Noisy Signaling Games

Jelle Zuidema (ILLC)

Outline

1. Evolutionary theory, Evolutionary Games

2. Communication with noise

3. Evolutionary Stable Strategies

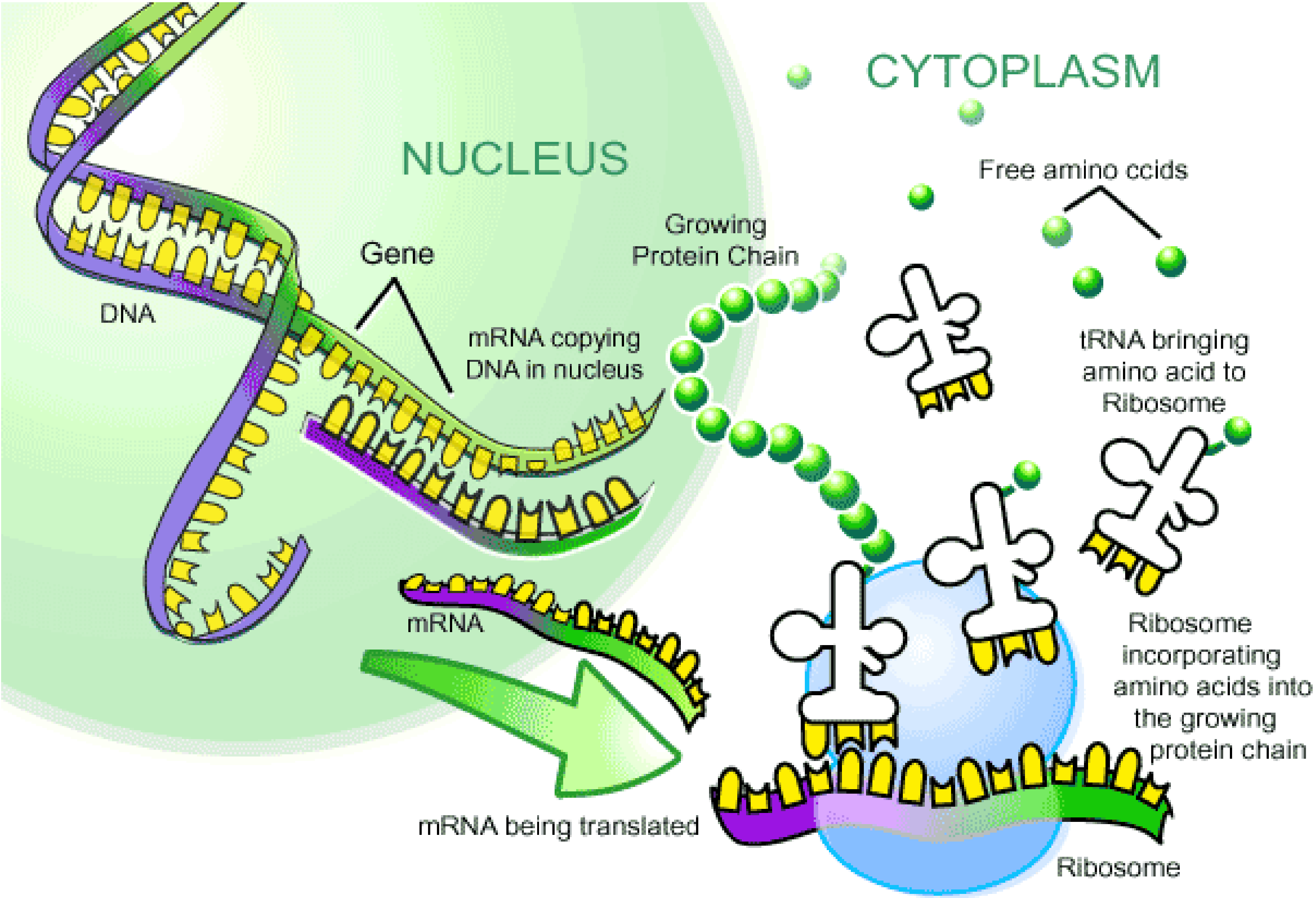
4. Numerical Simulations

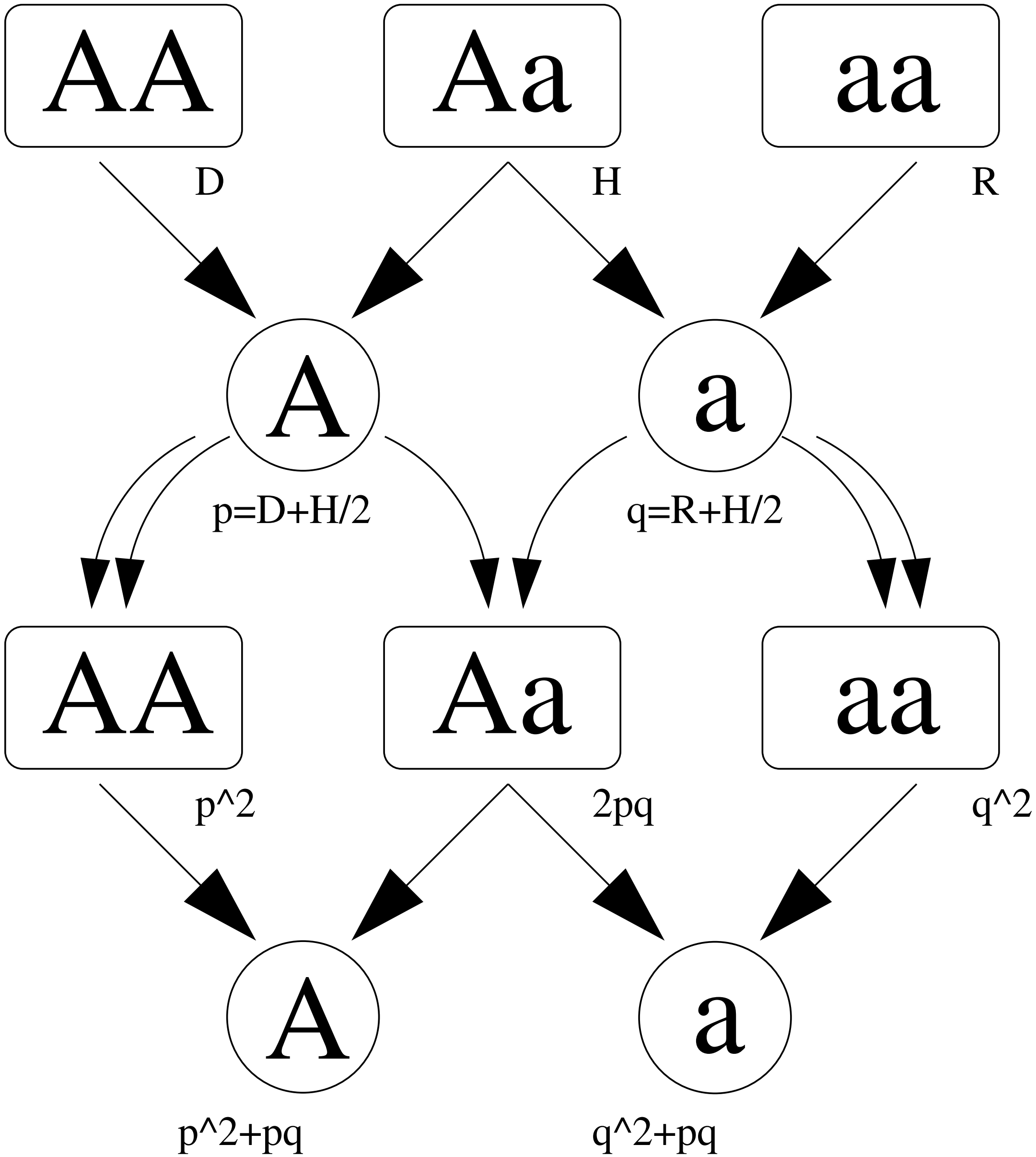
Evolution by Means of Natural Selection (Darwin, 1859)

