J_2^{ij}$$

for $p \in [0,1]$.

- *Any* minimum point in the matrix J_p^{ij} is called a Pareto-optimal solution for the bimatrix game (and note that there may be several minimum points for any one value of p).
- If $p = 1$ ($p = 0$), all the emphasis is placed on the two players collaborating to minimize the losses of P_1 (P_2).

Defining the Pareto Cost and Finding Pareto Solutions

- Problems in defining Pareto costs and computing Pareto solutions...

➙ Example: Figure 222(b) and analgous to the above example, let

$$J_p(\theta^1, \theta^2) = pJ_1(\theta^1, \theta^2) + (1-p)J_2(\theta^1, \theta^2)$$

for $p \in [0, 1]$.

- To give insight into the shape of the cost surface J_p see Figure 225 which is the case for $p = 0.5$.

- Finding the global minimum can in general be challenging.
- Also, note that the p parameter will in this case scale the "depth" of the two minima.

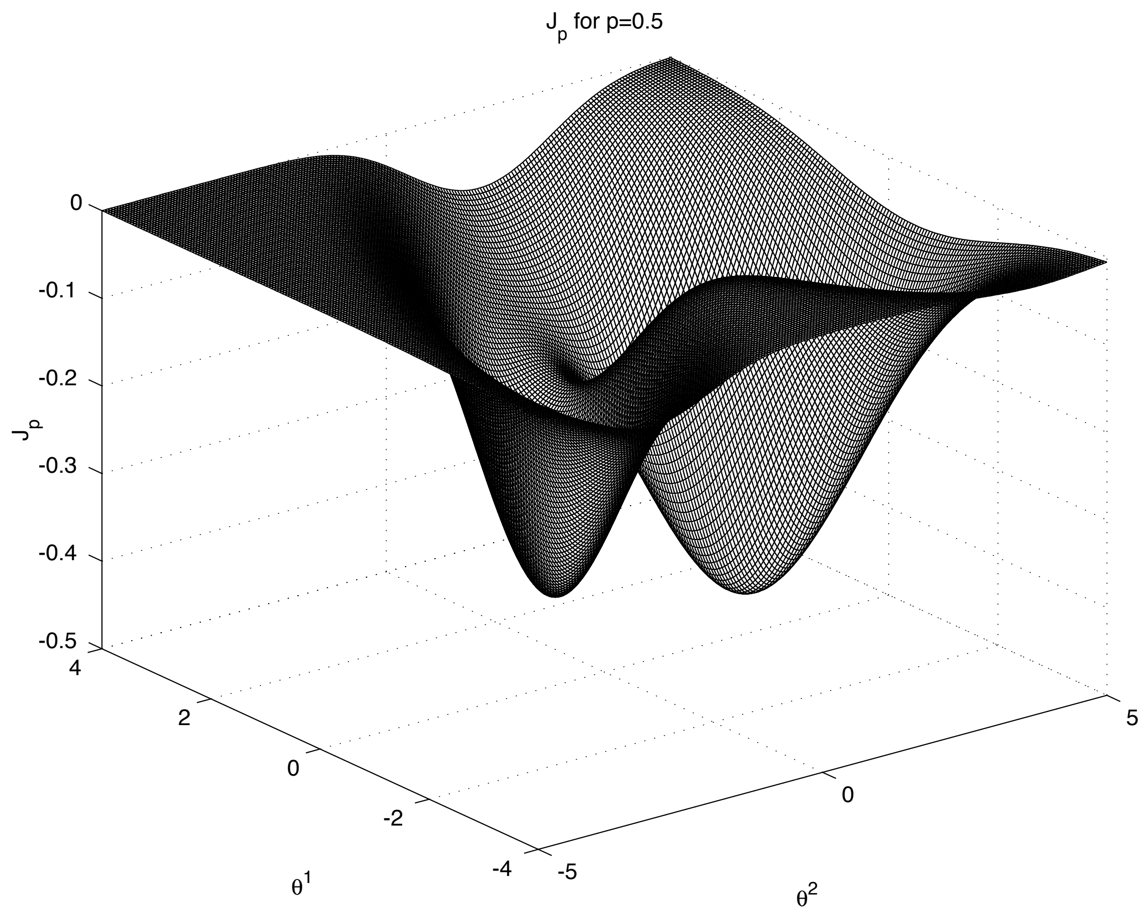

Figure 225: Pareto cost $J_p(\theta^1, \theta^2)$ for $p = 0.5$.

Design Example: Static Foraging Games

- Study the basics of competition and cooperation in static "foraging games"

Static Foraging Game Model

➙ Two-forager ($N = 2$), static, discrete, full-information "foraging game on a line."

- Resources are distributed in "cells" (bins) along the real line.
- M different types of resources in Q cells and denote the intial distribution of resources of type m to be $r^m(q)$, $q = 1, 2, \ldots, Q$, $m = 1, 2, \ldots, M$.
- Here, we assume that $r^m(q) \geq 0$, $q = 1, 2, \ldots, Q$, but the model is easily extended to the negative resource case (where one could think of moving to avoid regions where resources are lost).

- Let D_1 (D_2) be the number of decisions that forager 1 (2) can make and $\theta^1 \in \{1, 2, \ldots, D_1\}$ ($\theta^2 \in \{1, 2, \ldots, D_2\}$) be those decisions, which correspond to forager 1 (2) moving to a cell q if $\theta^1 = q$ ($\theta^2 = q$), $q = 1, 2, \ldots, Q$.
- $D_1 = D_2 = Q$ so each forager can move to any available cell.

Effort and Resource Consumption

➙ Let z_1 (z_2) denote the effort allocated by forager 1 (2) to consume resources.

- For simplicity, we assume that the same amount of effort is expended for consumption of each resource type $m = 1, 2, \ldots, M$ when a forager goes to a cell.

➙ Let $P(q)$ be the set of foragers that decide to go to the same position q to consume resources there,

$$P(q) = \left\{ i : \theta^i = q \right\}$$

- $0 \leq |P(q)| \leq N$ for all q and $\sum_{i \in P(q)} z_i$, the total consumption effort at q, is zero if $|P(q)| = 0$.
- Assume α^m, $m = 1, 2, \ldots, M$, is used to model the depletion rate of resource m in the presence of consumption effort.

➙ We model the amount of resource of type m remaining at cell q after one play (one unit of expenditure of effort) as

$$r^m(q)e^{-\alpha^m \sum_{i \in P(q)} z_i}$$

- Initial expenditures of effort in a cell yield more resources than later ones.
- Define the amount of *consumption* given that a strategy pair (θ^1, θ^2) is played by foragers 1 and 2.
- If both foragers are in the same cell expending effort to consume the same resource, then they have to split the resource since there is a type of competition for it.

- Here, assume that they split the resources evenly.

➙ Let the amount of consumption of resource m for decision pair (θ^1, θ^2) for foragers 1 and 2 be defined as follows:

1. Foragers at different locations: If $\theta^1 \neq \theta^2$, then for $i = 1, 2$,

$$C_i^m(\theta^1, \theta^2) = r^m(\theta^i)\left(1 - e^{-\alpha^m z_i}\right)$$

2. Foragers at the same location: If $\theta^1 = \theta^2 = \bar{\theta}$, then for $i = 1, 2$,

$$C_i^m(\theta^1, \theta^2) = \frac{1}{|P(\bar{\theta})|} r^m(\bar{\theta})\left(1 - e^{-\alpha^m \sum_{i \in P(\bar{\theta})} z_i}\right)$$

$$= \frac{1}{2} r^m(\bar{\theta})\left(1 - e^{-\alpha^m (z_1 + z_2)}\right)$$

- So, in cases where forager 1 (2) goes to a cell that forager 2 (1) does not go to, $r^m(\theta^1)$ $(r^m(\theta^2))$ is the initial amount of resource of type m and $r^m(\theta^1)e^{-\alpha^m z_1}$ $(r^m(\theta^2)e^{-\alpha^m z_2})$ is the amount remaining after consumption.

