Evolutionary games explain efficient language organization

Robert van Rooy, ILLC, Amsterdam, R.A.M.vanRooij@uva.nl

Recently, evolutionary game theory (EGT) has been used (e.g. Nowak) to study the emergence of *syntactic* (e.g. *compositionality*) and *semantic* (lexical entries) features of natural language. Here it is used to explain *pragmatic* linguistic principles.

Consider the case where two meanings m_1 and m_2 can be expressed by two linguistic signals s_1 and s_2. In principle this gives rise to two possible codings: $\{\langle m_1, s_1\rangle, \langle m_2, s_2\rangle\}$ and $\{\langle m_1, s_2\rangle, \langle m_2, s_1\rangle\}$. In many communicative situations, however, the underspecification does not really exist, and is resolved (e.g. by the use of pronouns) due to the general pragmatic principle that a lighter form will be interpreted by a more salient, or stereotypical, meaning. If we can explain this principle, we can also explain why language is organized so efficiently. To do so, however, we need, first, to explain why one way of resolving the underspecification is more natural than the other, and second, to show why underspecification of meaning is useful in the first place.

To explain both, we will make use of *signaling games* as introduced by David Lewis (1969) to account for linguistic conventions, and developed further in economics and theoretical biology. In this framework, signals have an underspecified meaning, and the actual interpretation the signals receive depend on the equilibria of sender and receiver strategy combinations of such games. Recently, these games have been looked upon from an *evolutionary* point of view to study the evolution of language. According to it, a coding (or signaling) convention can arise according to which signal s means m if and only if the pair $\langle s, m\rangle$ is part of an *evolutionary stable strategy* (ESS). Unfortunately, one can show (Wärneryd, 1993) that the ESSs of signaling games always give rise to 1-1 mappings between signals and meanings. But this predicts false: underspecification (or homonymy) of meaning is predicted not to exist, though in fact it is the rule rather than the exception in natural languages. So, if evolutionary game theory is to be a useful tool to investigate the evolution of language, it better is able to explain why and how we make use of expressions with incompletely specified conventional meanings.

It is. The solution is based on three ideas. First, and obviously: *underspecification* makes sense because speaker and hearer share a common *context* which helps resolving what is intended. We will show that languages that make 'smart' use of contexts are evolutionary stable. However, they are not the only ones. To select the 'smart' ones, we use a second idea and take into account (i) the *costs* of sending signals, and (ii) the *probabilities* of the meanings. As a result, of all evolutionary stable strategies, only the 'smart' ones are Pareto optimal. Still, standard evolutionary game theory gives no reason why only those should emerge. As the third idea, I propose two possible solutions: *correlation* (or clustering) and *mutation*. The first assumes that agents tend to speak more with others that use similar strategies (languages). One can show that assuming correlation in EGT gives rise to the emergence of strategies with the highest expected utility, are Pareto optimal. The second proposal assumes that the evolutionary transition from one generation to the next is *stochastic* in nature. One natural way to think of this is as being due to imperfect language acquisition. General game theoretical results (e.g. Young, 1990) show that such an evolutionary process gives rise to risk-dominant equilibria, which in cooperative games are equal to the Pareto optimal ones.

If time permits, I will discuss the naturalness of those two solutions and give evolutionary motivations of other pragmatic interpretation principles (such as the Gricean maxims of quantity and quality) as well.

.