Evolutionary games and language

Gerhard Jäger
Gerhard.Jaeger@uni-bielefeld.de

August 19, 2005

ESSLLI 2005

Cognitive semantics

Gärdenfors (2000):

- meanings are arranged in **conceptual spaces**
- conceptual space has geometrical structure
- dimensions are founded in perception/cognition

Cognitive semantics

Gärdenfors (2000):

- meanings are arranged in **conceptual spaces**
- conceptual space has geometrical structure
- dimensions are founded in perception/cognition

Convexity

A subset C of a conceptual space is said to be *convex* if, for all points x and y in C, all points between x and y are also in C.

Cognitive semantics

Gärdenfors (2000):

- meanings are arranged in **conceptual spaces**
- conceptual space has geometrical structure
- dimensions are founded in perception/cognition

Convexity

A subset C of a conceptual space is said to be *convex* if, for all points x and y in C, all points between x and y are also in C.

Criterion P

A *natural property* is a convex region of a domain in a conceptual space.

Examples

- spatial dimensions: *above, below, in front of, behind, left, right, over, under, between ...*
- temporal dimension: *early, late, now, in 2005, after, ...*
- sensual dimenstions: *loud, faint, salty, light, dark, ...*
- abstract dimensions: *cheap, expensive, important, ...*

The naming game

- two players:
  - **S**peaker
  - **H**earer
- infinite set of **M**eanings, arranged in a finite metrical space
  distance is measured by function $d : M^2 \mapsto R$
- finite set of **F**orms
- sequential game:
  1. nature picks out $m \in M$ according to some probability distribution p and reveals m to S
  2. S maps m to a form f and reveals f to H
  3. H maps f to a meaning m'

The naming game

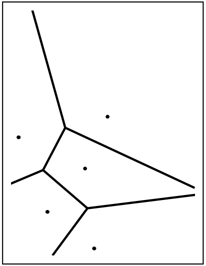
- **Goal:**
  - optimal communication
  - both want to minimize the distance between m and m'
- **Strategies:**
  - speaker: mapping S from M to F
  - hearer: mapping H from F to M
- **Average utility:** (identical for both players)

$$u(S,H) = \int_M p_m \times \exp(-d(m, H(S(m)))^2)dm$$

vulgo: average similarity between speaker's meaning and hearer's meaning

Voronoi tesselations

- suppose H is given and known to the speaker: which speaker strategy would be the best response to it?
    - every form f has a "prototypical" interpretation: $H(f)$
    - for every meaning m: S's best choice is to choose the f that minimizes the distance between m and $H(f)$
    - optimal S thus induces a **partition** of the meaning space
    - Voronoi tesselation, induced by the range of H



Voronoi tesselation

Lemma

The Voronoi tessellation based on a Euclidean metric always results in a partioning of the space into convex regions.

ESSs of the naming game

- best response of H to a given speaker strategy S not as easy to characterize
- general formula

$$H(f) = \arg\max_m \int_{S^{-1}(f)} p_{m'} \times \exp(-d(m, m')^2) dm'$$

- such a hearer strategy always exists
- linguistic interpretation: H maps every form f to the **prototype** of the property $S^{-1}(f)$

ESSs of the naming game

Lemma

In every ESS $\langle S, H \rangle$ *of the naming game, the partition that is induced by* S^{-1} *on* M *is the Voronoi tesselation induced by* $H[F]$.

ESSs of the naming game

Lemma

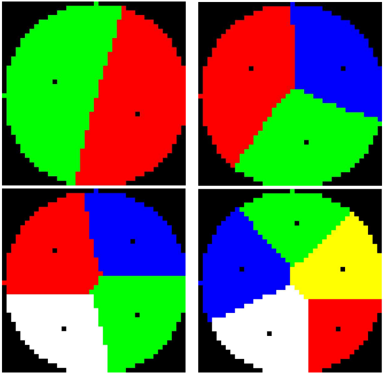
In every ESS $\langle S, H\rangle$ of the naming game, the partition that is induced by S^{-1} on M is the Voronoi tesselation induced by $H[F]$.

Theorem

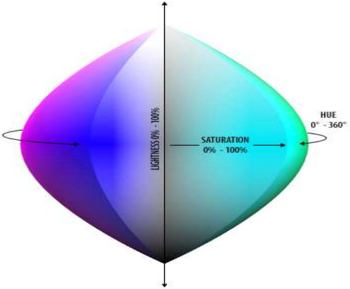
For every form f, $S^{-1}(f)$ is a convex region of M.