- When both foragers go to the same cell, then they both expend effort, but they have to split the returns in half.

➙ This results in a resource conservation property of: "all that is consumed plus what is remaining is equal to what was initially there."

Forager Payoffs: Consumption, Energy, and Danger Avoidance

➙ We assume that each forager has certain priorities to consume different resources.

- We denote these by p_1^m (p_2^m) for forager 1 (2), $m = 1, 2, \ldots, M$.

➙ You can think of these priorities as representing preferences or "tastes" for resources.

- One aspect of the cost to forager 1 (2) that it wants to minimize is given by the negative total consumption weighted by the priorities

$$
\begin{aligned}
J_{1c}^{ij} &= J_{1c}(\theta^1, \theta^2) = -\sum_{m=1}^{M} p_1^m C_1^m(\theta^1, \theta^2) \\
J_{2c}^{ij} &= J_{2c}(\theta^1, \theta^2) = -\sum_{m=1}^{M} p_2^m C_2^m(\theta^1, \theta^2)
\end{aligned}
$$

where $\theta^1 = i$, $\theta^2 = j$, and J_{1c}^{ij} and J_{2c}^{ij} constitute a matrix representation of the game.

- The problem for forager 1 (2) is how to pick θ^1 (θ^2).

➙ The adversarial nature of the foraging game will dictate what to choose (e.g., in a competitive game each forager may get less than if they cooperate).

- Think of the foragers as being located at position "0" (i.e., on one edge outside the foraging area) initially.

- Model the cost to move along to line to go to position i (j) for forager 1 (2) as

$$J_{1e}^{i} = J_{1e}(\theta^1) = w_{e1} i \quad \left(J_{2e}^{j} = J_{2e}(\theta^1) = w_{e2} j \right)$$

where $\theta^1 = i$, $\theta^2 = j$, $w_{e1} \geq 0$ and $w_{e2} \geq 0$ represent the unit amount of energy expenditure to move one unit (e.g., from cell 1 to cell 2).

- Can be location-dependent dangers for forager 1 (2) represented with

$$J_{1d}^{i} \geq 0 \; \left(J_{2d}^{j} \geq 0 \right)$$

where bigger values of the costs represent worse areas to be in and actions of the other forager do not affect the danger to a forager.

- We can think of forager 1 (2) trying to minimize

$$J_1^{ij} = J_{1c}^{ij} + J_{1e}^{i} + J_{1d}^{i} \quad \left(J_2^{ij} = J_{2c}^{ij} + J_{2e}^{j} + J_{2d}^{j} \right)$$

so that it maximizes the amount of resources it gets and minimizes the energy expenditure and exposure to dangers to get them.

➙ Here, each forager knows everything about the game (e.g., the payoffs, costs of movement, dangers, the other forager's objectives, etc.).

Competition and Cooperation for a Resource

- Choose $D_1 = D_2 = Q = 21$, $M = 1$, $z_1 = z_2 = 1$, $\alpha^1 = 1$, $p_1^1 = p_2^1 = 1$, and $w_{e1} = w_{e2} = 0$ (no energy required for foraging).
- Assume that $J_{1d}^i = J_{2d}^j = 0$ for all i and j.
- The initial resource distribution is shown in Figure 226.

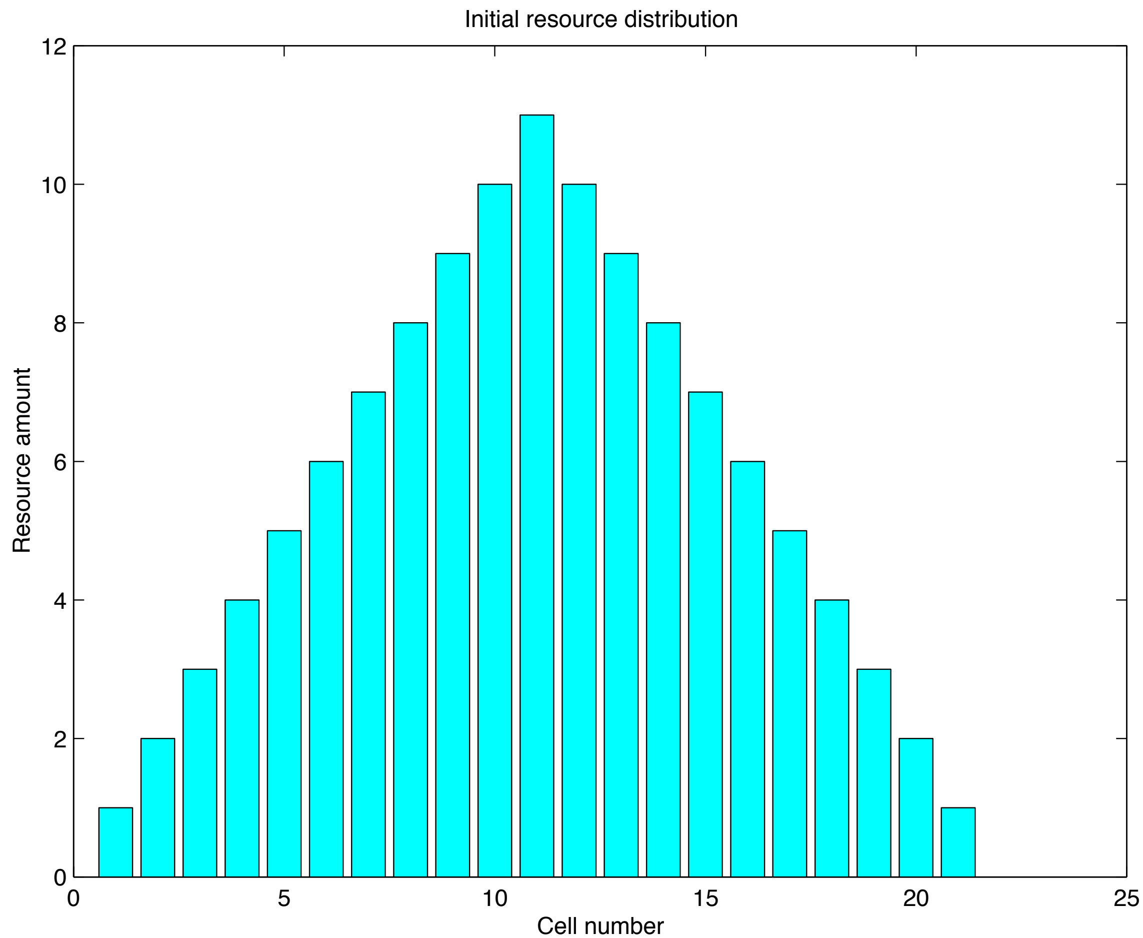

Figure 226: Example initial resource distribution.

- The cost functions J_1^{ij} and J_2^{ij} are plotted in Figures 227 and 228.

➙ Notice in Figure 227 that if you hold j constant, then forager 1 generally gets more consumption and hence more payoff if it moves to where there are more resources; however, if both foragers move to the same location they get less since they will then compete for resources at that cell.

- This competition is represented by the ridges of increased cost.