- Universal Common descent, Descent with modification
- Fitness: survival, reproduction
- Natural Selection, Sexual Selection
- Mutation, Heritable variation
- Genetic drift

Criticisms

- Young earth, Genesis, Blasphemy, Human uniqueness...
- Inheritance of acquired characteristics (Lamarck)
- Blending inheritance
- Origin of Species


NUCLEUS
CYTOPLASM
Free amino ccids
Growing
Protein Chain
Gene
DNA
mRNA copying
DNA in nucleus
tRNA bringing
amino acid to
Ribosome
mRNA
Ribosome
incorporating
amino acids into
the growing
protein chain
mRNA being translated
Ribosome

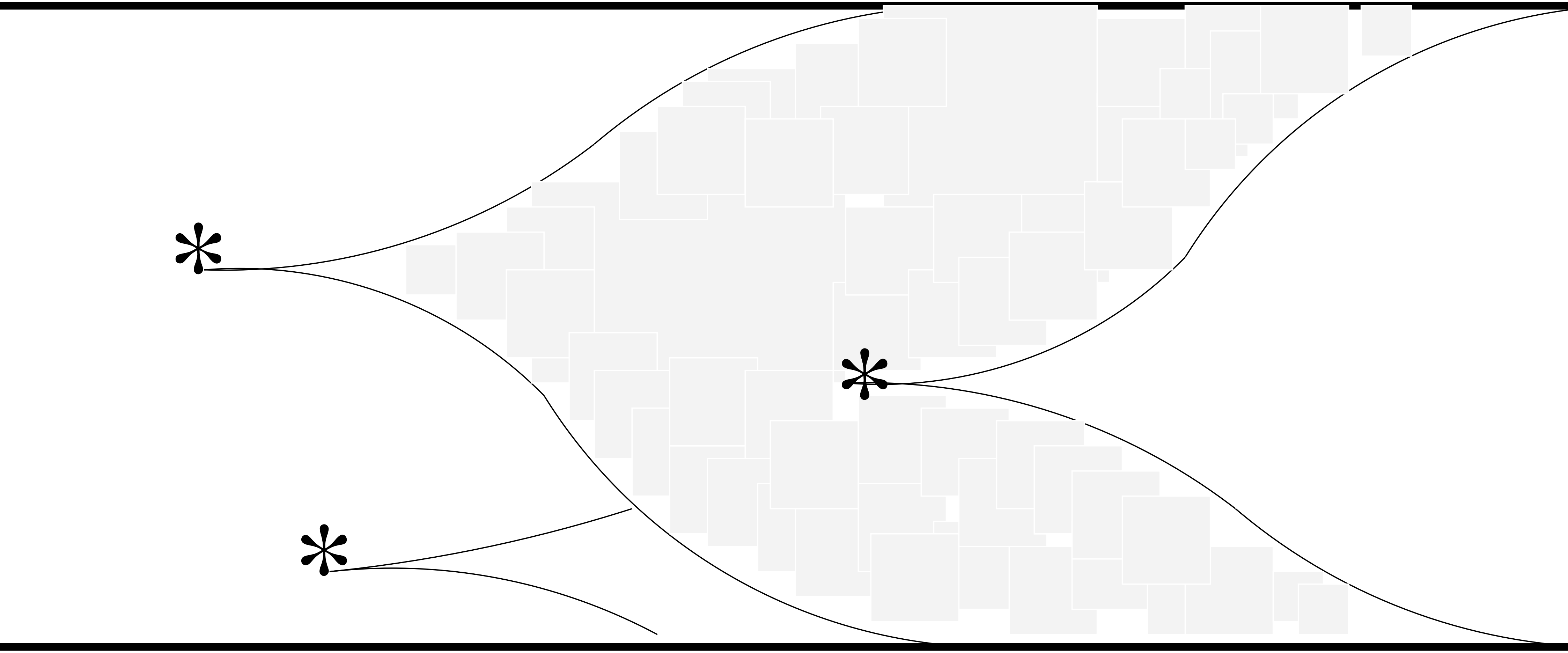

AA
Aa
aa
D
H
R
A
a
p=D+H/2
q=R+H/2
AA
Aa
aa
p^2
2pq
q^2
A
a
p^2+pq
q^2+pq

Evolution as Gene-Frequency Change

$$p' = \frac{p^2 w_{AA} + pq w_{Aa}}{\overline{w}}, \tag{1}$$

where $\overline{w}$ is the average fitness and given by:

$$\overline{w} = p^2 w_{AA} + 2pq w_{Aa} + q^2 w_{aa} \tag{2}$$



Evolution as Adaptation (Wright 1931, Fisher 1930)

$$\begin{aligned}\Delta p &= p' - p \\ &= \frac{p^2 w_{AA} + pq w_{Aa}}{\overline{w}} - p \end{aligned} \tag{3}$$

This equation can, with a bit of algebra, be rewritten as follows:

$$\Delta p = \frac{pq}{\overline{w}} \left(p \left(w_{AA} - w_{Aa} \right) - q \left(w_{aa} - w_{Aa} \right) \right) \tag{4}$$

The change of fitness with every change of p is:

$$\frac{\mathrm{d}\overline{w}}{\mathrm{d}p} = 2 \left(p \left(w_{AA} - w_{Aa} \right) - q \left(w_{aa} - w_{Aa} \right) \right) \tag{5}$$

And thus:

$$\Delta p = \frac{pq}{\overline{w}} \left(\frac{1}{2} \right) \frac{\mathrm{d}\overline{w}}{\mathrm{d}p}. \tag{6}$$

Details for the skeptics...

