Probe and Adjust and Signaling Chains Model

January 15, 2015

This model description follows the ODD (Overview, Design concepts, Details) protocol [4, 5]

1 Purpose

This model shows the evolution of a signaling convention on a Lewis' signaling chain. Lewis' signaling chains are a particular case of signaling networks in which a sender tries to send a single bit of information to a receiver trough a chain of players that do not share a common signaling system. Signaling, chains and especially signaling networks have interesting applications in organization science were they can be used to model small group learning [2], and in computer science were they can be used in machine learning and pattern recognition [7]. Similarly chains networks may be used to understand how simple organism, such as Dictyostelium colonies are able to organize from unicellular amoebae into a multicellular slug through the use of complex chains of chemical signals [9], or how eukaryote cells use chemical and mechanical interactions to generate and maintain tissues the tissues involved for example in the vasculatory and in the skeletal system [1]. Consequently, this may help us understand how humans and organizations can coordinate each other and solve complex collective challenges for example design complex trade and transportation networks, such us the salt trading routes or the silk road. Furthermore, probe and adjust is a learning strategy that demands minimal memory requirements and is for this reason suited to model simple organism or humans in large organizations [8, 6].

2 Entities, state variables, and scales

Both Lewis' signaling games and signaling chains are extensive form games that are repeated at each time $t = 0, 1, \ldots$. In signaling games, there are three fundamental elements. A sender, a receiver, and a state of Nature which provides random events that are independent of the players behavior. In signaling chains, there are n players and four entities Nature which chooses its state with some probability $x_{0t} \in \{1, 2\}$, the sender which sends a signal $x_{1t} \in \{1, 2\}$, the receiver which must choose an action $x_{nt} \in \{1, 2\}$, and all other players $2 \leqslant i < n$ are transmitters which sends a signal $x_{it} \in \{1, 2\}$.

In both Lewis' signaling games and signaling chains nature selects its state according to a Bernoulli probability distribution where probability of drawing $x_{0t} = 0$ is p_0, and $x_{0t} = 1$ is $1 - p_0$. Than each player behavior is determined from its stimulus-response association, $p_{it} : X_{i-1} \longmapsto X_i$. At each time interval t any player may decide to probe. Finally, it uses a payoff memory m_{is_i}, an association proposal p'_{it}

3 Process overview and scheduling

For each time interval $t = 1, 2, \ldots$ the model follows the following three steps:

Probe Decision Decide if and which player probes, this is done by a central scheduler because we assume just one agent at the time probes.

Game Session The players participate at a signaling game session, more in general this may be any extensive form game.

Adjust Decision All players adjust according to the game session resolt.

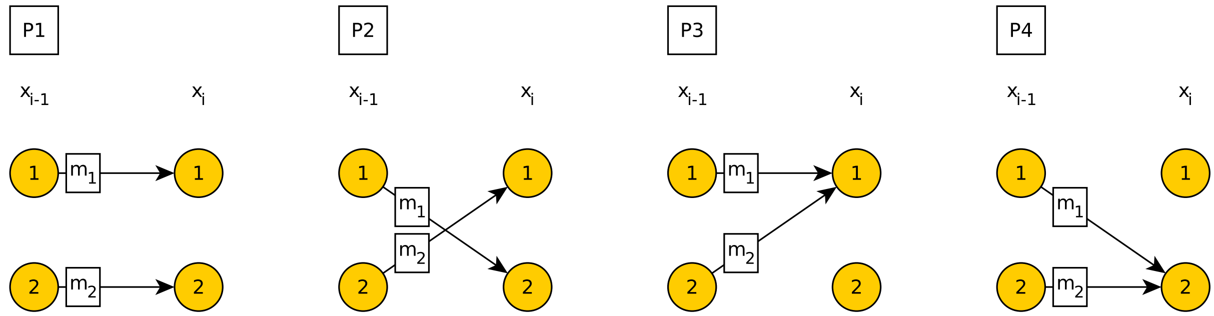

Figure 1: These diagrams are a representation of the player i's state. Where for the sender $i = 1$ and for the receiver $i = 2$. The arrows represent the stimulus-response association for the player, and the $m_{ix_{i-1,t}}$ is the last payoff the player i got when it responded to the stimulus $x_{i-1,t}$.

4 Design Concept

4.1 Basic principles

The signaling game represents a context in which the ability of the players to coordinate using a shared signaling convention may represent an opportunity to improve their fitness. At each model run the signaling games are repeated multiple time to simulate the natural recurrence of opportunities that signaling network have to improve the fitness of the participates, which allows the evolution of a signaling convention.

4.2 Emergence

If the agent are able to adapt successfully to their environment, after a certain number of game interaction they should be able to develop a signaling system that allows the receiver of the message to chose the correct act.

4.3 Objectives

The objective of the agents is to maximize their utility. In the Lewis' signaling game the agents maximize their utility when the receiver understands its signal, thus its act matches the state of Nature. Payoff is $u(x_{0t}, x_{nt}) = 1$ if the act matches the state $x_{0t} = x_{nt}$, and $u(x_{0t}, x_{nt}) = 0$ otherwise.

4.4 Learning

The agents use probe and adjust learning dynamics [8]. Probe and adjust is a minimal form of learning or adaptation dynamics, in which most of the time players continue acting in the way they have done in the past, and every once in awhile a player experiments, or "probes", a new act. If the experiment is "successful" she adapts [8]. There are two major implementations of probe and adjust. Players can either probe over pure strategies [6], or can probe acts taking in account the information set they are in [8]. Here the later is discussed.