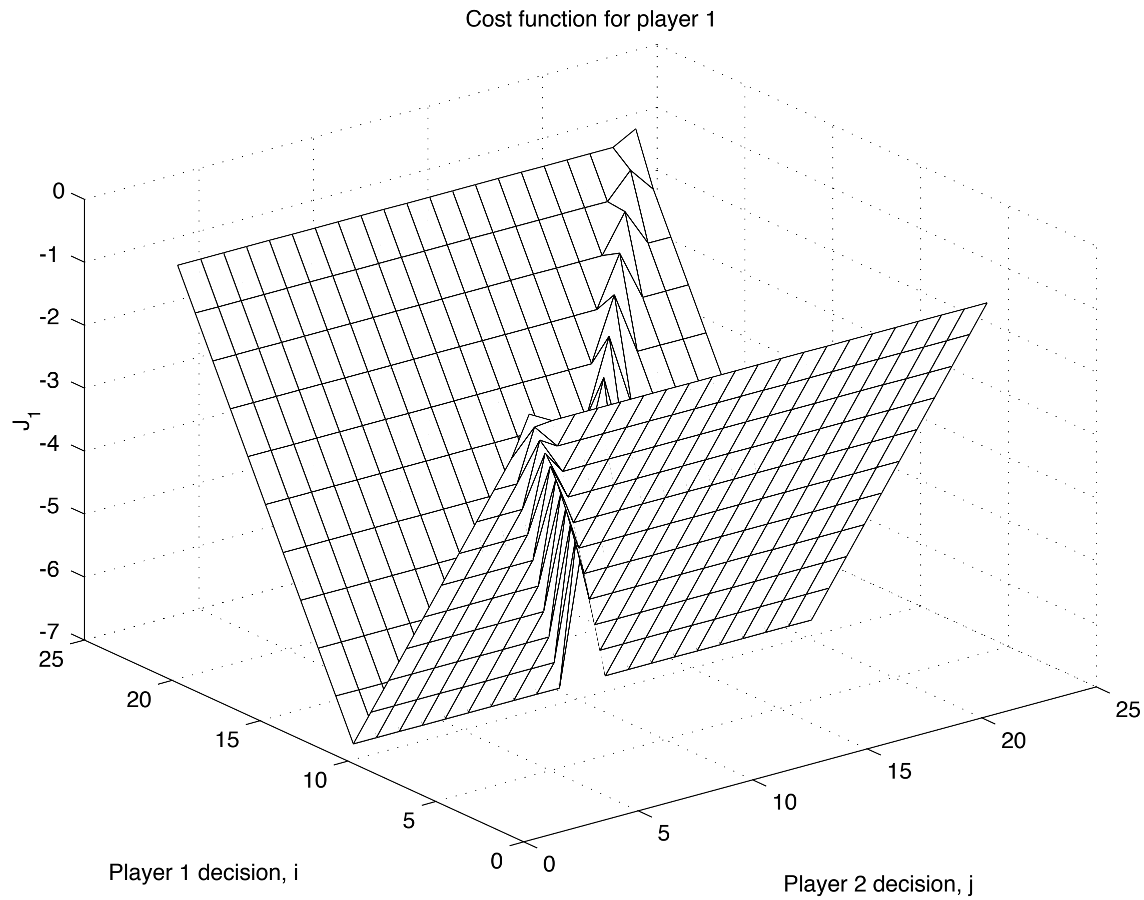

Figure 227: Cost for forager 1, J_1^{ij}.

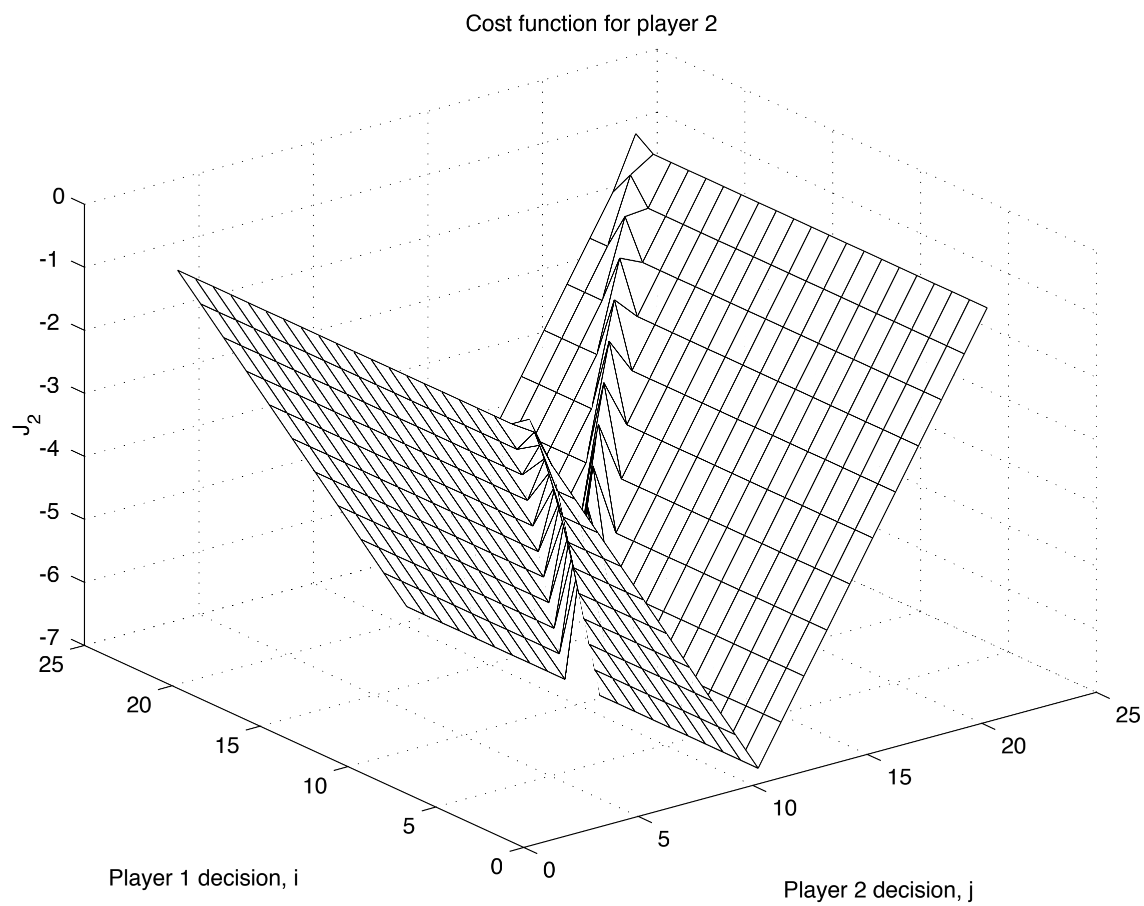

Figure 228: Cost for forager 2, J_2^{ij}.

- First, suppose that we have an adversarial (noncooperative) game so that the foragers do not coordinate where to go to forage.

- There are four Nash solutions

$$(10, 11), (11, 10), (11, 12), (12, 11)$$

- Does this make sense?
- From Figure 226 the cell with the most resources is cell 11.
- Note, however, that the problem of nonunique Nash solutions arises. How?

➙ A cooperative foraging game...

- Suppose that the two foragers cooperate by using a Pareto cost found via scalarization as $J_p^{ij} = pJ_1^{ij} + (1-p)J_2^{ij}$ with p as the Pareto parameter.
- Pareto points found in this case are shown in Figure 229.

- We get Pareto points (which are also Nash solutions) $(10, 11)$ or $(11, 10)$ depending on the Pareto parameter p.
- The two foragers would communicate to decide who goes to which location, which as opposed to the Nash game, is possible since the two foragers are cooperating.
- The one that goes to position 11 will get the most resources.

➙ When p is close to zero it favors forager 2 so forager 2 goes to position 11, and when p is close to one it favors forager 1 so forager 1 goes to position 11.

- The p parameter can be used to balance the cooperation to favor one forager or the other.