$$\overline{w} = p^2 w_{AA} + 2pq w_{Aa} + q^2 w_{aa} \tag{7}$$

Because $p + q = 1$ this expression can be rewritten as:

$$\begin{aligned} \overline{w} &= p^2 w_{AA} + 2p(1-p) w_{Aa} + (1-p)^2 w_{aa} \\ &= p^2 w_{AA} + 2p w_{Aa} - 2p^2 w_{Aa} + w_{aa} - 2p w_{aa} + p^2 w_{aa}. \end{aligned} \tag{8}$$

The derivative of $\overline{w}$ with respect to p is now (provided the fitness coefficients are independent of p):

$$\begin{aligned} \frac{\mathrm{d}\overline{w}}{\mathrm{d}p} &= 2p w_{AA} + 2 w_{Aa} - 4p w_{Aa} - 2 w_{aa} + 2p w_{aa} \\ &= 2\left(p w_{AA} + w_{Aa} - 2p w_{Aa} - w_{aa} + p w_{aa}\right) \\ &= 2\left(p w_{AA} + (1-p) w_{Aa} - p w_{Aa} - (1-p) w_{aa}\right) \\ &= 2\left(p w_{AA} + q w_{Aa} - p w_{Aa} - q w_{aa}\right) \\ &= 2\left(p\left(w_{AA} - w_{Aa}\right) - q\left(w_{aa} - w_{Aa}\right)\right). \end{aligned} \tag{9}$$

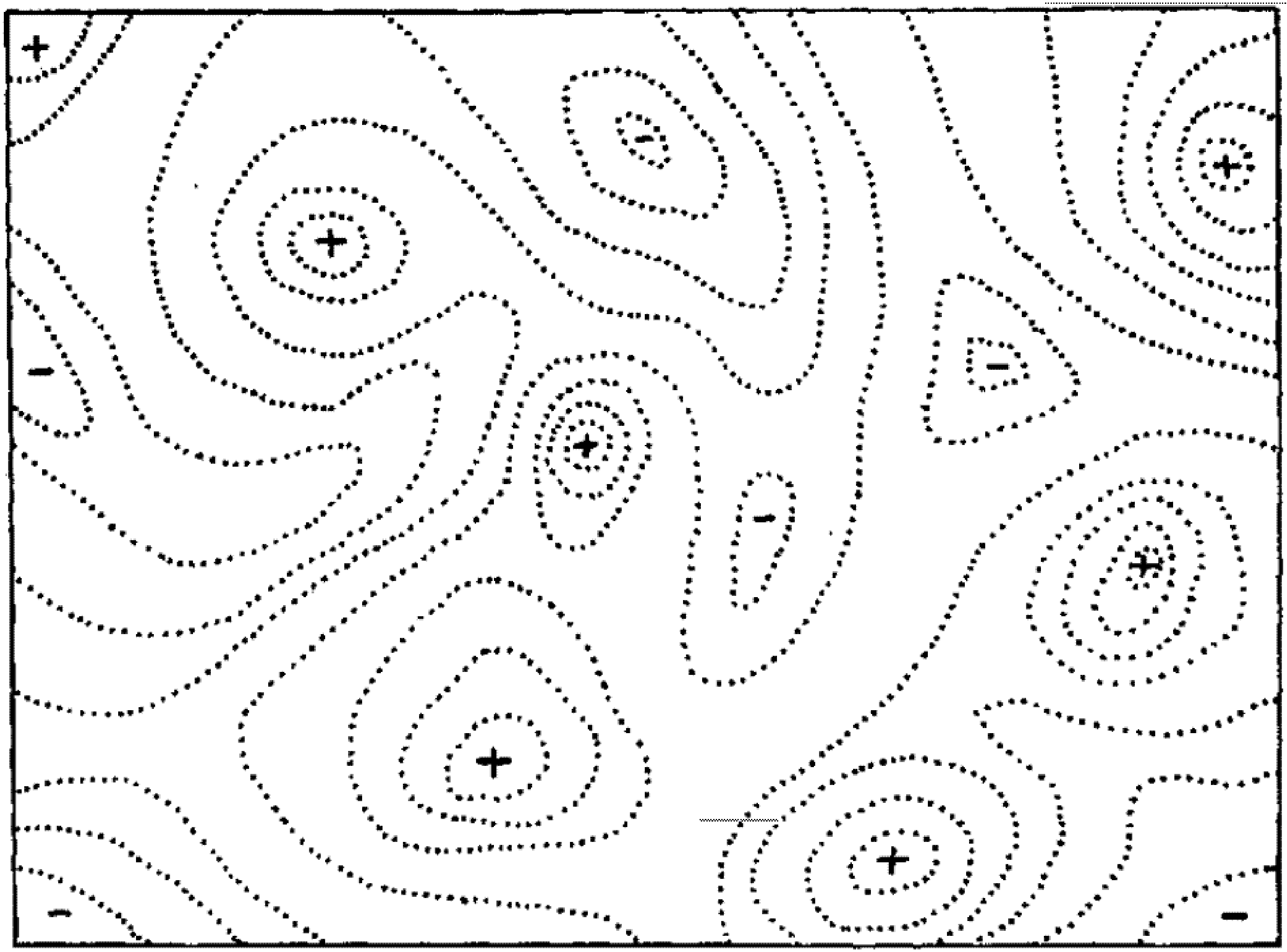
Now, recall the expression for the change in p (equation (3)), which can in a few steps be rewritten as:

