2 × 2 Signaling Game with Perfect Coincidence of Interests

\( \pi_1(t) \) is Sender's (Player 1's) payoff in a given play of the game at period \( t \).

\( \pi_2(t) \) is Receiver's (Player 2's) payoff in a given play of the game at period \( t \).

In the special case where the states are equiprobable, this signaling game has 11 Nash equilibria, only two of which are efficient (signaling systems).

A Conditional form of Reinforcement Learning

The state space is \( \Omega = \{\omega_1, \omega_2\} \).

At her given information set \( \omega \in \Omega \), Sender (Player 1) plays \( A \) with probability \( \frac{\theta_A^s(\omega)}{\theta_A^s(\omega) + \theta_A^r(\omega)} \) and \( B \) with probability \( \frac{\theta_B^r(\omega)}{\theta_A^s(\omega) + \theta_A^r(\omega)} \).

At his given information set \( M \in \{A, B\} \), Receiver (Player 2) plays \( S_1 \) with probability \( \frac{\theta_{S_1}(M)}{\theta_{S_1}(M) + \theta_{S_2}(M)} \) and \( S_2 \) with probability \( \frac{\theta_{S_2}(M)}{\theta_{S_1}(M) + \theta_{S_2}(M)} \).

Updating Rules

For the Sender (Player 1):

\[
\theta_A^{t+1}(\omega) = \theta_A^s(\omega) + 1_{[\pi_1(t)=1]}
\]

\[
\theta_B^{t+1}(\omega) = \theta_B^r(\omega) + 1_{[\pi_1(t)=1]}
\]

For the Receiver (Player 2):

\[
\theta_{S_1}^{t+1}(M) = \theta_{S_1}^s(M) + 1_{[\pi_2(t)=1]}
\]

\[
\theta_{S_2}^{t+1}(M) = \theta_{S_2}^s(M) + 1_{[\pi_2(t)=1]}
\]

where \( 1_E \) is the indicator function.

For this particular game (with perfect coincidence of interests) these updating rules are a conditional form of Roth Erev reinforcement (1996).

Urn Interpretation: Each Player has a Polya-type urn at each information set.
Augmented Signaling Game (after a successful interaction accompanying a draw of the “mutator”)?

In the special case where the states are equiprobable, this signaling game has 68 Nash equilibria, only 6 of which are efficient.

Questions for further consideration
(1) Can we get convergence to signaling systems at a rate that seems closer to the rate at which real humans learn to signal? (Perhaps?)
(2) Is there more than one way to interpret the mutator effect?

A drawback of the specific model?
At least in the 2-player case, we expect that eventually Sender and Receiver will lock into a pattern that mimics a signaling system (possibly after the introduction of sufficiently many signals). But this limiting result might take a very long time to realize.
Can we get there faster?

Note that Sender (Player 1) has his original pair of urns, but Receiver (Player 2) has a new urn for her new information set.
A variation on the Polya Urn Model?

New Updating Rules

For the Sender (Player 1):

\[
\begin{align*}
\theta_{A}^{t+1}(\omega) &= \theta_{A}^{t}(\omega) + r_{1k}1[\sigma_{1}(t) = 1] \\
\theta_{B}^{t+1}(\omega) &= \theta_{B}^{t}(\omega) + r_{1k}1[\sigma_{1}(t) = 1]
\end{align*}
\]

For the Receiver (Player 2):

\[
\begin{align*}
\theta_{S_{1}}^{t+1}(M) &= \theta_{S_{1}}^{t}(M) + r_{2k}1[\sigma_{2}(t) = 1] \\
\theta_{S_{2}}^{t+1}(\omega) &= \theta_{S_{2}}^{t}(\omega) + r_{2k}1[\sigma_{2}(t) = 1]
\end{align*}
\]

where \(1_{E}\) is the indicator function and \(r_{ik}\) is a weight associated with the \(k\)th time this parameter has been reinforced. (Urn interpretation: \(r_{ik}\) is the number of balls added of a particular color the \(k\)th time this ball is drawn and the players have a "success".)

Result for Polya Urns (Klenke 2007): Let \(w_{n}\) be the total number of balls of a given color after \(n\) draws of a ball of this color. If \(\sum_{n=0}^{\infty} \frac{1}{w_{n}} < \infty\), then almost surely eventually only balls of this color will be drawn.

Appropriate versions of this updating rule might converge more rapidly to a pure strategy Nash equilibrium, and possibly a signaling equilibrium. But new signals might emerge even less frequently with this dynamic than they do with Alexander-Skyrms-Zabell.

Looking at the mutator effect again.

Another interpretation

This game is a true signaling game, but it requires new acts at each of Sender’s information sets, not just the set where the “mutator” appeared.

This is possibly more faithful to the urn interpretation, but the generated game is technically not a signaling game.