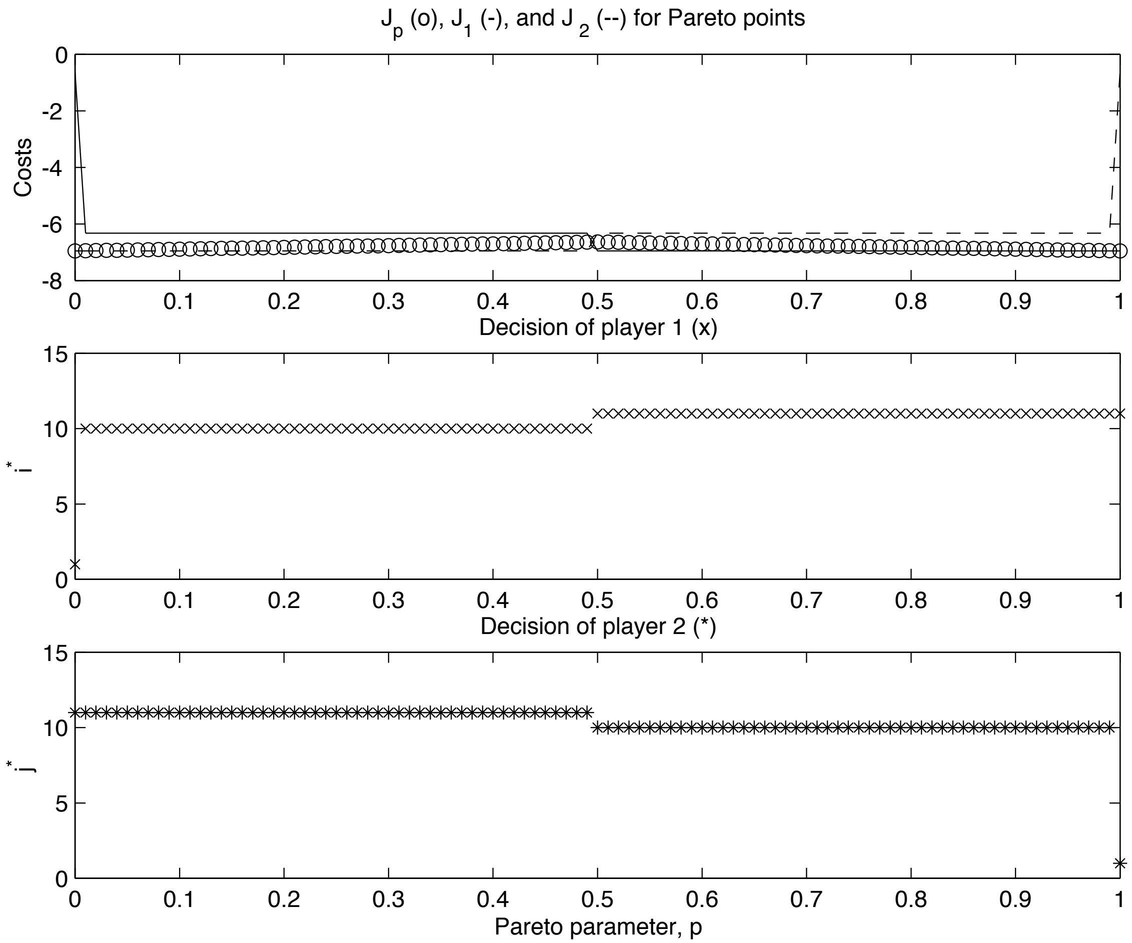

Figure 229: Set of all Pareto points for a cooperative foraging game, scalarized Pareto cost.

- The scalarization approach is, however, only one way to form the Pareto cost.
- The set of all Pareto points for the game is shown in Figure 230 and you can see that the ones that arise from the above scalarization approach are a subset of all possible Pareto points.
- These other Pareto points represent different ways to balance the payoffs to each of the two foragers.

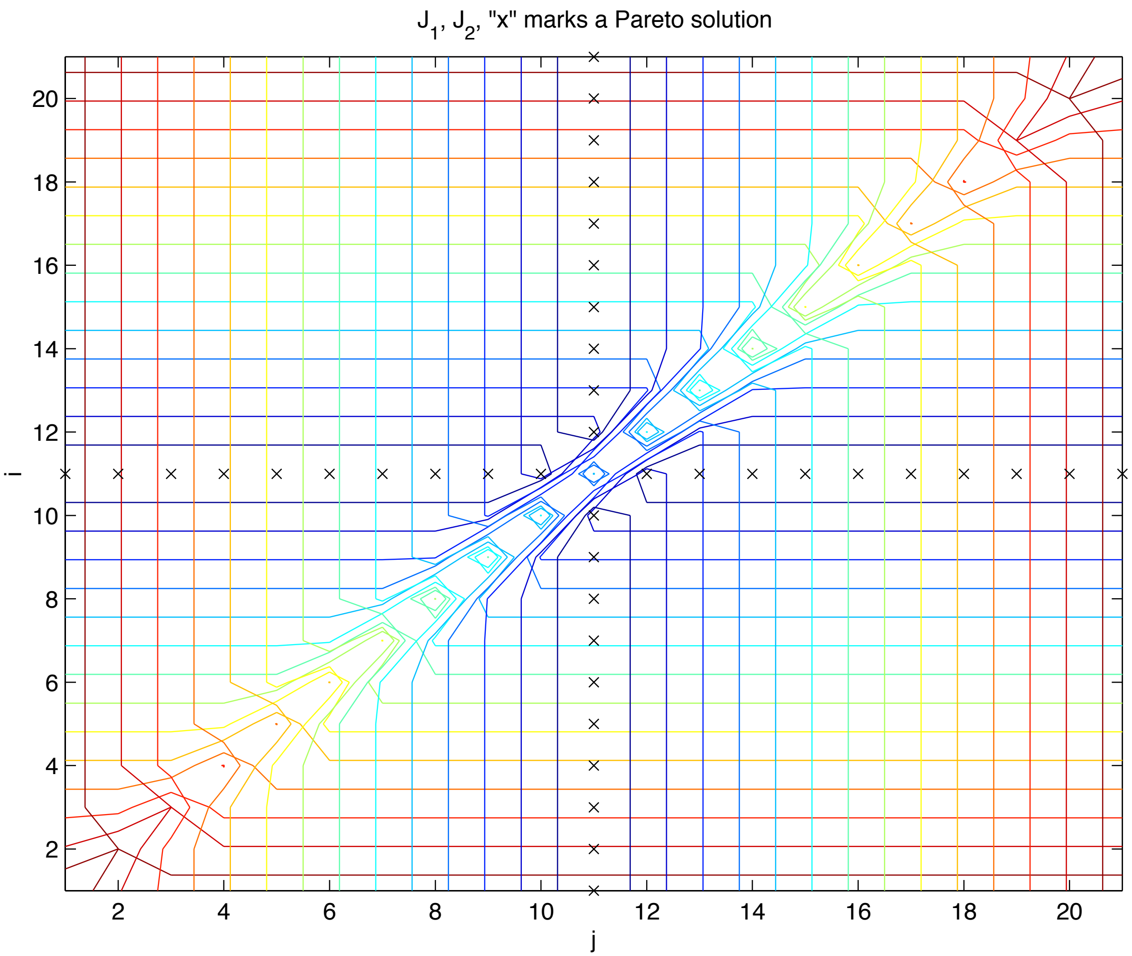

Figure 230: Set of all Pareto points for a cooperative foraging game.

➙ You can view a cooperative foraging game as one where you try to *allocate* resources to all the foragers so that everyone "wins," with the relative payoffs given by which Pareto points you choose.

Energy-Constrained Competition and Cooperation for Two Resources

- Choose $M = 2$, $p_1^1 = p_1^2 = 1$, $p_2^1 = 1$, and $p_2^2 = 2$ so that forager 2 places a high priority on getting resource 2.
- We let $w_{e1} = w_{e2} = 0.1$ so that moving to cell locations with higher values of q costs more energy.
- As before we have $D_1 = D_2 = Q = 21$, $z_1 = z_2 = 1$, $\alpha^1 = \alpha^2 = 1$, and $J_{1d}^i = J_{2d}^j = 0$ for all i and j.
- The initial resource distribution is in Figure 231.