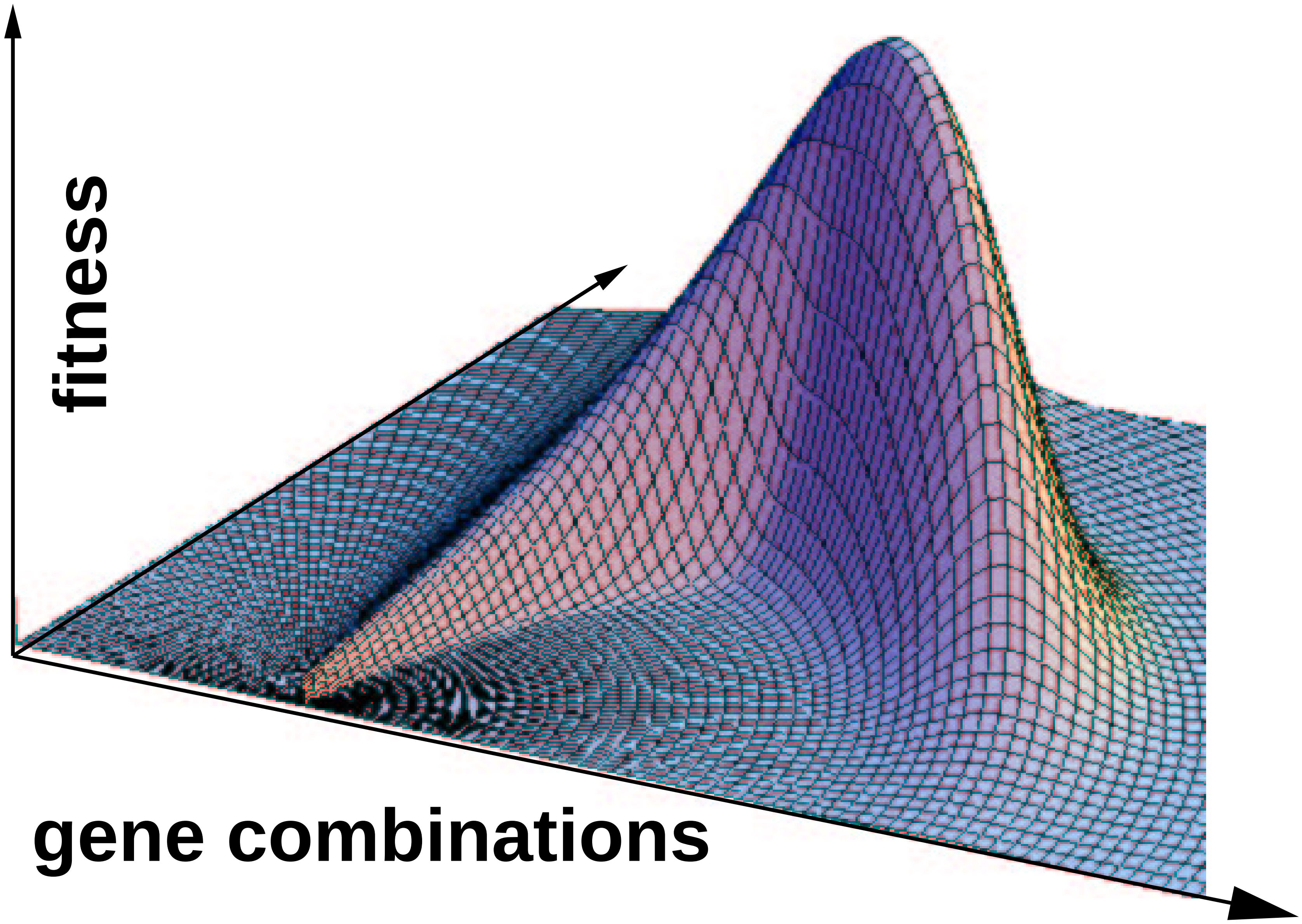
$$\begin{aligned}
\Delta p &= p' - p \\
&= \frac{p(pw_{AA} + qw_{Aa})}{\overline{w}} - p \\
&= \frac{p(pw_{AA} + qw_{Aa})}{\overline{w}} - \frac{p\overline{w}}{\overline{w}} \\
&= \frac{p}{\overline{w}}(pw_{AA} + qw_{Aa} - \overline{w}). \qquad (10)
\end{aligned}$$

Inserting equation (7) into equation (10), and rearranging using the fact that $q = 1 - p$, gives:

$$\begin{aligned}
\Delta p &= \frac{p}{\overline{w}}(pw_{AA} + qw_{Aa} - p^2 w_{AA} - 2pqw_{Aa} - q^2 w_{aa}) \\
&= \frac{p}{\overline{w}}(p(w_{AA} - pw_{AA}) + qw_{Aa} - 2pqw_{Aa} - q^2 w_{aa}) \\
&= \frac{p}{\overline{w}}(pqw_{AA} + qw_{Aa} - 2pqw_{Aa} - q^2 w_{aa}) \\
&= \frac{pq}{\overline{w}}(pw_{AA} + w_{Aa} - 2pw_{Aa} - qw_{aa}) \\
&= \frac{pq}{\overline{w}}(p(w_{AA} - w_{Aa}) - q(w_{aa} - w_{Aa})). \qquad (11)
\end{aligned}$$

The Adaptive Landscape (Wright, 1932)




fitness
gene combinations

Limits to Optimality

"Natural selection tends only to make each organic being as perfect as, or slightly more perfect than, the other inhabitants of the same country with which it comes into competition. And we see that this is the standard of perfection attained under nature" (Darwin, 1872, p 163)

- biophysical and genetic constraints
- the speed of evolution
- mutational load
- fluctuating fitness
- frequency-dependent fitness
- correlation, levels of selection

Evolutionary Game Theory (Maynard Smith & Price, 1973)

An Evolutionary Stable Strategy (ESS) is a strategy that cannot be invaded by any other strategy, because all other strategies get either a lower payoff when playing against the ESS, or if their payoff is equal, they get a lower payoff when playing against themselves.

That is, if $W(i, j)$ gives the payoff for a player playing strategy i against an opponent playing strategy j, then i is an ESS iff:

$$\forall j \,(W(i,i) > W(j,i) \vee W(i,i) = W(j,i) > W(j,j))$$

Problem of cooperation Why would senders be willing to send honest signals, and hearers be willing to receive and believe the signal?

Honest signaling theory (Zahavi, Maynard Smith, Grafen, Bergstrom)

Problem of coordination How is, after each innovation, a shared code established and maintained? And which code?

Coordination games (Lewis, Skyrms, Nowak, this talk)

A formalism for communication under noisy conditions

- Assume that there are M different meanings that an individual might want to express, and F different signals (forms) that it can use for this task.

- The communication system of an individual is represented with a production matrix S (S gives for every meaning m and every signal f, the probability that the individual chooses f to convey m);

- and an interpretation matrix R. (R gives for every signal f and meaning m, the probability that f will be interpreted as m).

- Signals can be more or less similar to each other and there is noise on the transmission of signals which depends on these similarities (confusion matrix U).

- Meanings can be more or less similar to each other, and the value of a certain interpretation depends on how close it is to the intention (value matrix V)

Example: alarm calls (vervets, squirls, ...)

Three different types of predators: from the air (eagles), from the ground (leopards) and from the trees (snakes).

The monkeys are capable of making a number (say 5) of different sounds that range on one axis (e.g. pitch, from high to low) and are more easily confused if they are closer together.

If one makes a mistake, typically not every mistake is equally bad.

$S =$

intention ↓	sent signal 1*kHz*	2*kHz*	3*kHz*	4*kHz*	5*kHz*
eagle	1.0	0.0	0.0	0.0	0.0
snake	0.0	0.0	1.0	0.0	0.0
leopard	0.0	0.0	0.0	0.0	1.0

$U =$

sent signal ↓	received signal 1*kHz*	2*kHz*	3*kHz*	4*kHz*	5*kHz*
1*kHz*	0.7	0.2	0.1	0.0	0.0
2*kHz*	0.2	0.6	0.2	0.0	0.0
3*kHz*	0.0	0.2	0.6	0.2	0.0
4*kHz*	0.0	0.0	0.2	0.6	0.2
5*kHz*	0.0	0.0	0.1	0.2	0.7

$R =$

received signal ↓	interpretation		
	eagle	*snake*	*leopard*
1*kHz*	1.0	0.0	0.0
2*kHz*	1.0	0.0	0.0
3*kHz*	0.0	1.0	0.0
4*kHz*	0.0	0.0	1.0
5*kHz*	0.0	0.0	1.0

