The conceptual alignment problem is not described effectively in the game theoretic framework. This is because of the representation properties of game theoretic matrices. A conceptually misaligned signaling system, represented game theoretically, is indistinguishable from a perfectly conceptually aligned one. This is because there exists no mechanisms in game theory to coordinate the signal meanings between agents. In a one-shot signaling game, the pay-offs are given and dependent only upon the state of affairs, not the signals. They can be made dependent on the signals by instituting a relation between states of affairs and signals, but this relation is not observable by the

The only mechanism is a meta game theoretic one and involves observation of effects or pay-offs and learning of pay-off structure for future rounds. This is the interesting component of coordination and communication and it is this that we will investigate.

Or maybe not??