Meaning, Evolution and the Structure of Society

Roland Mühlenbernd

November 7, 2014
Accounts differ in *psychology* of language users
What role does the *sociology* of language users play?
EMERGENCE OF HUMAN LANGUAGE

“[w]e can hardly suppose a parliament of hitherto speechless elders meeting together and agreeing to call a cow a cow and a wolf a wolf.”

Russell, The Analysis of Mind

“Two savages, who had never been taught to speak, but had been bred up remote from the societies of men, would naturally begin to form that language by which they would endeavor to make their mutual wants intelligible to each other, by uttering certain sounds, whenever they meant to denote certain objects.”

Smith, Considerations Concerning the First Formation of Languages
DEFINITION OF A SIGNALLING GAME (Lewis game)

Signaling Game $SG = \langle \{S, R\}, T, M, A, Pr, C, U \rangle$

- set of players $\{S, R\}$
- set of states $T \ (T = \{t_1, t_2\})$
- state probability function $Pr \in \Delta(T) \ (Pr(t_1) = Pr(t_2) = .5)$
- set of messages $M \ (M = \{m_1, m_2\})$
- cost function $C : M \rightarrow \mathbb{R} \ (C(m_1) = C(m_2) = 0)$
- set of actions $A \ (A = \{a_1, a_2\})$
- utility function $U(t_i, m_j, a_k) = \begin{cases} 
1 - C(m_j) & \text{if } i = k \\
0 - C(m_j) & \text{else} 
\end{cases}$

$\begin{array}{c|cc}
        & a_1 & a_2 \\
\hline
  t_1  & 1,1 & 0,0 \\
  t_2  & 0,0 & 1,1 \\
\end{array}$
PURE STRATEGIES

- the players’ “actions” can be represented as pure strategies
- set of pure sender strategies $S = \{s | s \in [T \rightarrow M]\}$
- set of pure receiver strategies $R = \{r | r \in [M \rightarrow A]\}$
- for the Lewis game there are 4 strategies for each player
- $S = \{s_1, s_2, s_3, s_4\}$
- $R = \{r_1, r_2, r_3, r_4\}$
are combinations of pure strategies. The Lewis game has two: \( L_1 = \langle s_1, r_1 \rangle \) and \( L_2 = \langle s_2, r_2 \rangle \)

\[
\begin{align*}
L_1: & \quad t_L \rightarrow m_1 \rightarrow a_L \\
\quad & \quad t_s \rightarrow m_2 \rightarrow a_S \\
L_2: & \quad t_L \times m_1 \times a_L \\
\quad & \quad t_s \times m_2 \times a_S
\end{align*}
\]

are strict *Nash equilibria* over expected utilities

associate messages to states in an unique way

are *evolutionary stable states*
EXPECTED UTILITIES OF THE LEWIS GAME

The expected utility for a combination of strategies is given as:

\[
EU(s, r) = \sum_{t \in T} Pr(t) \times U(t, s(t), r(s(t)))
\]  

(1)

<table>
<thead>
<tr>
<th></th>
<th>(r_1)</th>
<th>(r_2)</th>
<th>(r_3)</th>
<th>(r_4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(s_1)</td>
<td>1</td>
<td>0</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td>(s_2)</td>
<td>0</td>
<td>1</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td>(s_3)</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td>(s_4)</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
</tbody>
</table>

Table: The static Lewis game: Expected utilities for all strategy combinations of the dynamic Lewis game
**The Horn Game**

\[ SG = \langle \{S, R\}, T, M, A, Pr, C, U \rangle \]