$V =$

interpretations ↓	intentions		
	eagle	*snake*	*leopard*
eagle	9	5	1
snake	2	9	2
leopard	1	5	9

$S =$

intention ↓	sent signal 1*kHz*	2*kHz*	3*kHz*	4*kHz*	5*kHz*
eagle	1.0	0.0	0.0	0.0	0.0
snake	0.0	0.0	1.0	0.0	0.0
leopard	0.0	0.0	0.0	0.0	1.0

$U =$

sent signal ↓	received signal 1*kHz*	2*kHz*	3*kHz*	4*kHz*	5*kHz*
1*kHz*	0.7	0.2	0.1	0.0	0.0
2*kHz*	0.2	0.6	0.2	0.0	0.0
3*kHz*	0.0	0.2	0.6	0.2	0.0
4*kHz*	0.0	0.0	0.2	0.6	0.2
5*kHz*	0.0	0.0	0.1	0.2	0.7

$R =$

received signal ↓	interpretation *eagle*	*snake*	*leopard*
1*kHz*	1.0	0.0	0.0
2*kHz*	1.0	0.0	0.0
3*kHz*	0.0	1.0	0.0
4*kHz*	0.0	0.0	1.0
5*kHz*	0.0	0.0	1.0

$V =$

interpretations ↓	intentions *eagle*	*snake*	*leopard*
eagle	9	5	1
snake	2	9	2
leopard	1	5	9

Payoff (Utility)

The payoff with a given S and R is:

$$W(S,R) = \sum_m \sum_f \sum_{f'} \sum_{m'} S_{mf} U_{ff'} R_{f'm} V_{mm'}$$

A language L is

$$L = \{S, R\}$$

In an symmetric, cooperative game, the payoff for player 1 with L communicating with player 2 with L' is:

$$w(L, L') = \frac{1}{2}\left(W(S, R') + W(S', R)\right)$$

Categorical meanings and signals: $U = I, V = I$

$$I = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

Categorical meanings and signals: $U = I, V = I$

$M < F$

$$S = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \qquad R = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0.3 & 0.7 \\ 1 & 0 \end{pmatrix}$$

- maximum payoff: F
- many Nash equilibria
- no ESSs

Categorical meanings and signals: $U = I, V = I$

$M = F$

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

- maximum payoff: $M = F$
- many Nash; e.g. all values in S and R at $\frac{1}{3}$
- ESSs: S is row-permutation of I

Categorical meanings and signals: $U = I, V = I$

$M > F$

$$S = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0.3 & 0.7 \\ 1 & 0 \end{pmatrix} \qquad R = \begin{pmatrix} 0.6 & 0 & 0 & 0.4 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

- maximum payoff: M
- many Nash equilibria
- no ESSs

Noisy signaling, uniform payoffs: $V = I$

$M < F$

$$S = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix} \qquad U = \begin{pmatrix} 0.7 & 0.2 & 0.1 & 0 & 0 \\ 0.15 & 0.7 & 0.1 & 0.05 & 0 \\ 0.05 & 0.1 & 0.7 & 0.1 & .05 \\ 0 & 0.05 & 0.1 & 0.7 & 0.15 \\ 0 & 0 & 0.1 & 0.2 & 0.7 \end{pmatrix} \qquad R = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

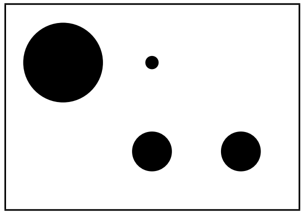
- maximum payoff: channel capacity (Shannon, 1948)
- few Nash equilibria
- At least $M!$ ESSs

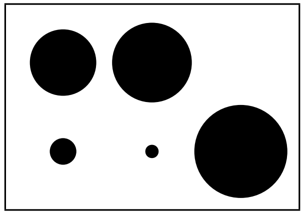
Visualising S and R

$$S = \left(\begin{array}{c|ccc} & f_1 & f_2 & f_3 \\ \hline m_1 & 0.9 & 0.1 & 0.0 \\ m_2 & 0.0 & 0.5 & 0.5 \end{array} \right)$$

$$R = \left(\begin{array}{c|cc} & m_1 & m_2 \\ \hline f_1 & 0.7 & 0.3 \\ f_2 & 0.9 & 0.1 \\ f_3 & 0.0 & 1.0 \end{array} \right)$$

$$R^T = \left(\begin{array}{c|ccc} & f_1 & f_2 & f_3 \\ \hline m_1 & 0.7 & 0.9 & 0.0 \\ m_2 & 0.3 & 0.1 & 1.0 \end{array} \right)$$

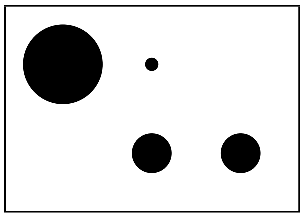


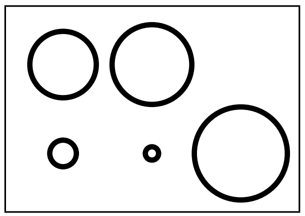


Visualising S and R

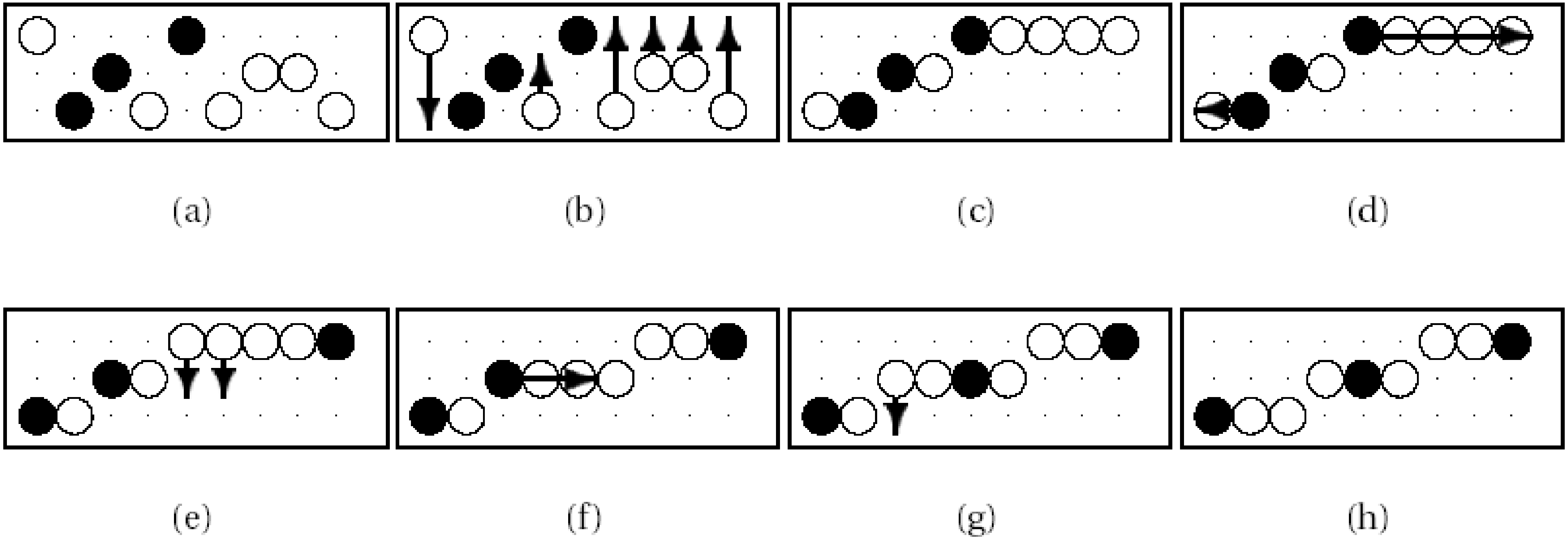
$$S = \left(\begin{array}{c|ccc} & f_1 & f_2 & f_3 \\ \hline m_1 & 0.9 & 0.1 & 0.0 \\ m_2 & 0.0 & 0.5 & 0.5 \end{array}\right) \qquad R = \left(\begin{array}{c|cc} & m_1 & m_2 \\ \hline f_1 & 0.7 & 0.3 \\ f_2 & 0.9 & 0.1 \\ f_3 & 0.0 & 1.0 \end{array}\right)$$

