Evolutionary Game Theory and Linguistics

Gerhard Jäger
Gerhard.Jaeger@uni-bielefeld.de

February 23, 2007

University of Tübingen

Cognitive semantics

Gärdenfors (2000):

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

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**ender
  - **R**eceiver
- 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 R
  3. R maps f to a meaning m'

The naming game

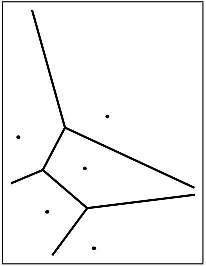
- **Goal:**
  - optimal communication
  - both want to minimize the distance between m and m'
- **Strategies:**
  - speaker: mapping S from M to F
  - hearer: mapping R from F to M
- **Average utility:** (identical for both players)

$$u(S,R) = \sum_m p_m \times \exp(-d(m, R(S(m)))^2)$$

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

Voronoi tesselations

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



Voronoi tesselation

Okabe et al. (1992) prove the following lemma (quoted from Gärdenfors 2000):

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 R to a given speaker strategy S not as easy to characterize
- general formula

$$R(f) = \arg\max_{m} \sum_{m' \in S^{-1}(f)} p_{m'} \times \exp(-d(m, m')^2)$$

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

ESSs of the naming game

Lemma

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

ESSs of the naming game

Lemma

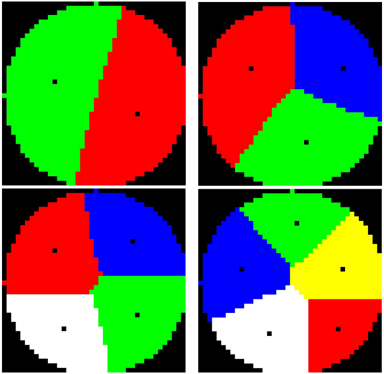
In every ESS $\langle S, R\rangle$ *of the naming game, the partition that is induced by* S^{-1} *on* M *is the Voronoi tesselation induced by* $R[F]$.

Theorem

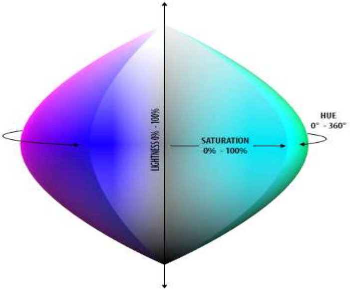
For every form f, $S^{-1}(f)$ *is a convex region of* M.

Simulations

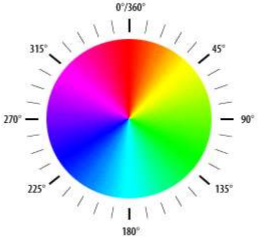
- two-dimensional circular meaning space
- discrete approximation
- uniform distribution over meanings
- initial stratgies are randomized
- update rule according to (discrete time version of) replicator dynamics



The color space

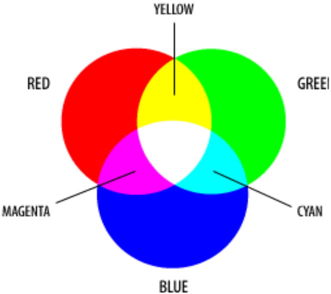


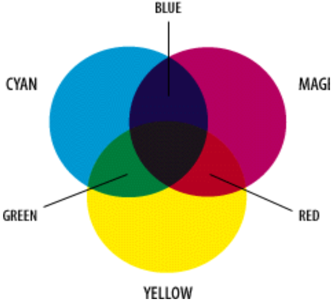
- physical color space is of infinite dimensionality
- psychological color space has only three dimensions:
  1. brightness
  2. hue
  3. saturation



