Evolutionary stability of games with costly signaling

Gerhard Jäger
Gerhard.Jaeger@uni-bielefeld.de

December 17, 2007

Amsterdam Colloquium

Overview

- signaling games with costly signaling
- Evolutionary Game Theory
- conditions for evolutionary stability of signaling games

Signaling games

general setup

- two players, the sender and the receiver.
- sender has private information about an event that is unknown to the receiver
- event is chosen by nature according to a certain fixed probability distribution
- sender emits a signal which is revealed to the receiver
- receiver performs an action, and the choice of action may depend on the observed signal
- utilities of sender and receiver may depend on the event, the signal and the receiver's action

Signaling games

specific assumptions

- the utility of sender and receiver are identical,
- set of events $\mathcal{E}$, set of events $\mathcal{F}$, and set of actions $\mathcal{A}$ are finite,
- $\mathcal{E} = \mathcal{A}$ (the receiver's action is to guess an event)

Signaling games

costly signaling

- production/reception of signals may incur costs
- examples:
  - length, processing complexity etc. of natural language expressions
  - advertising costs in economics
  - “handicap” signaling in biology
  - ...
- can be represented as negative utility

Signaling games

- let e be the event to be communicated, σ the signal and a the receiver's action
- c_σ is the cost of using signal σ
- partnership game: S and H have identical utility function

utility function (extensive form)

$$u(e, \sigma, a) = \delta_{e,a} + c_\sigma \tag{1}$$

Signaling games

further constraints

- costs are normalized such that $\max_i c_i = 0$
- all events have positive probability
- no event has costs ≤ -1—otherwise use of that signal would never be rationalizable

structural stability

- no two events have identical probability
- no two signals have identical costs
- all signals have costs strictly > -1

Signaling games

matrix representation

- let $n = |\mathcal{E}|$ be number of events
- $m = |\mathcal{F}|$ is number of signals
- (pure) strategies can be represented as matrices with one 1 per row and else 0
- sender strategy S: $n \times m$-matrix
- receiver strategy R: $m \times n$-matrix
- $\vec{e}$: nature's probability distribution over events
- $\vec{c}$: costs of signals $1, \ldots, m$

Signaling games

An example

- $\vec{e} = \langle .75, .25 \rangle$
- $\vec{c} = \langle 0, -.1 \rangle$
- Horn strategy:

$$S = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad R = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

compiling $\vec{e}$ and $\vec{c}$ into the matrix representation

$$P = \begin{pmatrix} .75 & 0 \\ 0 & .25 \end{pmatrix} \quad Q = \begin{pmatrix} 1 & 0 \\ -.1 & .9 \end{pmatrix}$$

$$u(S, R) = \mathrm{tr}(PQ) = .975$$

Signaling games

An example

- $\vec{e} = \langle .75, .25 \rangle$
- $\vec{c} = \langle 0, -.1 \rangle$
- Anti-Horn strategy:

$$S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad R = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$P = \begin{pmatrix} 0 & .75 \\ .25 & 0 \end{pmatrix} \quad Q = \begin{pmatrix} 0 & 1 \\ .9 & -.1 \end{pmatrix}$$

$$u(S, R) = \mathrm{tr}(PQ) = .925$$

Evolution of communication

Universal Darwinism

- Darwinian Evolution:
  - population of replicating individuals
  - heritable traits
  - differential replication
  - leads to "natural" selection → survival of the fittest
- not confined to biology
- imitation is form of selection in cultural sphere

Evolution of communication

Evolutionary Game Theory

- utility is interpreted as *fitness* (expected replicative success)
- evolutionary game theory models *frequency-dependent selection*
- "frequency dependent": fitness of some traits depends on quantitative composition of surrounding population
- focus is not on details of the evolutionary dynamics, but on long-term behavior
- basic idea: natural selection is deterministic, but actual systems are subject to (infinitesimal) noise
- Big question: What is the long-run behavior of such a system?

Evolution of Communication

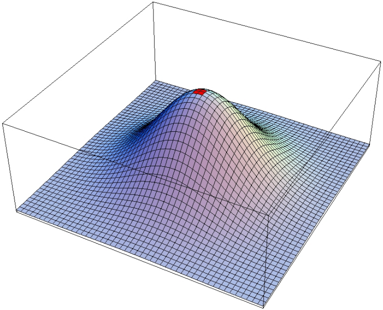
Evolution of partnership games

- each trajectory converges to some rest point ($\approx$ Nash equilibrium)
- average fitness is a strict Lyapunov function $\rightarrow$ every change comes with an increase in average fitness
- rest points are flat points in the fitness landscape

Types of rest points

local fitness maxima

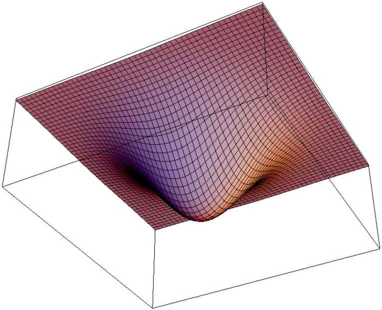
- basin of attraction has positive measure
- evolutionarily stable state: after a small random shock, the system will be pushed back into equilibrium