$$R^T = \left(\begin{array}{c|ccc} & f_1 & f_2 & f_3 \\ \hline m_1 & 0.7 & 0.9 & 0.0 \\ m_2 & 0.3 & 0.1 & 1.0 \end{array}\right)$$





Only distinctive lexicons are ESSs


(a) (b) (c) (d)
(e) (f) (g) (h)

Noisy signaling, heterogeneous payoffs

Numerical simulations

- The values in the S and R matrices are all either 1 or 0
- Distributed hill-climbing:
  1. Random speaker (i) and hearer (j) are picked, and $W(S_i, R_j)$ is measured;
  2. A random change is made in a random matrix of the speaker (or hearer), and $W(S_i, R_j)$ is measured again;
  3. If the new $W(S_i, R_j)$ is better, the change is kept; otherwise, it is reverted.

Motivation

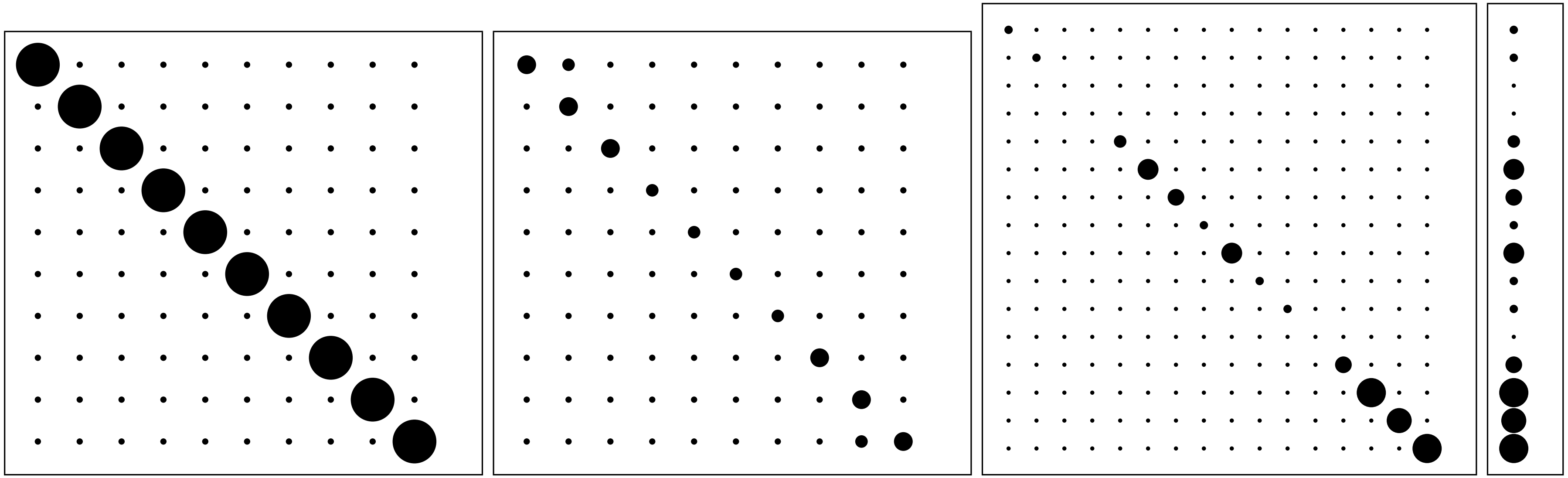
for this style of optimization:

1. it is fast and straightforward to implement;

2. it works well, and gives, if not the optimum, a good insight on characteristics of the optimal communication system;

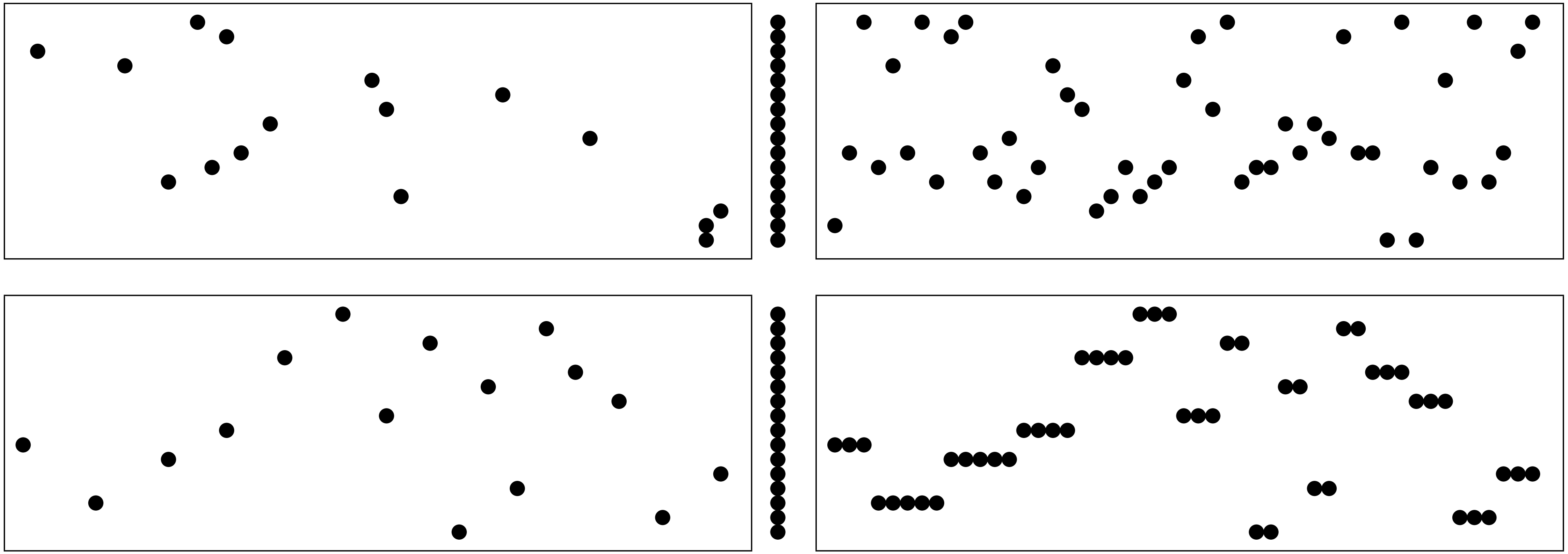
3. it shows possible routes to (near-) optimal communication systems, and in a sense forms an abstraction for both learning and evolution.