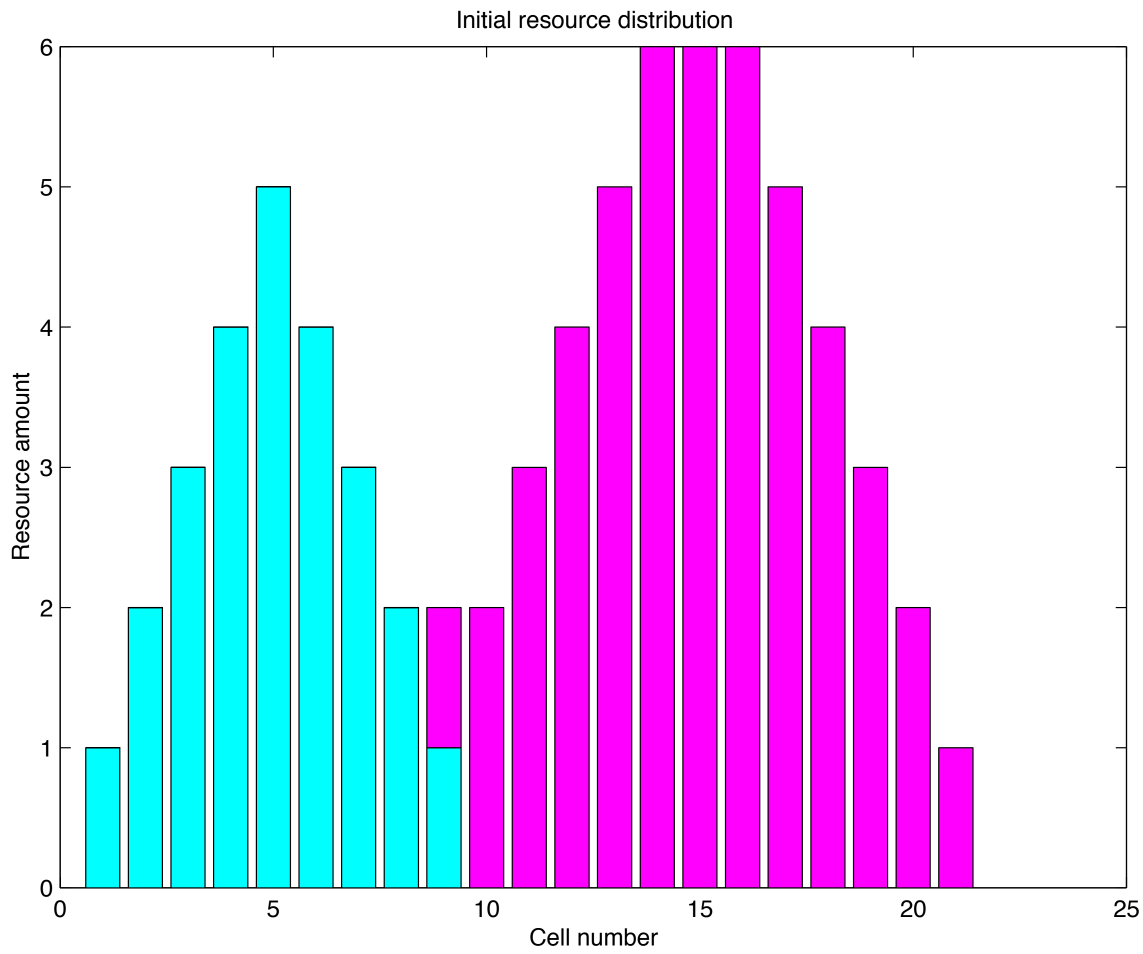

Figure 231: Initial resource distribution (darker shaded bars on the right are the second resource) with an "overlap" of resources in the middle designated by "stacking" the plots).

- The cost functions J_1^{ij} and J_2^{ij} are plotted in Figures 232 and 233.
- Focus on Figure 232 and notice that even though it sets an equal priority for both resources the costs generally increase as i increases (ignoring the ridge) due to the presence of the J_{1e}^i term that represents the energy needed to forage at each position.
- This raises the cost of the second resource.

➙ Notice that in Figure 233 we have the presence of this same effect, *and* the effect of the higher priority of resource 2 for forager 2 so that for forager 2, even though it has to travel further to get resource 2, since it is higher priority it may be willing to do that.

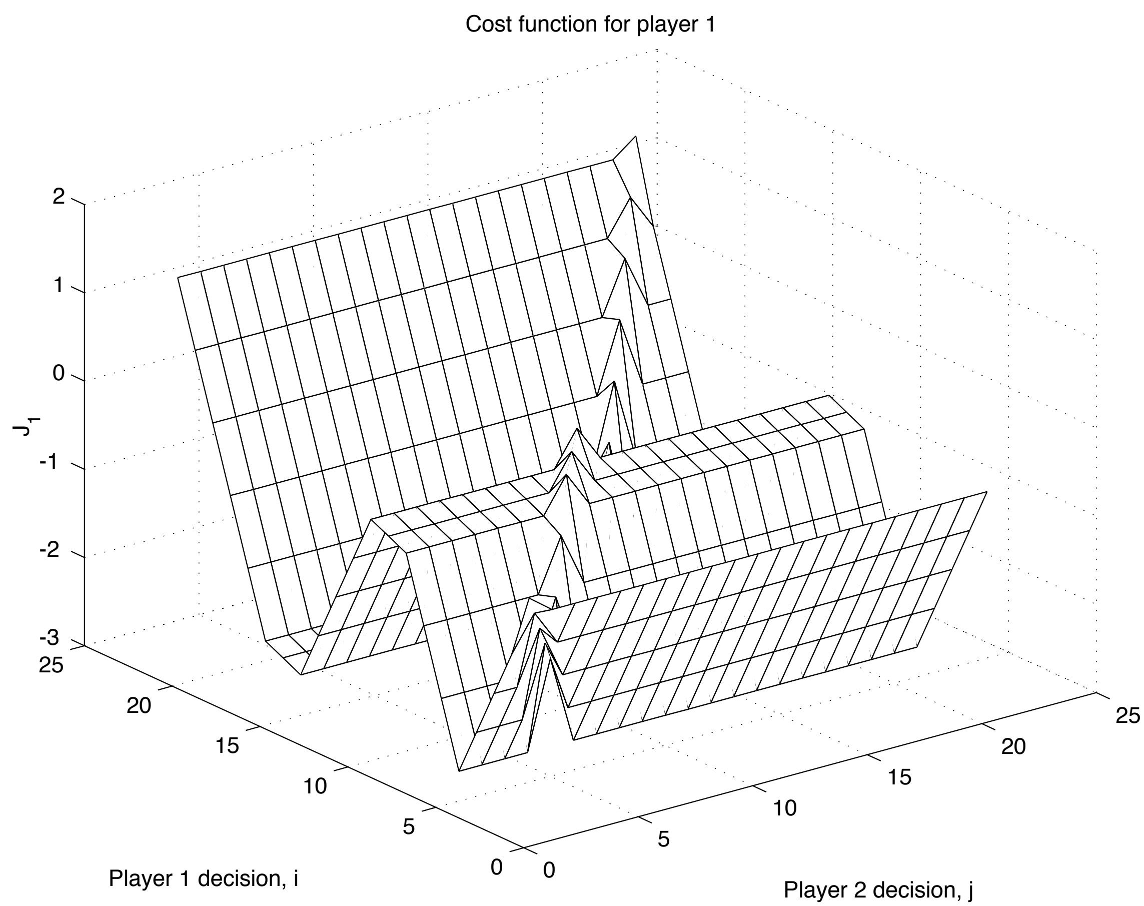

Figure 232: Cost for forager 1, J_1^{ij}.