Types of rest points

local fitness minima

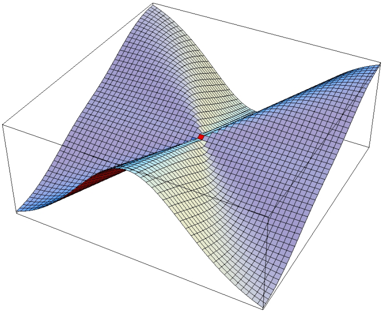
- basin of attraction has measure 0
- instable: random noise will push the system out of equilibrium



Types of rest points

saddle points

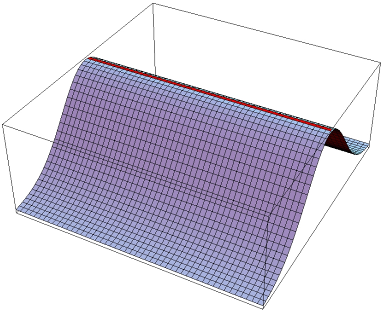
- basin of attraction has measure 0
- instable: random noise will push the system out of equilibrium



Types of rest points

ridge

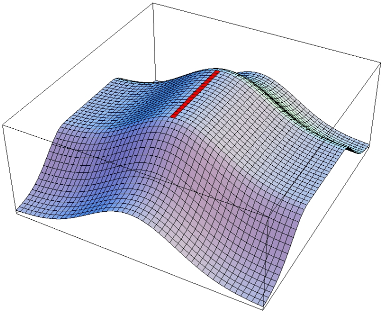
- basin of attraction has positive measure
- neutrally stable: small deviations remain local for some time



Types of rest points

extended peak

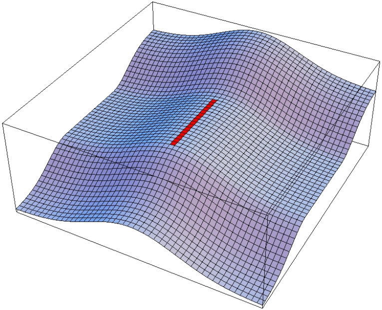
- basin of attraction has positive measure
- evolutionarily stable set: system cannot leave the extended peak once it is attained



Types of rest points

extended saddle

- basin of attraction has positive measure
- instable in the long run: random noise will eventually push the system out of the basin of attraction



Evolutionary stability of signaling games

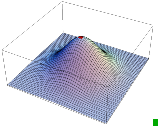
Theorem

(S, R) is an ESS if and only if

1. $m \leq n$,
2. *the first column of S has $n - m + 1$ positive entries,*
3. *each other column of S has exactly one positive entry, and*
4. $r_{ji} = 1$ *iff* $i = \min(\{i' : s_{i'j} > 0\})$, *otherwise* $r^x_{ji} = 0$.

Corollary

If $n = m$, the ESSs are exactly the states where S and R are bijective and the inverse of each other.

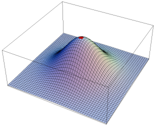


Evolutionary stability of signaling games

an ESS with $m < n$

$$S = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \quad R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$P = \begin{pmatrix} .5 & 0 \\ .3 & 0 \\ 0 & .2 \end{pmatrix} \quad Q = \begin{pmatrix} 1 & 0 & 0 \\ -.1 & -.1 & .9 \end{pmatrix}$$

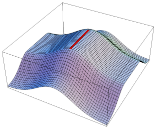


Evolutionary stability of signaling games

Theorem

A set of strategy pairs A is an ESSet (possibly extended peak) iff for each $(S, R) \in A$, (S, R) is an ESS or

1. $m > n$,
2. *the restriction of S to the first n columns and the restriction of R to the first n rows form an ESS, and*
3. *for each R' such that R and R' agree on the first n rows: $(S, R') \in A$.*

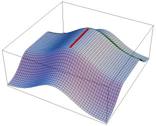


Evolutionary stability of signaling games

a non-singleton ESSet ("extended peak")

$$\left\{ x : \; S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad R = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \alpha & 1-\alpha \end{pmatrix} \; \& \; \alpha \in [0,1] \right\}$$

$$\left\{ x : \; P = \begin{pmatrix} .8 & 0 & 0 \\ 0 & .2 & 0 \end{pmatrix}, \quad Q = \begin{pmatrix} 1 & 0 \\ -.1 & .9 \\ \alpha - .2 & .8-\alpha \end{pmatrix} \; \& \; \alpha \in [0,1] \right\}$$



Evolutionary stability of signaling games

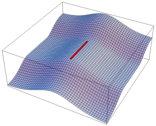
Theorem

Every game with $n, m \geq 2$ has an extended saddle.

In Horn games, Nash-Smolensky strategies form extended saddle:

$$S = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \quad R = \begin{pmatrix} 1 & 0 \\ \alpha & 1-\alpha \end{pmatrix}$$

$$P = \begin{pmatrix} .75 & 0 \\ .25 & 0 \end{pmatrix} \quad Q = \begin{pmatrix} 1 & 0 \\ \alpha - .1 & .9 - \alpha \end{pmatrix}$$



for $\alpha \in (.9, 1]$.

Evolutionary stability of signaling games

Summary

- natural selection + noise:
  - always leads to optimal communication (the maximally possible number of events is reliably communicated)
  - does not necessarily lead to optimal strategy (where average costs are minimized)
- natural selection without noise:
  - may lead (with positive probability) to a sub-optimal state where some events cannot be communicated

$\Rightarrow$ Evolution does lead to optimal communication, but it may take a very long time to reach that state.