Control parameters: V and U-matrices



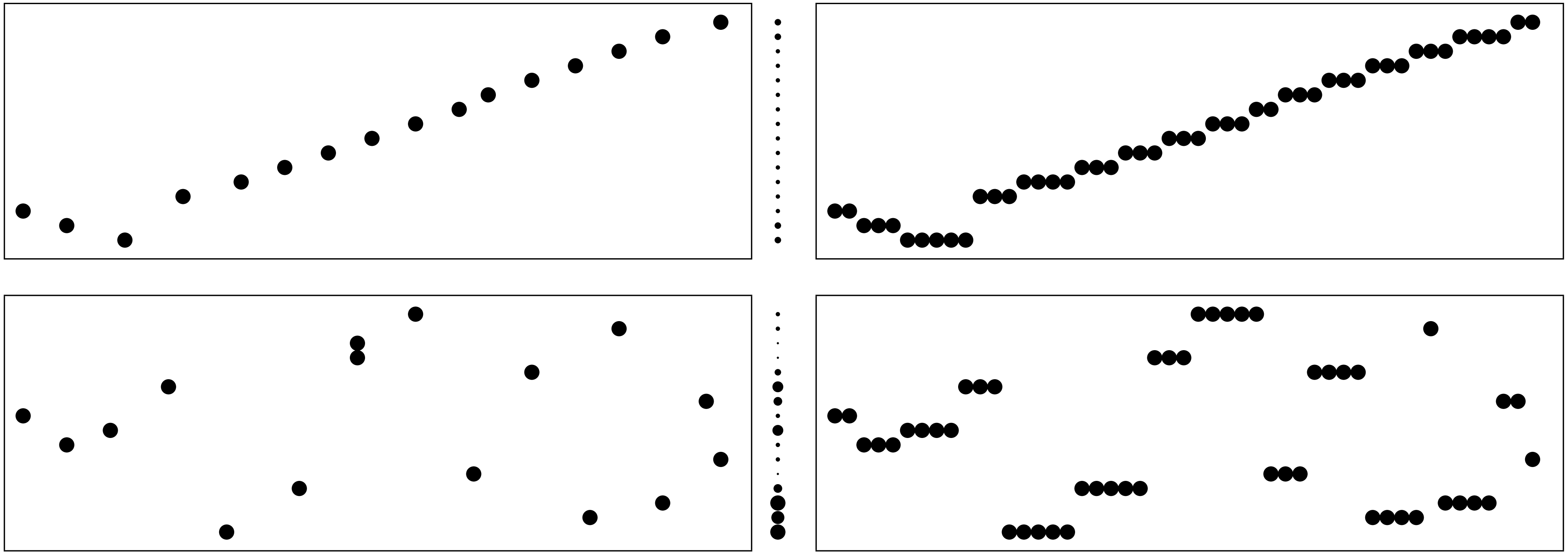
V:0d, 1d homogeneous, 0d heterogeneous

Specificity, Coherence, Distinctiveness



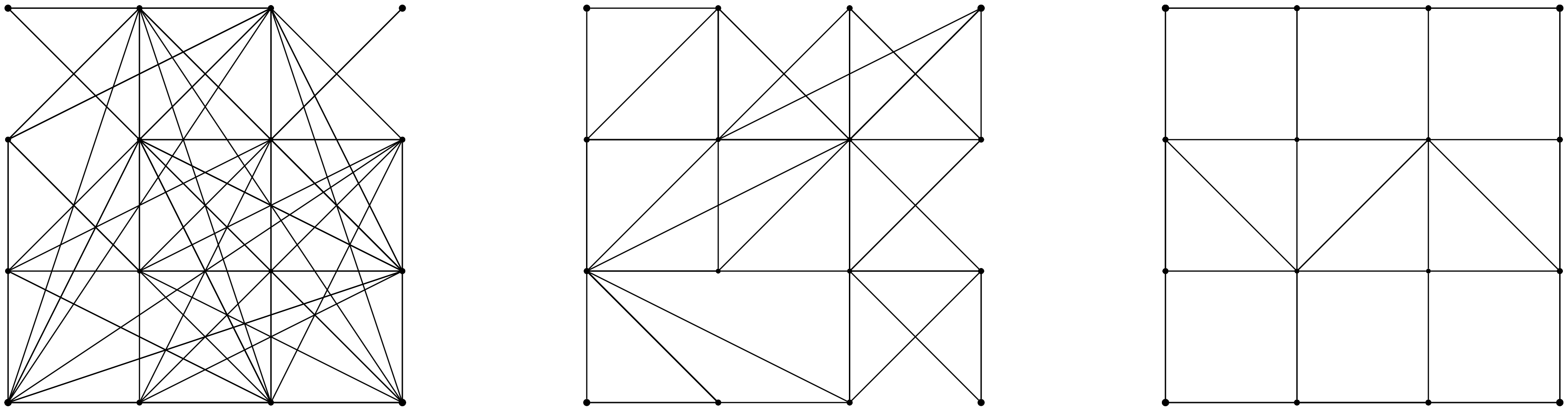
U:1d, V:0d homogeneous, $t = 0$, $t = \infty = 2 \times 10^8$