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

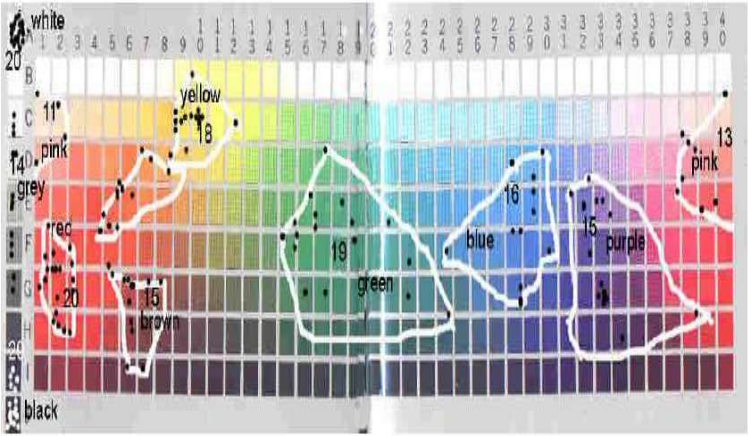




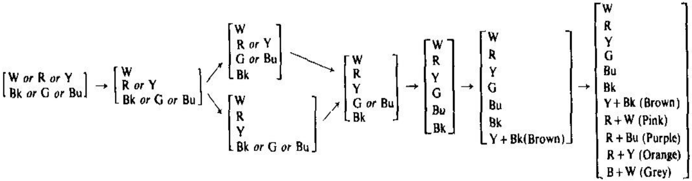
Color words

- 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



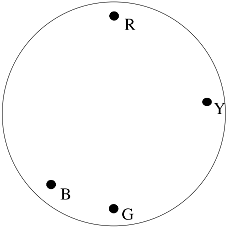
Implicational hierarchies



A toy example

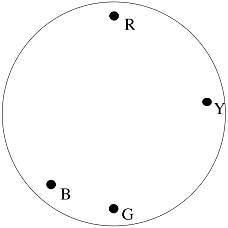
- suppose
  - circular two-dimensional meaning space
  - four meanings are highly frequent
  - all other meanings are negligibly rare
- let's call the frequent meanings
  Red, Green, Blue and Yellow

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



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

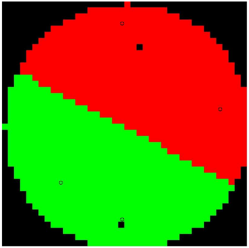


$$p_i(\text{Red}) > p_i(\text{Green}) > p_i(\text{Blue}) > p_i(\text{Yellow})$$

Yes, I made this up without empirical justification.

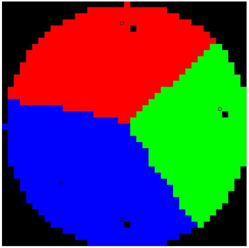
Two forms

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



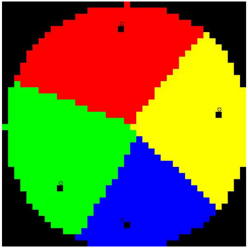
Three forms

- 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}



Measure terms

Krifka's observations

- measure terms are vague
- some measure terms are ambiguous between different degrees of vagueness
- usually only simple expressions are ambiguous in this way
- complexifying an expression may reduce ambiguitiy

vagueness

95 m: between 94.5 and 95.5 m

ambiguity

- *The water has a temperature of 40°:* $38° < T < 42°$
- *His body temperature is 40°:* $39.95° < T < 40.05°$

simple and complex expression

His body temperature is 39°: cannot mean $37° < T < 41°$

complexification

The water has a temperature of exactly 40°: $39.9° < T < 40.1°$

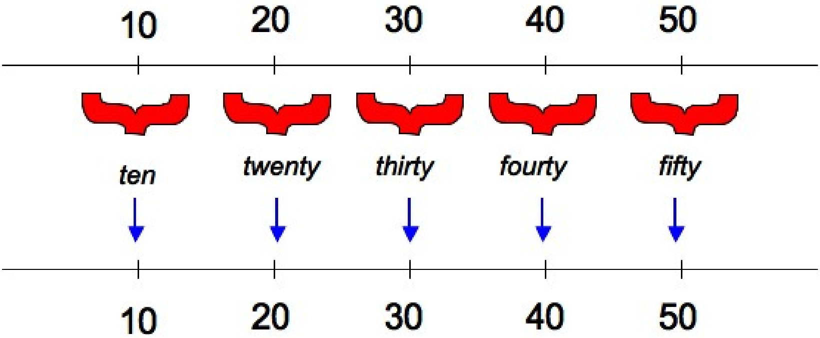
General considerations

- Suppose the game setup is as before, with arithmetic difference as distance function

ESS

- Sender:
  - meaning space is partitioned into continuous intervals of equal length
  - each interval is correlated with one signal
- Receiver:
  - each signal is mapped to the center of the corresponding interval

General considerations



Costly signaling

- suppose signals incur a cost for both sender and receiver
- modified utility function

$$u(S, R) = \sum_m p_m \exp(-(m - R(S(m)))^2) - c(S(m))$$

- intuitive idea:

$$c(\text{thirty-nine}) > c(\text{forty})$$

etc.

