Grounding Communication Without Prior Structure

Eric M. Meisner and Selma Šabanović
Department of Computer Science at
Johns Hopkins University, Baltimore MD 21218
School of Informatics and Computing
University of Indiana, Bloomington IN 47408

Abstract

This work describes an approach to time-series modeling of social interactions between human and robot, which is motivated by the social psychology concept of social grounding. In this model, the goal of the agents is to establish and use patterns of communication, rather than rely on existing patterns. Our goal is to allow an artificial agent to construct a pattern of shared meaning with a human or other agent through shared experience rather than relying a model provided a priori. We describe a preliminary human robot interaction study which illustrates the proposed approach.

The flow of time is a situated, context-dependent experience. Face-to-face with a snarling dog, you may feel time slow down to a trickle, seemingly giving you a chance to react. Spending time with friends, it may surprise you to find that the hours have flown by "in no time." Technological artifacts, such as clocks and electricity, have had a rationalizing effect on our representation of the passage of time, turning our situated reactions and diurnal rhythms into mechanically standardized minutes, hours, and days. These technologies have also allowed for the construction of analytical concepts with which to objectively describe and measure the passage of time, such as speed, acceleration, tempo, and delay. In the moment of action, however, our experience of time emerges through our interactions with the environment and the common understanding that we develop with other actors, rather than being perceived as a quantity denoting a particular speed or rate of acceleration.

In the case of social interaction, the fundamental significance of the temporal dimension is evidenced by research showing that rhythmic synchrony is essential to the establishment and success of short- as well as long-term interactions (Condon 1986, Kendon 1990, Trevarthen 2000). Most people masterfully establish temporal entrainment in their encounters with other social actors; their abilities, however, do not rely on a high-level representation of the passage of time within the interaction. In fact, the perception-action times of humans are so short–a few hundred milliseconds–that people do not have the time to consciously consider what they will do next or how and when to do it. We can therefore think of interaction partners as subconsciously embodying the temporal aspects of interaction, which emerge out of their mutually constrained actions and reactions. Entrainment could in some sense even be seen as allowing interaction participants to be one step ahead in time, as the rhythmic flow of their movements anticipates the next beat in the interaction. Through the combination of entrainment and contextual situatedness, a blink of an eye can become a meaningful wink.

Social interaction is composed of "countless patterns and natural sequences of behavior occurring whenever persons come into one another's presence" (Goffman 1967). Learning to perceive, predict, and evoke particular patterns and responses from others is an important part of social learning. If our experience of time is emergent and situated, rather than objectively predefined, how then can robots and other machines driven by a more discrete representation of time participate in the adaptive dance of social interaction? If we approach social intelligence as situated action, we shift the focus of our analysis and modeling from knowledge stored within the artifact to knowledge as constructed capability-in-action, emerging through physical and social performances(Clancey 1997). Social engagement triggers an embodied, situated system that is sensitive to recognizing socially relevant patterns in our everyday behavior, such as interaction rhythms, imitation, and joint attention. This allows us to predict what others will do and coordinate our behaviors. As individuals respond dynamically to the actions of others, behavior is used to regulate one's own state and the behavior of other individuals and enables the attunement of intentions (or what we are calling here a common understanding) among interaction partners (Barrett 2007).

Common ground in humans is based on certain shared characteristics–the sharing of physical space, of a certain type of embodiment and its constraints, and the assumption that the other is a thinking, social being. Even in cases where language and other cultural forms of establishing common ground (such as particular symbols, gestures, or combinations thereof) are not previously shared by interaction partners, two people can come to some agreement through mutual trial and error and develop their own communicative shorthand. Two people who do not speak the same language are capable of creating a mutual understanding, and to then use this understanding as a channel of communication.