Topology preservation, Heterogeneity



U:1d, V:0d homogeneous/heterogeneous, $t = \infty$

Results: 2d meaning spaces

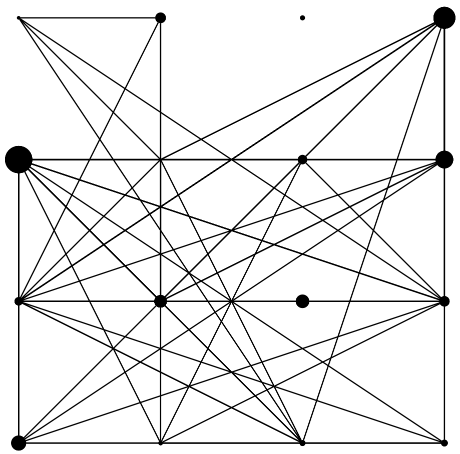


(a) t=0 (b) t=10^6 (c) t=∞

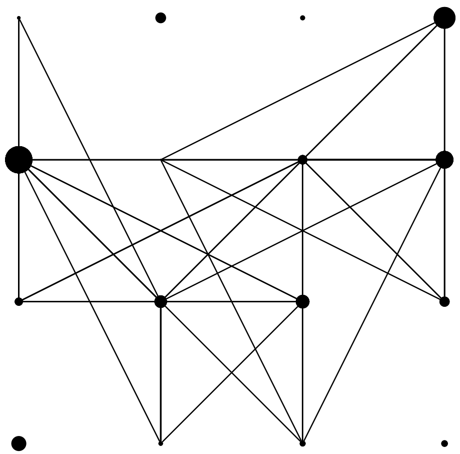
U:2d, V:2d homogeneous, $t = 0, 10^6, \infty$

(Points in meaning space are connected, if their preferred forms are neighbours in form space)

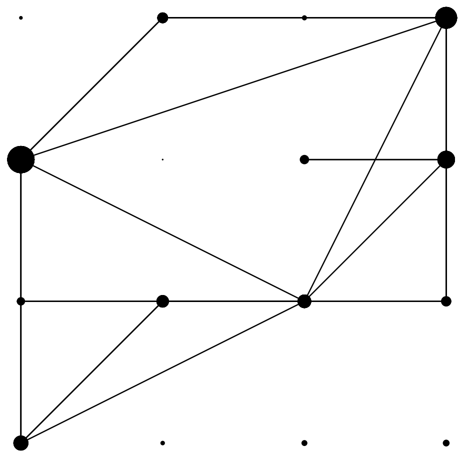
Heterogeneity: sacrificing low-valued meanings



(a) t=0



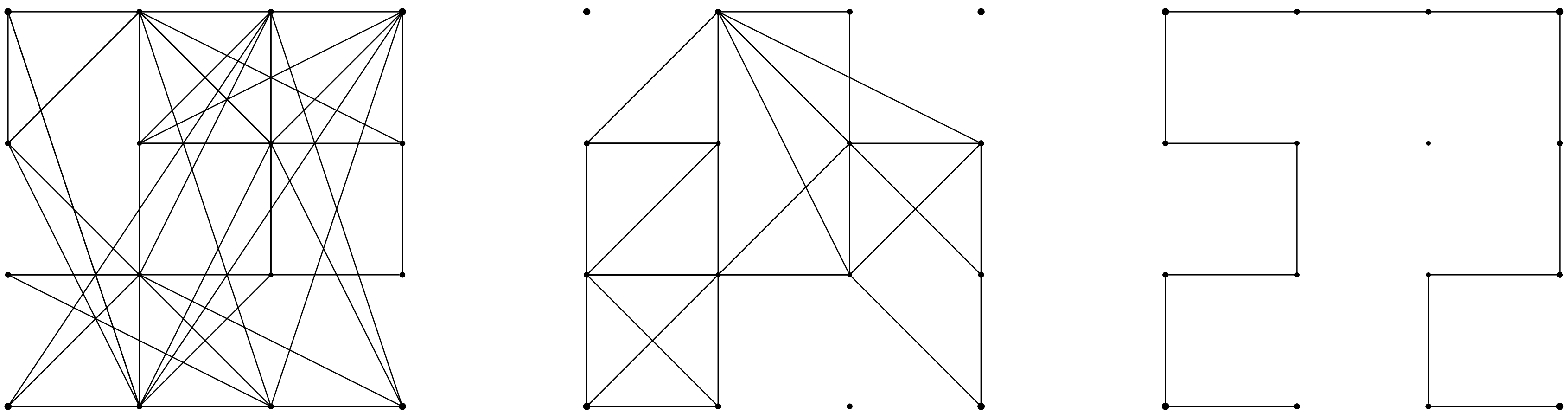
(b) $t=10^6$



(c) $t=\infty$

U:2d, V:2d heterogeneous, $t = 0, 10^6, \infty$

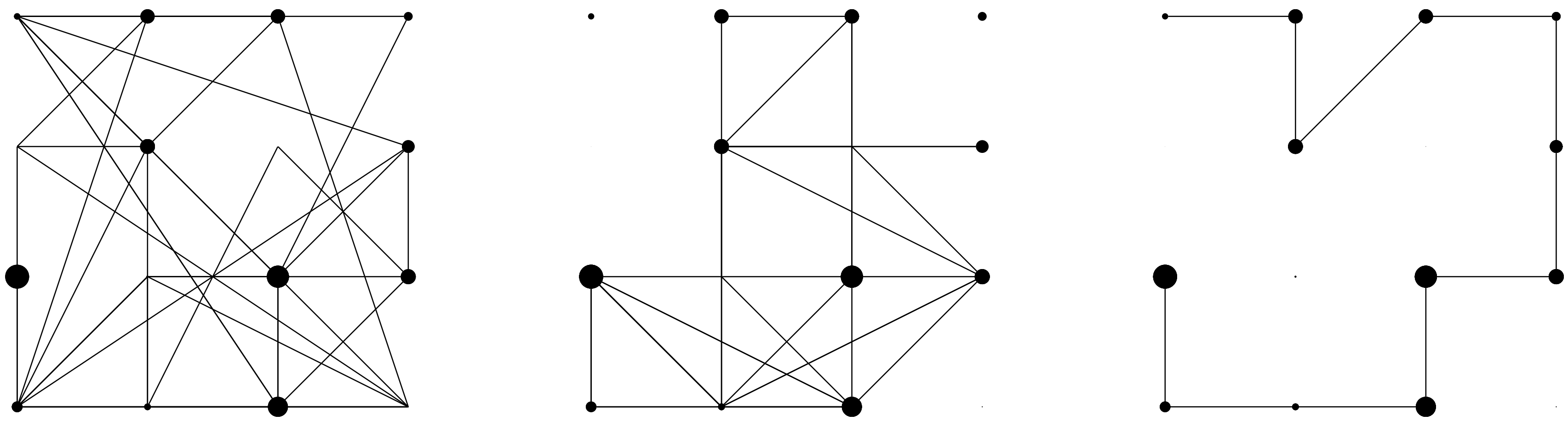
Dimensionality mismatch



(a) t=0 (b) t=10^6 (c) t=∞

U:1d, V:2d homogeneous, $t = 0, 10^6, \infty$

Dimensionality mismatch & Heterogeneity



(a) t=0 (b) $t=10^6$ (c) $t=\infty$

U:1d, V:2d heterogeneous, $t = 0, 10^6, \infty$

Summary of results

- Specificity
- Coherence (Lewis, 1969; Steels, 1996; Oliphant, 1996)
- Distinctiveness (De Boer, 1999; Nowak & Krakakauer, 1999)
- Topology preservation (Zuidema & Westermann, 2003)
- High-value meanings first
- Low-value meanings sacrificed