Costly signaling

ESSets

- general pattern as before
- additional constraint: in an ESS (S, R), we have

$$\forall m : S(m) = \arg_f \max[\exp(-(m - R(f))^2) - c(f)]$$

- simultaneous
  - minimizing distance between m and $R(S(m))$
  - minimizing costs $c(S(m))$
- in equilibrium (ESSet), distance between m and $R(S(m))$ need not be minimal

Variable standard of precision

Assessment

- this setup
  - predicts the possibility of vague interpretation: **good**
  - fails to predict the ambiguity between precise and vague interpretations (or different degrees of vagueness): **bad**

Variable standard of precision

Proposal

- required degree of precision depends on context
- modeling as Bayesian game with different utility function
- both players still have same utility function and know that function

$$u(S,R) = \sum_{m,\sigma} p_{m,\sigma} \exp(-(m - R(S(m)))^2/\sigma^2) - c(S(m))$$

- high value of σ: precision doesnt matter very much
- low value of σ: precision is more important than economy of expression

An example

- Suppose:
  - just two meanings: 39, 40
  - just two forms: *thirty-nine*, *forty*

$$c(\textit{thirty-nine}) - c(\textit{forty}) = \mathbf{c} > 0$$

- two standards of precision, σ_1 and σ_2

$$\begin{aligned}
\sigma_1 &< \sigma_2 \\
\exp(-(1^2/\sigma_1^2)) &= d_1 \\
\exp(-(1^2/\sigma_2^2)) &= d_2 \\
1 - d_1 &> \mathbf{c} \\
1 - d_2 &< \mathbf{c} \\
\forall m, \sigma : p_{m,\sigma} &= .25
\end{aligned}$$

An example

Intuitive characterization

- two standards of precision
- utility loss under vague interpretation is $1 - d_i$
- utility loss due to usage of more complex expression is $\mathbf{c}$
- under σ_1 precision is more important
- under σ_2 economy of expression is more important
- uniform probability distribution over states

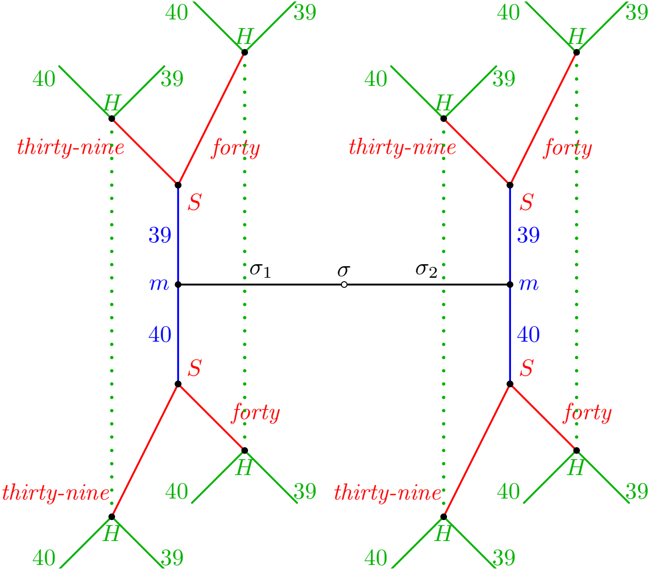
meanings/signals

S		R
39	*thirty-nine*	39
40	*forty*	40

strategies

- S_1/R_1 : 
- S_2/R_2 : 
- S_3/R_3 : 
- S_4/R_4 : 

Extensive form



Utility matrices

σ_1

				
	$1 - \frac{c}{2}$	$d_1 - \frac{c}{2}$	$\frac{1+d_1-c}{2}$	$\frac{1+d_1-c}{2}$
	$d_1 - \frac{c}{2}$	$1 - \frac{c}{2}$	$\frac{1+d_1-c}{2}$	$\frac{1+d_1-c}{2}$
	$\frac{1+d_1}{2}$	$\frac{1+d_1}{2}$	$\frac{1+d_1}{2}$	$\frac{1+d_1}{2}$
	$\frac{1+d_1}{2} - c$	$\frac{1+d_1}{2} - c$	$\frac{1+d_1}{2} - c$	$\frac{1+d_1}{2} - c$

σ_2

				
	$1 - \frac{c}{2}$	$d_2 - \frac{c}{2}$	$\frac{1+d_2-c}{2}$	$\frac{1+d_2-c}{2}$
	$d_2 - \frac{c}{2}$	$1 - \frac{c}{2}$	$\frac{1+d_2-c}{2}$	$\frac{1+d_2-c}{2}$
	$\frac{1+d_2}{2}$	$\frac{1+d_2}{2}$	$\frac{1+d_2}{2}$	$\frac{1+d_2}{2}$
	$\frac{1+d_2}{2} - c$	$\frac{1+d_2}{2} - c$	$\frac{1+d_2}{2} - c$	$\frac{1+d_2}{2} - c$