- \( T = \{t_f, t_r\}, M = \{m_u, m_m\}, A = \{a_f, a_r\} \)
- \( Pr(t_f) > Pr(t_r) \)
- \( C(m_u) < C(m_m) \)
- \( S = \{s_h, s_a, s_s, s_y\} \)
- \( R = \{r_h, r_a, r_s, r_y\} \)

\[ s_h: \quad t_f \rightarrow m_u \quad t_r \rightarrow m_m \]
\[ s_a: \quad t_f \quad m_u \quad t_r \quad m_m \]
\[ s_s: \quad t_f \rightarrow m_u \quad t_r \rightarrow m_m \]
\[ s_y: \quad t_f \rightarrow m_u \quad t_r \rightarrow m_m \]

\[ r_h: \quad m_u \rightarrow a_f \quad m_m \rightarrow a_r \]
\[ r_a: \quad m_u \quad a_f \quad m_m \quad a_r \]
\[ r_s: \quad m_u \rightarrow a_f \quad m_m \rightarrow a_r \]
\[ r_y: \quad m_u \rightarrow a_f \quad m_m \rightarrow a_r \]
EXPECTED UTILITIES OF THE HORN GAME

Exemplary parameters for the Horn Game:

- $Pr(t_f) = .7$
- $C(m_u) = .1, C(m_m) = .2$

<table>
<thead>
<tr>
<th></th>
<th>$r_h$</th>
<th>$r_a$</th>
<th>$r_s$</th>
<th>$r_y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_h$</td>
<td>.87</td>
<td>-.13</td>
<td>.57</td>
<td>.17</td>
</tr>
<tr>
<td>$s_a$</td>
<td>-.17</td>
<td>.83</td>
<td>.53</td>
<td>.13</td>
</tr>
<tr>
<td>$s_s$</td>
<td>.6</td>
<td>.2</td>
<td>.6</td>
<td>.2</td>
</tr>
<tr>
<td>$s_y$</td>
<td>.1</td>
<td>.5</td>
<td>.5</td>
<td>.1</td>
</tr>
</tbody>
</table>

Table: Payoff table of the asymmetric static Horn game with $Pr(t_f) = .75, C(m_u) = .1$ and $C(m_m) = .2$
The Program of Today

▶ Which strategy emerges under evolutionary (population) dynamics like the replicator dynamics?
▶ How might the outcome differ for agent-based dynamics like imitation or learning?
▶ How does the outcome vary for different typologies?
▶ What role might structural aspects play for the emergence of specific strategies?
**REPLICATOR DYNAMICS**

Given a very large (effectively infinite) population of agents playing a symmetric static game $\langle \{P_1, P_2\}, S, U : S \times S \rightarrow \mathbb{R} \rangle$ randomly against each other. Then we can define

- $p(s_i)$: proportion of agents in the population playing strategy $s_i$
- $U(s_i) = \sum_{s_j \in S} p(s_j)U(s_i, s_j)$: expected utility for agents playing $s_i$
- $U = \sum_{s_i \in S} p(s_i)U(s_i)$ the average fitness of the whole population

**Definition (Replicator Dynamics)**

*The RD is defined by the following differential equation:*

$$\frac{dp(s_i)}{dt} = p(s_i)[U(s_i) - U]$$
**Evolutionary Stable Strategy**

**Definition (Evolutionary stable strategy)**

For a symmetric static game \( \langle \{P_1, P_2\}, S, U : S \times S \rightarrow \mathbb{R} \rangle \) a strategy \( s_i \in S \) is said to be a evolutionary stable strategy if and only if one of the following two conditions holds:

1. \( U(s_i, s_i) > U(s_i, s_j) \) for all \( s_j \neq s_i \)
2. if \( U(s_i, s_i) = U(s_i, s_j) \) for some \( s_j \neq s_i \), then \( U(s_i, s_j) > U(s_j, s_j) \)

A ESS \( s_i \) has the following properties:

- \( s_i \) has an invasion barrier: resistant against the invasion of a small proportion of mutants with \( s_j \neq s_i \)
- \( s_i \) is a attractor under the replicator dynamics
- \( \text{Strict NE} \subset \text{ESS} \subset \text{NE} \)
HORN GAME UNDER REPLICATOR DYNAMICS

Strategies: *Horn*, *anti-Horn*, *Smolensky*, *anti-Smolensky*

Start population: \([.25, .25, .25, .25]\)
**Basin of Attraction and Game Parameters**

Horn game 1:
- \( Pr : [.6, .4] \)
- \( C : [.05, .1] \)
- \( L_h: 55\% \), \( L_a: 43\% \), \( L_s: 2\% \)

Horn game 2:
- \( Pr : [.7, .3] \)
- \( C : [.1, .2] \)
- \( L_h: 59\% \), \( L_a: 35\% \), \( L_s: 6\% \)

Horn game 3:
- \( Pr : [.9, .1] \)
- \( C : [.1, .3] \)
- \( L_h: 64\% \), \( L_a: 22\% \), \( L_s: 14\% \)
AND NOW?