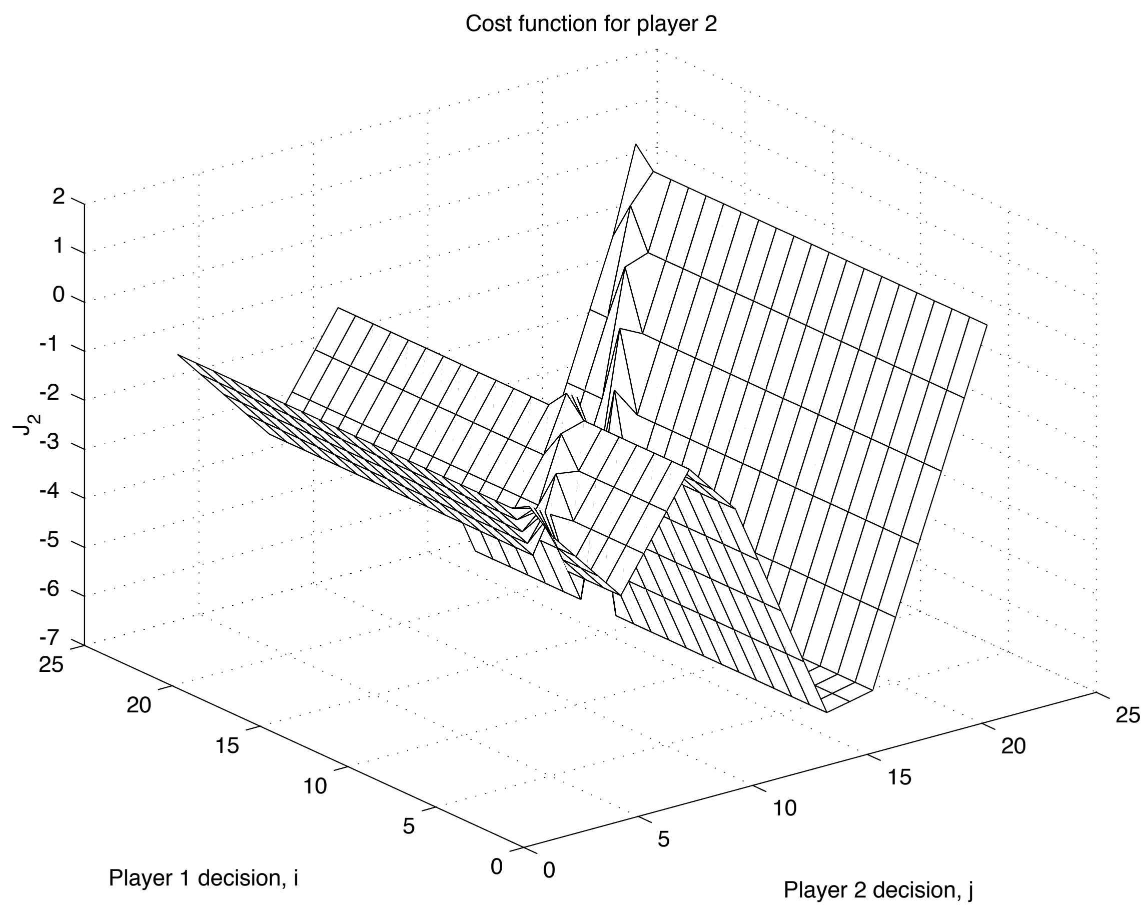

Figure 233: Cost for forager 2, J_2^{ij}.

- Consider the competitive case first.

- The unique Nash solution is $(5, 14)$.
- Essentially, with the above choices, forager 1 chooses resource 1 since it is close, but forager 2 picks resource 2 since its level of priority is high so it is willing to expend the energy to get it.
- Notice that the maximum for the second resource is achieved at three contiguous positions, but forager 2 it picks the smallest of these to minimize energy.
- If the foragers enter into a cooperative game, with a scalarized Pareto cost $J_p^{ij} = pJ_1^{ij} + (1-p)J_2^{ij}$, then we get the Pareto solutions all at $(5, 14)$ for all p.
- The two forager's objectives are so different that there is nothing to be gained by cooperation (and nothing to be lost by competition) and hence there is really no need for communication.

Dynamic Games

➛ Dynamic games consider repeated decisions, actions, and observations by the players.

Modeling the Game Arena and Observations

- N players, a discrete time formulation.
- Let
$$x(k) \in X \subset \Re^{n_x}$$
denote the state of the game at time k, $k \geq 0$.
- The admissible controls (actions) by player i are for $k \geq 0$
$$u^i(k) \in U^i(k) \subset \Re^{n_u}$$
- The outputs (measurements of what is happening in the game) are, for $k \geq 0$,
$$y^i(k) \in Y^i(k) \subset \Re^{n_y}$$

- Let

$$u(k) = \left[(u^1(k))^\top, (u^2(k))^\top, \ldots, (u^N(k))^\top\right]^\top$$

and

$$y(k) = \left[(y^1(k))^\top, (y^2(k))^\top, \ldots, (y^N(k))^\top\right]^\top$$

➙ Define the "arena" in which the game is played as f where

$$x(k+1) = f(x(k), u(k), k) \tag{112}$$

and suppose that the intial state of the game is $x(0) \in X$.

- This is a deterministic game model, but it can be time varying.
- A stochastic game: $x(k+1) = f(x(k), u(k), w(k), k)$

➙ The observations that player i can make about the arena of the game are $y^i(k) = h^i(x(k), k)$ for $k \geq 0$, and if we let $h(x(k), k) = \left[(h^1)^\top, (h^2)^\top, \ldots, (h^N)^\top\right]^\top$, then

$$y(k) = h(x(k), k) \tag{113}$$

- You could view h as part of the representation of the arena of the game as it models what can be observed by each player while the game is played.
- Observations lead to decisions, which lead to actions, which lead to observations, and so on.
- Let $J_i(x(k), u(k))$ denote the loss (cost) function of of the i^{th} player at the k^{th} stage of play.

➙ Multiple stages of play:

$$J_i^{N_s} = \sum_{k=0}^{N_s-1} J_i(x(k), u(k)) \tag{114}$$

- Players choose sequences $u^i(k)$ to minimize $J_i^{N_s}$

- Another sequence

$$J_i^{N_s} = \sum_{k=1}^{N_s-1} J_i(x(k+1), x(k), u(k)) \tag{115}$$

- In general player i does not know $J_i(x(k), u(k))$ since it may not know $x(k)$ and $u(k)$.

Information Space and Strategies

- How are player's strategies defined?
- More complicated than in the static game case. Why?

➙ Each player may make decisions based on "what they know and when they know it" and hence it is not assumed that each player knows everything at one time and only one action is taken by each player at that time.

- If a player has memory it can store and recall past observations.

- Then, for any player i, at the k^{th} stage of play, it may base its decisions to choose $u^i(k)$ on a subset of

$$\Big\{y^1(0), \ldots, y^1(k); \cdots; y^N(0), \ldots, y^N(k);$$
$$u^1(0), \ldots, u^1(k-1); \cdots; u^N(0), \ldots, u^N(k-1)\Big\}$$

➙ Each such subset is called an "information structure"

- Let

$$\begin{aligned} I^i(k) \quad \subset \quad & (Y^1(0) \times \cdots \times Y^1(k)) \times \cdots \times (Y^N(0) \times \cdots \times Y^N(k)) \times \\ & (U^1(0) \times \cdots \times U^1(k-1)) \times \cdots \times (U^N(0) \times \cdots \times U^N(k-1)) \end{aligned}$$

denote the "information space" of player i at time $k \geq 0$

- $I^i(k)$ is implemented via an appropriately-defined communication network between the players and memory within each player to store past values.

- In an adversarial game there may be no communication links between the players, but memory.

- Cooperation requires communications? Cues (e.g., via environment)? Signals?
  - Game constraints may specify $I^i(k)$, but other times the designer may be able to choose it.
- Example: If all players only make decisions based on their own current observations of the arena of play and the previous actions of all other players and itself, then

$$I^i(k) \subset (Y^i(k)) \times (U^1(k-1) \times \cdots \times U^N(k-1))$$

- A strategy for a player i at the k^{th} stage of play is G_k^i, $k \geq 0$,

$$G_k^i : I^i(k) \rightarrow U^i(k)$$

  - The design of strategies involves designing *both* $I^i(k)$ and G_k^i.
- Example: If each player i can observe at stage k, only $y^i(k)$ (i.e., its only observation) and all actions $u^i(k-1)$,

$i = 1, 2, \ldots, N$, the strategies of the players are defined via a G_k^i mapping for each player that specifies its actions,

$$u^i(k) = G_k^i(y^i(k), u^1(k-1), \ldots, u^N(k-1))$$

(and at $k = 0$ there are no elements in the u^i slots).