When trying to computationally describe what is going on in human-human interaction and design robots that can participate in it, the natural approach therefore seems to be a decision-theoretic one that models a dynamical system that is not completely observable, such as a partially observable Markov decision process (POMDPs) or other types of dynamic Bayes networks. The descriptive power of these frameworks is attractive for real-world robot tasks because it enables agents to reason about the result of future actions by interacting with the world and receiving feedback using imperfect perception. Typically, POMDPs have proved useful in cases where the state and observation spaces of a system can be described in very few terms (Hsiao, Kaelbling, & Lozano-Perez 2007).

POMDPs, along with much of the computational machinery used in modeling time-series decision and estimation problems, closely resemble the 'sense-think-act' model of cognitive architectures. Conversely, Semin (Semin 2007) suggests that language and other high level aspects of communication are based on synchronization or parity of behaviors. The processes that generate these behaviors are non-cognitive and more strongly connected to physical experience than to high-level reasoning. This is incompatible with sense-think-act learning models, and suggests that we need to use models that more immediately couple actions and observations.

A different approach to enabling human-robot interaction is suggested by "emergent" theories of social cognition, which say that understanding, meaning, and rules of social interaction are not a predefined property of the world, but rather something that is agreed upon in interaction. For the specific problem of social interaction, the structure and patterns are developed through parity and closed loop feedback. This is a process that requires both (or all) agents to adjust their behavior and understanding based on the behavior of other agents. The result and structure of this process is not well defined and this is what seriously limits the utility of using sequential decision/control problem formulations in this domain. In order to make social agents, our aim should be to instrument them with the ability to form patterns of interaction with other agents, rather than simply adjust a predefined model of communication.

We propose an approach that allows an agent to construct shared understanding with another agent by processing perception and action information, rather than using an existing model to assign specific meaning to actions and perceptions. In our proposed approach, an agent uses a predictive state representation (PSRs, (Littman, Sutton, & Singh 2002; McCracken & Bowling 2005)) to identify important sequences in the exchange of behaviors with a human. Ihe agent seeks to identify patterns in the stream of information that help predict future actions and observations. Determining these patterns can be thought of as learning the modes of interaction. These patterns can be of arbitrary length. The actions of the agent are driven by the classic trial and error approach of reinforcement learning, however the reward for the agent comes in the form of correctly predicting the responses of other agents to its actions. The notion of a context or mode is represented only as a joint sequence of agent actions that have no explicit meaning other than the resulting sequence itself. This is in contrast to the dynamic Bayes approach, which starts with defined causal, temporal relationships between modes and uses these relationships to learn and adapt.

A PSR can be described briefly as the following:

- A partitioning of actions and observations into sets $a \in A$ and $o \in O$.
- A set of test sequences $q \in Q$ composed of alternating action-observation pairs.
- A system dynamics matrix, $Pr(Q|H)$, in which each column represents a test sequence $q \in Q$ and each row $h \in H$ represents an instant of time. Each entry in this matrix $Pr(q|h)$ represents the probability that q will occur starting at time h.

The goal of this representation is the following. First, using frequency probability, correctly model $P(q|h)$ for all q. Second, determine a minimal set of test sequences $\bar{Q}$ that are needed to predict all other q. This means that for any q there exists $f_q(Pr(\bar{Q}|h)) == Pr(q)$. The second problem is addressed by analysis of the system dynamics matrix $Pr(Q|H)$, in particular by finding the linearly independent columns of $Pr(Q|H)$. One may think of this as a process in which the agent allows itself to be imprinted by experience, and then constructs meaning by examining and analyzing its own experience.

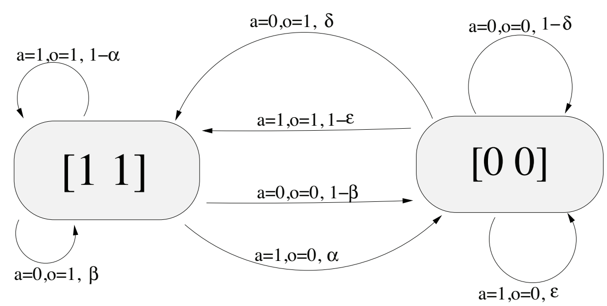

Figure 1: A 2-state POMDP describing the pattern of interaction with parameters $0 < \alpha, \beta, \delta, \epsilon \ll 1$. The agent can remain in either of the two states [1,1] and [0,0], with high probability, by selecting 1 or 0 respectively. It can cause a transition out of [1,1] or [0,0], with high probability, by selecting 0 or 1 respectively.