4.5 Stochasticity

Nature selects its state according to a Bernoulli probability distribution where probability of drawing $x_{0t} = 0$ is p_0, and $x_{0t} = 1$ is $1 - p_0$. Because simultaneous probes are assumed to not occur, players' probes S_t are chosen at random from a categorical distribution. In the simple sender-receiver case the categorical distribution has three events $S_t \in \{\text{no probes, sender probes, receiver probes}\}$, and parameters $(1-\epsilon, \epsilon_1, \epsilon_2)$, and in the signaling chain case it has $n+1$ events, and parameters $(1-\epsilon, \epsilon_1, \epsilon_2, \ldots, \epsilon_n)$. Let us define frustrated probes as two probes respectively at t_1 and t_2, where $t_1 < t_2$, such that there is a Nature's state is x_{0t} at times t_1 and t_2 but is not x_{0t} for any t such that $t_1 < t < t_2$ [3]. The model allows the possibility in which frustrated probes are removed, in this case whenever we get a frustrated probe the second probe is removed.

4.6 Collectives

At each time interval t after the Probe Decision step all the player participate in the Game Session which a extensive form game.

4.7 Observations

All Probe Decision S_t, all p_{it}, and all $x_{i,t}$ are measured. From these variables we can compute the actuals utility $u(x_{0t}, x_{2t})$, and the expected utility $E[u]$. Furthermore, from these variable we can determine if the system evolved a signaling system, add how long it stays in the signaling system.

5 Initiation

The model is initiated with assigned number of players n, values of p_{i0} for all $0 < i \leqslant n$, Nature state parameter p_0, and probing probability ϵ. This values are the only parameters of the model.

6 Submodules

6.1 Probe Decision

Suppose probe probability $\epsilon > 0$. Each player uses a payoff memory m_{is_i}, an association proposal p'_{it}.At each time interval t, either no player probes with probability $1-\epsilon$, or with probability ϵ one player i is chosen to probe uniformly at random.

1. If no player probes then for each player i the stimulus-response association is $p_{it} = p'_{it}$.
2. If player i is chosen to probe. All the players $j \neq i$ set the stimulus-response association according to the proposal $p_{jt}(s_j) = p'_{jt}(s_j)$, and player i selects a new stimulus-response association $p_{it}(s_i) \neq p'_{it}(s_i)$ chosen uniformly at random from $P_i \backslash \{p'_{it}(s_i)\}$.

In any case, in the game phase each player selects her act in order according to $x_{it} = p_{it}(s_i)$. Notice that the probability that any probe occurs is ϵ, and since we suppose that the prober is selected uniformly at random from the players' population then the probability that player i probes is $\epsilon_i = \frac{1}{n}\epsilon$. Clearly, probes are not required to be uniform and these specification could be made more general.

6.2 Game Session

In the signaling game case, first the Nature's state x_{0t} is selected, second the sender observes Nature's state and choses a message according to its stimulus-response association $x_{1t} = p_{1t}(x_{0t})$, third the receiver observes the message and selects an act according to its stimulus-response association $x_{2t} = p_{2t}(x_{1t})$, without knowing Nature's state. In the signaling chain case, first the Nature's state x_{0t} is selected, than in sequence each player i selects its state $x_{i,t}$ according to $x_{i-1,t}$ and $p_{it}(x_{i-1,t})$ In both cases, all players get the same payoff $u(x_{0t}, x_{2t})$.

6.3 Adjust Decision

At each time interval t, after the expensive form game is played all players i adjust. There are two cases either the player i did probe or it did not probe. If player i did not probe she updates $m_{is_i} = u(\mathbf{x}_t)$, and $p'_{it+1} = p_{it}$. If player i did probe, there are two cases

1. If $m_{is_i} > u(\mathbf{x}_t)$, she does not adjust, and sets $p'_{it+1} = p'_{it}$.

2. If $m_{is_i} < u(\mathbf{x}_t)$, she does adjust and sets $m_{is_i} = u(\mathbf{x}_t)$, and $p'_{it+1} = p_{it}$.
3. If $m_{is_i} = u(\mathbf{x}_t)$, she flips a fair coin and with probability $\frac{1}{2}$ she adjusts, or otherwise she does not adjust.

References

[1] Alberts, B. *Molecular biology of the cell.* Garland Science, New York, 2008.

[2] Carley, K. Organizational Learning and Personnel Turnover. *Organization Science 3*, 1 (Feb. 1992), 20–46.

[3] Gosti, G. *The Evolution of Naming Conventions on Networks.* University of California, Irvine, 2013.

[4] Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S. K., Huse, G., Huth, A., Jepsen, J. U., Jø rgensen, C., Mooij, W. M., Müller, B., Peer, G., Piou, C., Railsback, S. F., Robbins, A. M., Robbins, M. M., Rossmanith, E., Rüger, N., Strand, E., Souissi, S., Stillman, R. A., Vabø, R., Visser, U., and DeAngelis, D. L. A standard protocol for describing individual-based and agent-based models. *Ecological Modelling 198*, 1-2 (Sept. 2006), 115–126.

[5] Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., and Railsback, S. F. The ODD protocol: A review and first update. *Ecological Modelling 221*, 23 (Nov. 2010), 2760–2768.

[6] Huttegger, S. M., Skyrms, B., and Zollman, K. J. S. Probe and Adjust in Information Transfer Games. *Erkenntnis* (2013), 1–19.

[7] Rumelhart, D., and McClelland, J. *Parallel distributed processing: explorations in the microstructure of cognition. Volume 1. Foundations.* MIT Press,Cambridge, MA, Jan. 1986.

[8] Skyrms, B. Learning to Signal with Probe and Adjust. *Episteme 9*, 02 (July 2012), 139–150.

[9] Weijer, C. J. Collective cell migration in development. *Journal of cell science 122*, Pt 18 (Sept. 2009), 3215–23.