- "Full state feedback" case: $y^i(k) = x(k)$ for all i and k ("perfect information")
- The standard concepts of saddle point, Nash, and Stackelberg equilibrium solutions can be extended to dynamic games.
- There are additional solution concepts depending on the information space that is assumed (e.g., the "feedback Nash solution").
- Most common: linear systems with quadratic cost

Decision/Action Timing

- Finite number of N_s stages (time steps), or $k \rightarrow \infty$?

➙ Actions need not be *synchronous*, could be asyncrhonous ı

➙ tem All players above act at each time k (but can define a "null play")

Example: Modeling Dynamic Foraging Games

Dynamic Foraging Game Model

State and Inputs

- $N \geq 2$ foragers.
- The state $x \in \Re^{n_x}$ is composed of aspects of the foraging environment and the positions of the foragers in that environment.
- Assume that you have a two-dimensional foraging environment (a "foraging plane").

➙ The position of the i^{th} forager is given by

$$x^i(k) = \left[x_1^i(k), x_2^i(k)\right]^\top \in \{1, 2, \ldots, Q_1\} \times \{1, 2, \ldots, Q_2\} = F$$

with $x_1^i(k)$ its horizontal and $x_2^i(k)$ its vertical position on a discrete grid.

- The decisions by forager i are commands to move itself to each of the cells that are adjacent to the current position, and which resource type to consume there.
- That is, at time k, so long as the movement is in the valid foraging region so that $x^i(k) \in F$, we have that $u_p^i(k)$ is in the set

$$\left\{ \left[x_1^i(k), x_2^i(k) \right]^\top, \left[x_1^i(k)+1, x_2^i(k) \right]^\top, \ldots, \left[x_1^i(k)+1, x_2^i(k)+1 \right]^\top \right\} \bigcap F$$

which we will denote by $U_p^i(k)$.

- $u_p(k) = \left[(u_p^1(k))^\top, (u_p^2(k))^\top, \ldots, (u_p^N(k))^\top \right]^\top$.
- M resources indexed with m.
- We represent the choice of resource by player i at time k as $u_r^i(k)$, $i = 1, 2, \ldots, N$, where

$$u_r^i(k) \in U_r^i(k) \subset \{1, 2, \ldots, M\}$$

represents the resource type m that forager i chooses to consume at time k, and $U_r^i(k)$ can be used to model the set of resources that it can choose from.

- $u_r(k) = \left[u_r^1(k), u_r^2(k), \ldots, u_r^N(k)\right]^\top$.
- The decision $u^i(k)$ of forager i at time k is

$$u^i(k) = \left[(u_p^i(k))^\top, u_r^i(k)\right]^\top \in U_p^i(k) \times U_r^i(k)$$

  and $u(k) = \left[(u^1(k))^\top, (u^2(k))^\top, \ldots, (u^N(k))^\top\right]^\top$.
- The distribution of resources is also part of the state.
- Let

$$q = [q_1, q_2]^\top \in F$$

  denote a cell in the foraging plane.
- Let z_i^m denote the effort allocation to consume resource m by forager i.

- Let
$$P^m(q) = \left\{ i : u_p^i = q, u_r^i = m \right\}$$
be the set of foragers that decide to go to position q to consume resource m at time k.

- Notice that $0 \leq |P^m(q)| \leq N$, but below we will only use $P^m(q)$ for $q = u_p^i$ for some $i = 1, 2, \ldots, N$, so $|P^m(q)| > 0$.

- We use the depletion rate α^m, $m = 1, 2, \ldots, M$, for the m^{th} resource.

- The amount of resource at time k of type m at cell q is $r^m(q, k)$ with $r^m(q, 0)$ the intitial distribution.

- The resources change over time due to growth (e.g. plants), weather, disease, farming, and foraging.

- For foraging, resources may diminish due to consumption, and in some cases such consumption may result in the increase of other resources (e.g., since the resources may be living so foraging influences their competitive balance).

- In other cases foraging for one type of resource at one time may make it possible to forage for other resources later (e.g., if one forager eats one type of resource and this gives rise to other resources due to, for example, a forager leaving behind remains).
- Assume

$$r^m(q, k+1) \quad = \quad r^m(q, k) e^{-\alpha^m \sum_{i \in P^m(q)} z_i^m} \tag{116}$$

for all $q \in F$.

- For this equation notice that $P^m(q)$ is a function of u.
- Let

$$x_p(k) = \left[(x^1(k))^\top, (x^2(k))^\top, \ldots, (x^N(k))^\top \right]^\top$$

denote the vector of places where the foragers are located.

- Let

$$x_r(k) = \begin{bmatrix} r^1([1,1]^\top, k) \\ \vdots \\ r^1([Q_1, Q_2]^\top, k) \\ \vdots \\ r^M([1,1]^\top, k) \\ \vdots \\ r^M([Q_1, Q_2]^\top, k) \end{bmatrix}$$

be a vector that holds a vectorized representation of the resource distribution (maps).

- The state of the game is

$$x(k) = \begin{bmatrix} x_p(k) \\ x_r(k) \end{bmatrix}$$

➙ Define next state...

- $x_p(k+1) = u_p(k)$ (assuming no dynamics and kinematics for our forager, or a rate assumption on movement).
- $x_r(k+1)$ is defined via Equation (116)

- Assume that the real time between k and $k+1$ is fixed so that the real time is $t = kT$ where T is a sampling period (needed for how we model resource depletion).
- The real time at the next sampling instant is $t' = kT + T$.
- But, still asychronous via use of "null plays"

Sensing and Outputs

- The physiology of the animal constrains what sensing is possible.
- Some animals can only sense via sampling chemicals in their immediate surrounding environment (e.g., certain bacteria),

while others can sense light or sound and hence "see" for long distances.

➙ Possibilities:

1. Full observations: Forager, $i = 1, 2, \ldots, N$, and time k,

$$y^i(k) = x(k) \tag{117}$$

2. Resource observations and own position:

$$y^i(k) = h^i(x(k), k) = \begin{bmatrix} x^i(k) \\ x_r(k) \end{bmatrix}$$

3. Range-constrained sensing: Let $S(q)$ denote the set of cell locations that a forager can sense resources in, or other forager positions, when it is located at cell q. Suppose that

$$S(q) = \left\{ \bar{q} : \sqrt{(q - \bar{q})^\top (q - \bar{q})} \leq R_s \right\} \bigcap F$$

First, form a vector of the forager locations, for foragers that can be sensed, from elements of x_p, as $x_p^{s_i}$ with elements $x^j(k)$ where $x^j(k) \in S(x^i(k))$ for all $j = 1, 2, \ldots, N$. Second, form a new vector of the currently sensed cells from elements of x_r, as $x_r^{s_i}$ with elements $r^m(q, k)$ where $q \in S(x^i(k))$ for all $m = 1, 2, \ldots, M$. If

$$y^i(k) = h^i(x(k), k) = \begin{bmatrix} x_p^{s_i} \\ x_r^{s_i}(k) \end{bmatrix}$$

- Many other possibilities...

Consumption, Energy, and Payoff to Foragers

- Define the amount of consumption of resource m by forager i, $i = 1, 2, \ldots, N$, at time k for a set of forager decisions u^1, u^2, $\ldots$, u^N, as

$$C_i^m(u^1(k), u^2(k), \ldots, u^N(k))$$

$$= \frac{1}{|P^m(u_p^i(k))|}\left(r^m(u_p^i(k),k) - r^m(u_p^i(k),k+1)\right)$$

$$= \frac{1}{|P^m(u_p^i(k))|} r^m(u_p^i(k),k)\left(1 - e^{-\alpha^m \sum_{i\in P^m(u_p^i(k))} z_i^m}\right)$$

- Notice that $|P^m(u_p^i(k))| > 0$.
- The factor $\frac{1}{|P^m(u_p^i(k))|}$ is used to represent splitting resources

➜ The cost due only to consumption for one move is, given that forager i has priority p_i^m for resource m,

$$J_{ic}(x(k+1),x(k),u(k)) = -\sum_{m=1}^{M} p_i^m C_i^m(u^1(k),u^2(k),\ldots,u^N(k))$$

➜ Each forager must expend energy to forage, and we define this via

$$J_{ie}(x(k+1),x(k)) = w_{ie}\left(x^i(k+1) - x^i(k)\right)^\top \left(x^i(k+1) - x^i(k)\right)$$

where $w_{ie} \geq 0$ sets the amount of energy needed to move a certain distance.