In the context of reinforcement learning, the difference between this and classic approaches is subtle. This method does not promise any more mathematically efficient or effective approach to classical reinforcement learning problems. However, in the context of social dynamics, where rules are established in process rather than ahead of time, this method provides a way for the agent to create its own useful abstractions rather than have them pre-specified by an expert. This is an important step towards the type of flexibility required for social intelligence.

We present a proof of concept system, in which a human is asked to evoke a particular response from a robot that has no prior understanding of context, task or the meaning of

Algorithm 1 Exploration Schedule PSR Learning Algorithm(set <core tests> CT)

```
Initialize exploration schedule α = 1.0
while Change in One-step Prediction error is large do
    If( RAND(0:1) ≤ α ) EXPLORATION move
    Else EXPLOITATION move
    decay(α)
end while
DISCOVERY
Re-select core tests
loop
    If( RAND(0:1) ≤ α ) EXPLORATION move
    Else EXPLOITATION move
    decay(α)
end loop
```

actions. To start with, neither the human nor the robot have any notion of how the other will behave. The robot is driven by novelty of actions and responses as well as the the desire to correctly predict the results of its actions. We base our approach on the notion that in learning communication from the ground up, it makes sense for an agent to 1) predict how other agents will respond to its actions, and 2) show its understanding by taking actions that will result in predictable responses. This thinking leads to the reinforcement learning algorithm in 1.

We design our experiment by first analyzing a specific and important part of interaction: imitation. We construct a POMDP representation of the observed pattern of imitation between two people engaging in open-ended play (figure 1). We start with the notion that a robot would not have access to such a model and ask the question "Can an agent, using the controller learning algorithm, build a PSR that encodes the same understanding about the interaction as its human interaction partner?"

In the experiment, the human participant is given the instructions to "Get the robot to imitate/not imitate you", but no instruction as to how. The robot is provided with the ability to differentiate between the human's actions and the ability to perform the same actions. The robot observes when it is doing the same thing as the human and when it is not. We show that over time, the robot acts in accordance with the human's desire. To confirm this, we examine the encoded representations (the contents of $Pr(Q|H)$ and $\bar{Q}$) of the robot and show that it matches the goal of the human, thus resulting in a shared representation. In doing so, we have demonstrated the potential for artificial agents to learn and use interaction patterns with other agents, without being provided with a predefined time-series model. There are two possible directions for further work in this area. The first is to make the system deal more directly with raw data, rather than being provided with gestural primitives. The second, is to build towards more complex interactions.

References

Barrett, L. 2007. Too much monkey business. In *Expert Meeting Grounding Sociality: Neurons, Minds, and Culture*.

Clancey, W. J. 1997. *Situated Cognition: On Human Knowledge and Computer Representations*. Cambridge University Press.

Goffman, E. 1967. *Interaction Ritual*. New York: Pantheon.

Hsiao, K.; Kaelbling, L.; and Lozano-Perez, T. 2007. Grasping pomdps. In *IEEE International Conference on Robotics and Automation*, 4685–4692.

Littman, M. L.; Sutton, R. S.; and Singh, S. 2002. Predictive representations of state. In *In Advances In Neural Information Processing Systems 14*, 1555–1561. MIT Press.

McCracken, P., and Bowling, M. H. 2005. Online discovery and learning of predictive state representations. In *NIPS*.

Semin, G. R. 2007. Grounding communication: Synchrony. In *Social Psychology: A Handbook of Basic Principles*. New York: Guildford Publications. 630–649.