"...the replicator dynamics is a natural place to begin investigations of dynamical models of cultural evolution, but I do not believe that it is the whole story."

Skyrms, Evolution of the Social Contract
IMITATION DYNAMICS

The replicator dynamics describes the most likely path of strategy distribution in a virtually infinite and homogeneous population, when every agent updates her behavior by conditional imitation.

shown by e.g. Helbing 1996, Schlag 1998

Conditional Imitation: the probability of an agent playing strategy $s$ to switch to a randomly chosen participant’s strategy $s'$ is given as:

$$P(s \rightarrow s') = \begin{cases} 
0 & \text{if } U(s, s') \geq U(s', s') \\
\alpha \times \frac{U(s', s') - U(s, s')}{U_{\text{max}} - U_{\text{min}}} & \text{else} 
\end{cases}$$

(2)

with:

- $\alpha$: scaling factor
- $U_{\text{max}}$: maximal value of the utility table
- $U_{\text{min}}$: minimal value of the utility table
IMITATION VS REPLICATOR DYNAMICS

Figure: Basins of attraction for the replicator dynamics and the conditional imitation rule by numerical simulation for a population of 100 randomly interacting agents.

→ the Correspondence of basins of attraction is 93.2%
IMITATION ON DIFFERENT TYPOLORIES

Simulations with 100 agents on different typologies:

- $\beta$-graph (Watts-Strogatz algorithm, starting ring)
- scale-free networks (Holme-Kim algorithm)
- $10 \times 10$ toroid lattice

Figure: Comparing distributions of basin of attraction and RD-correspondence (percentage in brackets) for different network topologies
The Evolution of Regional Meaning

Zollman (2005) made experiments with agents

- playing the Lewis game
- updating by Imitation
- placed on a 100 × 100 toroid lattice

Result: the emergence of regional meaning

Agents play the dynamic game $SG = \langle (S, R), T, M, A, Pr, C, U \rangle$

- as probabilistic strategies $s : T \rightarrow \Delta(M)$, $r : M \rightarrow \Delta(A)$
- communicative outcome shifts probabilities in appropriate directions
- successful communication via $\langle t \in T, m \in M, a \in A \rangle$ makes the choices $s(m|t)$ and $r(a|m)$ more probable in subsequent rounds
- agents might learn a pure strategy profile that constitutes a signaling system
The Evolution of Regional Meaning

Experiments on a $30 \times 30$ toroid lattice:

The Lewis Game

The Horn Game

Regional Convention evolve not always for the Horn Game.
EXPERIMENTS ON A SOCIAL MAP

- The closer the distance on the lattice, the more probable the interaction
- Relationship controlled by degree of locality $\gamma$

Local interaction supports the emergence of local conventions.
Small World Networks

\( \beta \)-graph network
(Watts & Strogatz 1998)

scale-free network
(Holme & Kin 2002)
EXPERIMENTAL SETUP

Simulation runs for populations of 300 agents
  ▶ placed on small world network interaction structures
  ▶ ’interact’ via Lewis game and Horn game
  ▶ decide and update by learning dynamics: reinforcement learning or belief learning
  ▶ interact until 90% of all agents are language learners

Analysis of
  ▶ properties of language regions
  ▶ interrelationships between agents’ network position and learning behavior
  ▶ specific roles of agent in the emerging process of conventions/signaling languages
**Results: Language Regions**

Cliquishness supports the emergence of language regions.
CONCLUSION

Social structure matters to the evolution and establishment of conventional meaning!

- different interaction structures lead to shifts in the basin of attraction of communication strategies
- locality of structure supports regional variety
- particular network properties support the emergence of regional conventions