- We assume that energy is independent of resource type being sought and consumed.
- The "danger" aspect could be modeled as in the last section, but we ignore this possibility here.

➙ Our total payoff to forager i at time k is

$$J_i(x(k+1), x(k), u(k)) = J_{ic}(x(k+1), x(k), u(k)) + J_{ie}(x(k+1), x(k))$$

- If there are N_s steps in the game we have a payoff $J_i^{N_s}$ for playing the entire multistage game as given in Equation (115).
- Each forager wants to minimize $J_i^{N_s}$ and thereby maximize consumption with minimal energy expenditure.
- This can require considerable finesse as it may be a good strategy to give up payoffs at some points in time in order to

realize more benefits at some other later time.

Information Space and Strategy Design Challenges

- Consider one approach...
- Suppose that Equation (117) holds and each forager only uses $y^i(k)$ so
$$I^i(k) = Y^i(k)$$
and we need to choose G^i_k where $u^i(k) = G^i_k(x(k))$.
- Can see all forager positions and resources, but does not have memory
- Communication network? Type of game?
- Here, we will assume that the communication topology enables the sharing of sensed information for $I^i(k)$ above. This will allow each forager to compute all the decisions of all the other foragers so that they do not need to share information on $u(k)$.

Biomimcry for Foraging Strategies

Rules, Planning, Learning

- Simple rules (heuristics?)
- Planning: model predictive control approach with a dynamic programming solution—computational complexity problems!
- Learning, learning and planning?

A Generic Saltatory Strategy

- Model "saltatory search" (recall from early)
- Aim at getting a computationally tractable strategy
- Generic steps:
  1. Play a static matrix game to determine where each forager should go and what to consume (consider this a set of goals)
  2. If any forager achieves a goal position, then play another

static matrix game

3. Repeat

- How do the foragers move from one goal position to another?
  - What can they sense during movement? Should it set a path, then not deviate from it? Should it take into account how other foragers move? Do foragers consume while they move?
- Notice that there is an element of prediction in this approach
  - Could use abstractions (e.g., about profitability of regions)
  - Can add other “social” elements
- Coping with complexity is a key challenge:
  - Spatial abstractions (e.g., about profitability of regions)
  - Time abstractions

Intelligent Foraging

- Some additional psychology and biology foundations.

Planning, Attention and Learning for Foraging

- A surrogate model method simultaneously learns an approximation to the cost function and uses it to guide where to search the cost function:
  1. Pick a test point (or set of test points) for J and compute J at this point (these points). Note that the method can be "set-based" so that it computes in parallel the cost function at several test points (e.g., via parallel processing).
  2. Store the pairing(s) between the test point(s) and value(s) of J in a training data set G for an approximator f for J.
  3. Construct an approximator (interpolator) for the data in G (perhaps removing some points as others are added). This approximator retuning can be achieved via repeated

application of recursive least squares over a linear in the parameters approximator, or via application of a Levenberg-Marquardt method to training a nonlinear in the parameter appoximator.

4. Perform an optimization *over the approximator surface* (not the cost function) to find a minimum point on that surface (you may use gradient methods or pattern search methods to perform this optimization). Call this a new test point, compute J at this point (and for a set-based method perhaps at a pattern of points around it), and add this (these) to the training data set. Go back to step 3.

Intelligent Social Foraging

- Vehicles, Environment, and Objectives:
  1. Groups of Vehicles:
     - Vehicle dynamics, sensors/actuators, communications
     - Hierarchy and distribution in the group of vehicles
  2. Environment Model
     - Media
     - Predator/prey (or noxious substance/nutrient) characteristics
     - Environmental changes
  3. Goals
     - Energy consumption
     - Achieving goal positions
     - Gathering information
     - Changing the environment

Elements of Distributed Decision-Making

- Distributed Rule-Based Cooperative Foraging:
  - Using neighbor's information in rule antecedents
  - Rules for sending information to neighbors
- Distributed Planning for Cooperative Foraging:
  - Sharing models
  - Sharing plans or sets of plans
  - Sharing plan selection strategies
- Distributed Attention for Cooperative Foraging:
  - Distributed agreement on focus regions
  - Distributed dynamic attentional strategies
  - Leaders and hierarchical strategies

Distributed Learning

- Can use with all above distributed strategies
- Learning characteristics of the environment
- Learning foraging strategies from other group members
- Learn how to communicate
- Example: Distributed surrogate model method for intelligent social foraging.

Evolution of Foragers

- Evolving Cooperative Foraging Strategies:
  - Designing parameters of the decision-making elements
  - Achieving balance between decision-making functionalities
  - Evolving simple designs
  - Studying trade-offs between computational and communication resources
  - Co-evolution
  - Darwinian design of the software
- Evolving Vehicular Hardware?
  - Could you build an experiment that would illustrate hardware evolution in a foraging swarm, not just evolution of software?
  - How do you make the hardware replicate itself with

fecundity and variation, and instill inheritance into the process?

- How do you implement (un)natural selection via environmental influences?
- Could you make this emulate evolution of biological organisms, however simple they might be?
- Would this be useful for understanding biological evolution?
- Could you argue that your hardware is alive?
- What is the engineering utility of performing hardware evolution for populations? Could it be a way to make a group of vehicles more adaptable to changes in its environment?

➙ Via combined hardware-software evolution, learning, planning, attention, rule-based, and neural systems approaches, could you implement a truly “intelligent” controller?

References

[1] B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J.D. Watson. *Molecular Biology of the Cell*. Garland Publishing, NY, 2nd edition, 1989.

[2] T. Audesirk and G. Audesirk. *Biology: Life on Earth*. Prentice Hall, NJ, 5 edition, 1999.

[3] E.O. Budrene and H.C. Berg. Dynamics of formation of symmetrical patterns by chemotactic bacteria. *Nature*, 376:49–53, 1995.

[4] Neil A. Campbell. *Biology*. Benjamin Cummings Pub. Co., Menlo Park, CA, 1996.

[5] D. Chrisiansen. Pigeon-guided missiles that bombed out. *IEEE Spectrum*, page 46, August 1987.

[6] M. Domjan. *The Principles of Learning and Behavior*. Brooks/Cole Pub., NY, 4th edition, 1998.

[7] Michael S. Gazzaniga, Richard B. Ivry, and George R. Mangun. *Cognitive Neuroscience: The Biology of the Mind*. W. W. Norton and Co., NY, 1998.

[8] P. Gray. *Psychology*. Worth Pub., NY, 3 edition, 1999.

[9] M. Groening. *Simpsons Comics: Simpsorama*. HarperPerennial, Kansas City, 1995.

[10] E.R. Kandel, J.H. Schwartz, and T.M. Jessell, editors. *Essentials of neural science and behavior*. Appleton and Lange, Norwalk, CT, 1995.

[11] G. Larson. *The Far Side Gallery*. Andrews and McMeel, Kansas City, 1984.

[12] Irwin B. Levitan and Leonard K. Kaczmarek. *The Neuron: Cell and Molecular Biology*. Oxford University Press, Oxford, 2nd edition, 1997.

[13] R. Losick and D. Kaiser. Why and how bacteria communicate. *Scientific American*, 276(2):68–73, 199?

[14] M.T. Madigan, J.M. Martinko, and J. Parker. *Biology of Microorganisms*. Prentice Hall, NJ, 8 edition, 1997.

[15] K.M. Passino. *Biomimicry for Optimization, Control, and Automation*. ???, ???, ???

[16] J.A. Shapiro. Bacteria as multicellular organisms. *Scientific American*, 258:62–69, 1988.

[17] P. Whitfield. *From So Simple a Beginning*. Macmillan Pub. Co. NY, 1993.