Simulations

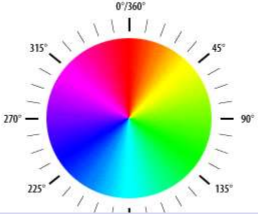
- two-dimensional circular meaning space
- discrete approximation
- uniform distribution over meanings
- initial stratgies are randomized
- update rule according to (discrete time version of) replicator dynamics



The color space

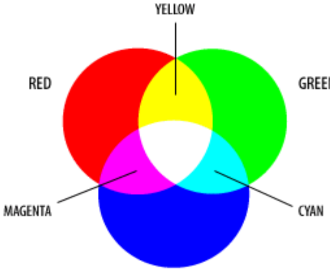


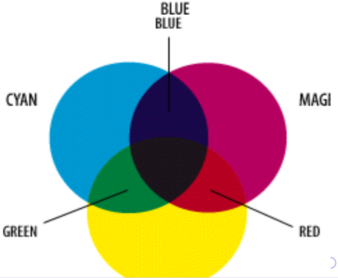
- physical color space is of infinite dimensionality
- psychological color space has only three dimensions:
  1. brightness
  2. hue
  3. saturation



The color space

- alternative axes (but maintaining dimensionality of three)
  1. black-white
  2. red-green
  3. yellow-blue
- yet another triple of dimensions ("additive"):
  1. red
  2. green
  3. blue
- "subtractive" color space:
  1. cyan
  2. magenta
  3. yellow

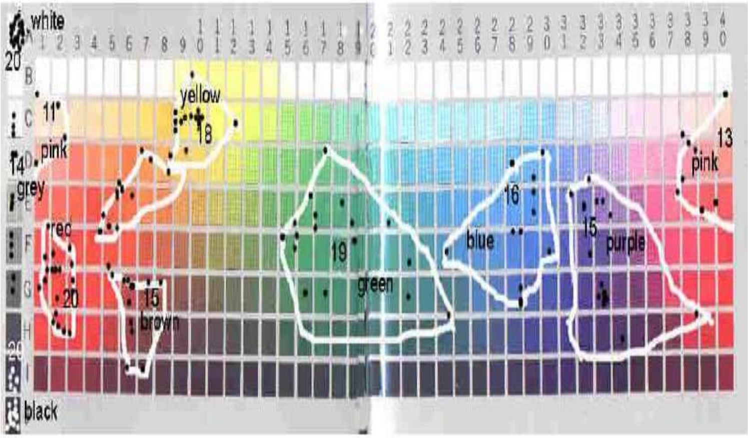




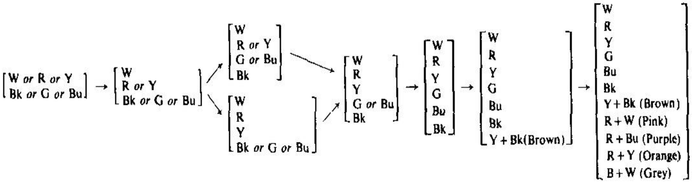
Color words

- Berlin and Kay (1969): study of the typology of color words
- subjects with typologically distant native languages
- subjects were asked about prototype and extension of the basic color words of their native language
- English: 11 basic colors

Berlin and Kay's study



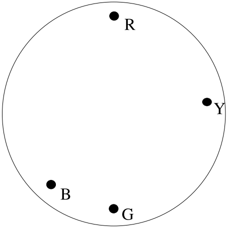
Implicational hierarchies



A toy example

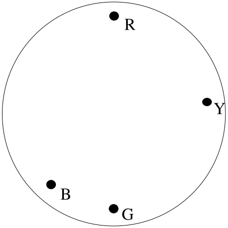
- suppose
  - circular two-dimensional meaning space
  - four meanings are highly frequent
  - all other meanings are negligibly rare
- let's call the frequent meanings
  Red, Green, Blue and Yellow

$$p_i(\text{Red}) > p_i(\text{Green}) > p_i(\text{Blue}) > p_i(\text{Yellow})$$



A toy example

- suppose
  - circular two-dimensional meaning space
  - four meanings are highly frequent
  - all other meanings are negligibly rare
- let's call the frequent meanings Red, Green, Blue and Yellow

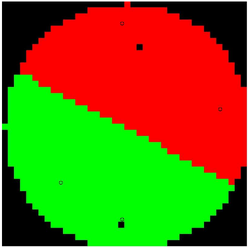


$$p_i(\text{Red}) > p_i(\text{Green}) > p_i(\text{Blue}) > p_i(\text{Yellow})$$

Yes, I made this up without empirical justification.

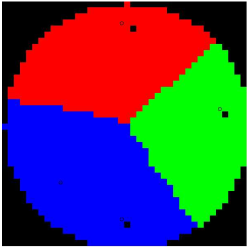
Two forms

- suppose there are just two forms
- only one Strict Nash equilibrium (up to permuation of the forms)
- induces the partition {**Red, Blue**}/{**Yellow, Green**}



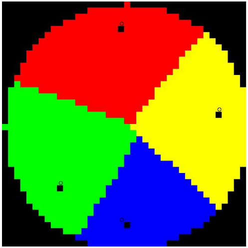
Three forms

- if there are three forms
- two Strict Nash equilibria (up to permuation of the forms)
- partitions {Red}/{Yellow}/{Green, Blue} and {Green}/{Blue}/{Red, Yellow}
- only the former is **stochastically stable** (resistent against random noise)