Results

Evolutionary stability

- first subgame (σ_1; precision is important): two ESS
  - S_1/R_1
  - S_2/R_2
- in either case, both expressions have a precise meaning and are interpreted exactly as intended
- second subgame (σ_2; economy of expression is important): one ESSet
  - consists of S_3 and all mixed strategies of R
- Bayesian game:
  - two ESSets
  - any combination of ESSets of the two sub-games

Assessment

- this setup
- predicts that
  - all number words receive a precise interpretation if precision is important
  - only short number words are used and receive a vague interpretation if economy is important
- **good**
- with larger dictionary prediction that there is no correlation between the interpretation of words between the different subgames
- for instance:
  - *forty* could mean 40 for σ_1 and $\{28...32\}$ for σ_2
- **bad**

Asymmetric information

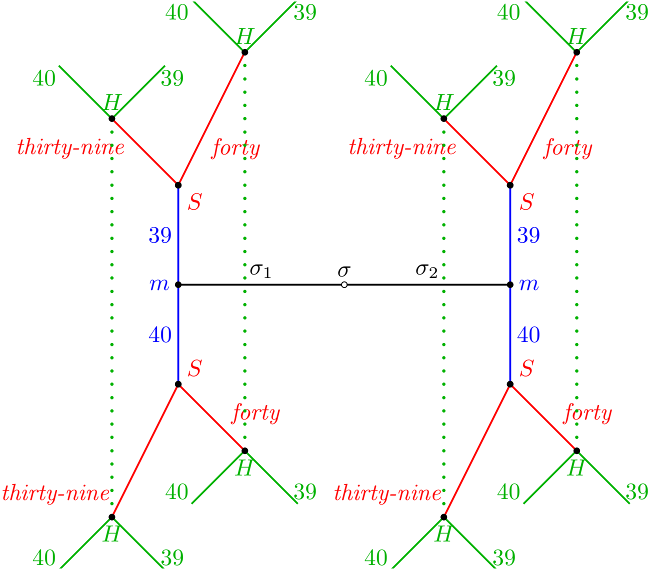
Modified information sets

- idea
  - S knows σ, but
  - R doesn't
- then R's interpretation of a word cannot depend on σ

Strategy space

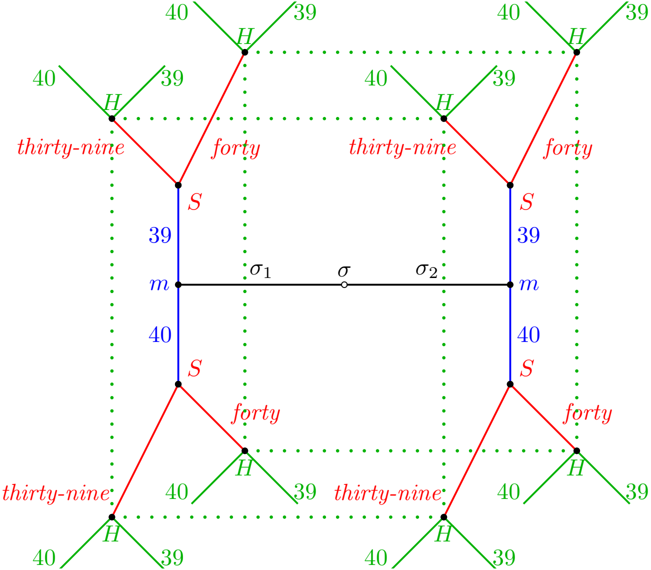
- Sender strategies:
  - functions from pairs (m, σ) to signals
  - in the example: $4 \times 4 = 16$ strategies, as before
- Receiver's strategies
  - functions from signals to meanings
  - in the example: only four such functions (as in the first version of the example)

old game:

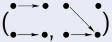
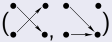

40
39
H
40
39
H
thirty-nine
forty
S
39
m
σ_1
σ
σ_2
40
S
forty
H
40
39
thirty-nine
H
40
39
40
39
H
40
39
H
thirty-nine
forty
S
39
m
40
S
forty
H
40
39
thirty-nine
H
40
39

Extensive form

new game:



ESS

- resulting game has only two ESSs
  - ESS 1:
    - S: 
    - R: 
  - ESS 2:
    - S: 
    - R: 
- in either case
  - R always assumes precise interpretation
  - S always chooses correct word if σ is low
  - S always chooses short word if σ is high

Loose ends

Open questions

- notion of ESS/ESSet only make sense for finite strategy space
- can results be maintained if meaning space is really continuous?
- S's signal gives information about value of σ
- perhaps R's guess about value of σ should enter the utility function
- would explain why
  - it can be rational for S to use excessively complex phrases like *exactly fourty* and short phrases like *fourty* synonymously
  - *exactly fourty* can only be interpreted precisely, while *fourty* is ambiguous