Four forms

- if there are four forms
- one Strict Nash equilibrium (up to permuation of the forms)
- partitions {Red}/{Yellow}/{Green}/{Blue}



Conclusion

Meaning spaces

- assumption: utility is correlated with similarity between speaker's meaning and hearer's meaning
- consequences:
  - convexity of meanings
  - prototype effects
  - uneven probability distribution over meanings leads to the kind of implicational universals that are known from typology of color terms

Don't talk to strangers:

Spatial EGT

Spatial EGT

- idealized assumption of standard EGT:
  - populations are infinite
  - each pair of individuals is equally likely to interact with each other
- Stochastic EGT gives up the first assumption
- What happens if you give up second assumption as well?

Spatial EGT

- one possible instantiation:
  - individuals are arranged in a spatial structure
  - every individual only interacts with its immediate neighbors

Spatial EGT

Suppose we have

- set of **positions** *pos*
- irreflexive **neighbourhood** relation n among *pos*
- **strategy function** *st* maps positions and time points random variable over strategies
- **density function** d maps positions/time points to positive real number
- **fitness function** f assigns fitness value (positive real) to positions/time points
- $Z(a,t)$: normalization variable; accumulated weighted fitness of the neighborhood of a at time t

Spatial EGT

$$
\begin{aligned}
f(a,t+1) &= \sum_{b:n(a,b)} u(st(a,t), st(b,t)) \\
d(a,t+1) &= d(a,t) \times f(a,t+1) \\
P(st(a,t+1)=i) &= \frac{1}{Z(a,t+1)} \times \\
&\quad \sum_{(b\in\{x:n(a,x)\}\cup\{a\})\cap\{x:st(x,t)=i\}} d(b,t+1) \times f(b,t+1) \\
Z(a,t+1) &= \sum_{b\in\{x:n(a,x)\}\cup\{a\}} d(b,t+1) \times f(b,t+1)
\end{aligned}
$$

Spatial structure

- two-dimensional chessboard like structure
- neighborhood: adjacent fields; each field has eight neighbors
- torus shape: upper and lower boundaries are neighbors, and likewise left and right boundaries

Spatial Prisoner's dilemma

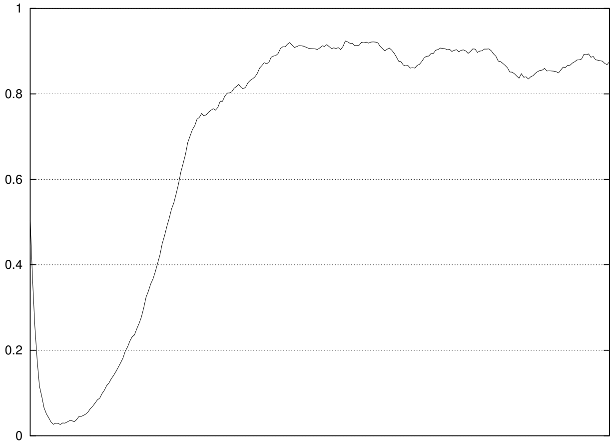
- one version of Prisoner's dilemma:

	C	D
C	5,5	1,6
D	6,1	2,2

- standard EGT: one ESS: (D, D)
- spatial EGT:
  - only interaction with neighbors
  - neighbors are likely to be "related" to each other
  - increased likelihood of interactions between individuals with identical strategies
  - favors strategies with high utility against itself, even if not NE

Spatial Prisoner's dilemma

- proportion of C-players in a spatial Prisoner's dilemma:



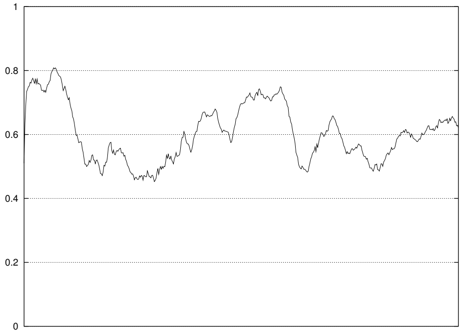
Spatial Hawks and Doves

- spatial evolution generally favors intra-strategy altruism
- should favor Doves over Hawks

	H	D
H	1	7
D	2	3

Spatial Hawks and Doves

- development of the proportion of hawks in spatial HaD
- proportion of doves is most of the time higher than in the ESS (20%)



Game of communication

- row strategies:
  - T: talk
  - S: remain silent
- column strategies
  - A: pay attention
  - I: ignore
- only one ESS: (S, I)

	A	I
T	1,2	0,1
S	1,0	1,1

Spatial game of communication

- symmetrized game of communication:

	(T, A)	(T, I)	(S, A)	(S, I)
(T, A)	3	2	1	0
(T, I)	2	1	2	1
(S, A)	3	3	1	1
(S, I)	2	2	2	2

- "cooperative" strategy pair (T, A) forms stable clusters