Interactively Learning Nonverbal Behavior for Inference and Production: A Machine Learning Approach

Jin Joo Lee
MIT Media Lab
Cambridge, MA 02142 USA
jinjoo@media.mit.edu

Cynthia Breazeal
MIT Media Lab
Cambridge, MA 02142 USA
cynthiab@media.mit.edu

***Abstract*—By designing socially intelligent robots that can more effectively communicate and interact with us, we can increase their capacity to function as collaborative partners. Our research goal is to develop robots capable of engaging in nonverbal communication, which has been argued to be at the core of social intelligence. We take a human-centric approach that closely aligns with how people are theorized to model nonverbal communication. We propose a unified computational approach to interactively learn the meaning of nonverbal behaviors for inference and production. More specifically, we use a partially observable Markov decision process model to infer an interactant's mental state based on their observed nonverbal behaviors as well as produce appropriate nonverbal behaviors for a robot to communicate its internal state. By interactively learning the connection between nonverbal behaviors and the mental states producing them, an agent can more readily generalize to new people and situations.**

I. Introduction

Robots have an immense potential to help people in domains such as education, healthcare, manufacturing, and disaster response. For instance, researchers have designed robots that take steps towards helping children learn a second language [1], assisting nurses with triage [2], and participating as part of a search and rescue team [3]. As such robots begin to collaborate with us, we believe the successful outcome of the human-robot team depends on the robot's ability to effectively communicate and socially interact with everyday people. The challenge thereby lies in imbuing robots with social intelligence, which involves the ability to both express and infer motives, intentions, and emotions to/from others.

Our research goal is to develop socially intelligent robots capable of engaging in nonverbal communication, which has been argued to be at the core of social intelligence [4]. We present a unified computational approach to interactively learn the *meaning* of nonverbal behaviors for inference and production. To understand how a robot could produce *meaningful* nonverbal behaviors, we begin with an investigation into literature with a focused attention to current state-of-the art approaches and theories on nonverbal production (Section II: Background). From these works, we gain insights into the underlying representation and design constraints when computationally modeling nonverbal communication (Section III: Model Representation). We detail our proposed approach in using a partially observable Markov decision process model to infer mental states based on observed nonverbal behaviors as well as to produce nonverbal behaviors based on a given mental state (Section IV: Proposed Work). Lastly, we conclude by comparing our approach to prior approaches to establish the novelty of our proposal (Section V: Related works).

(a) **Nonverbal Inference**: behaviors → mental state

(b) **Nonverbal Production**: mental state → behaviors

Fig. 1: Nonverbal communication includes the ability to both express and infer motives, intentions, and emotions (mental states) to and from others using nonverbal behavior.

II. Background

To understand how an agent (robotic or human) could produce *meaningful* nonverbal behaviors, we investigate prior approaches and theories on nonverbal production from diverse fields such as Embodied Conversational Agents, Social Signal Processing, and Social Psychology.

A. Embodied Conversational Agents

The field of Embodied Conversational Agents (ECA) has predominately lead the research on nonverbal production for virtual agents. Researchers working on ECAs focus on how a virtual agent can interact with humans using both language and nonverbal behavior. Frameworks, such as the Behavior Expression Animation Toolkit, enable animators to translate from input-text (i.e., a sentence to be spoken by a virtual agent) to synchronized nonverbal behaviors and synthesized speech [5]. Based on the input-text's linguistic and contextual analysis, the toolkit determines the appropriate behavior rules to apply when generating the co-verbal gestures. For example, a behavior rule would suggest a gaze-away behavior at the beginning of an utterance 70% of the time to establish conversational turn-taking.

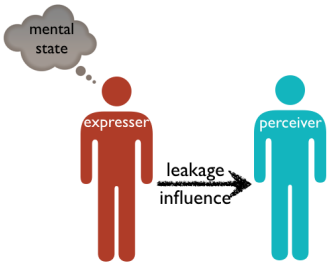
The ECA community has moved to formalize a generic behavior generation framework, which defines a Function Markup Language (FML) and a Behavior Markup Language (BML) [6], [7]. The FML describes an agent's intent or internal state (e.g., the agent is confused or wants to take a turn speaking), while the BML describes the physical behavior the agent should perform (e.g., eyebrows up or head-shake). This generic framework leaves open the possibility to design a module to transform the high-level FML to the low-level BML using planning algorithms, rule-based systems, etc. But many frameworks use hand-crafted behavior rules for this mapping [8], [9], [10].

B. Social Signal Processing

The Social Signal Processing (SSP) community, in comparison to the ECA community, has a more data-driven approach to behavior generation [11], [12]. Morency et. al. built probabilistic models that can learn from human-to-human interactions to predict when a listener should nod in a dyadic conversation [13]. Based on the speaker's nonverbal behaviors (e.g., their gaze behavior or pauses in their speech), the model learns the probability of a listener's nod occurring. Without attending to the literal spoken contents, the model is automatically learning associations between a speaker's nonverbal behaviors and a listener's nodding behavior. Once the prediction model is trained on the human-human interaction data, it is used to drive the nodding behavior of a virtual listening agent.

C. Social Psychology

From the field of Social Psychology, classical theories of human nonverbal communication fall between two major perspectives. The theory of *nonverbal leakage* views mental states as direct influencers of nonverbal behaviors. These nonverbal behaviors, which are thought to occur largely outside of people's conscious control, reveal or "leak" information about a person's underlying intentions or emotions [4], [14]. According to this perspective, a frown is a direct "hard-wired" production of a person being sad. The theory of *nonverbal influence* views nonverbal behaviors as expressive signals that communicate or influence others. These nonverbal behaviors primarily function to affect a perceiver. According to this perspective, a frown is produced to communicate sadness or to elicit a hug from a perceiver [15]. Although different in



mental state → nonverbal behaviors

Fig. 2: Theories of nonverbal leakage and influence share a common representation in which nonverbal behaviors, whether intentional or unintentional, are manifestations of an implicit or explicit mental state.

their definitions, these two theories have similar underlying representations. As illustrated in Figure 2, both theories have in common that nonverbal behaviors, whether intentional or unintentional, are produced from some given implicit or explicit mental state.

Modern theories of nonverbal production emphasize the contextual nature of nonverbal behaviors. Rather than assuming an one-to-one correspondence between a specific behavior and its underlying meaning (like a smile always meaning happiness), the meaning of nonverbal behaviors is dependent on the context in which they are embedded [16]. For example, an eye-roll in conjunction with a large grin can be more accurately interpreted as conveying humor, whereas observing an eye-roll in isolation could lead to the less accurate interpretation of contempt. The interpretation of the eye-roll is contingent upon the observation of the grin. Beyond temporal dependence, other contextual factors including setting, culture, gender, and personality affects how an individual encodes and decodes nonverbal behavior [17].

III. Model Representation

Given what we have learned from classical theories, modern theories, and prior frameworks of nonverbal communication, we define design parameters to constraint a computational model representation. We then detail our proposed nonverbal communication framework based on partially observable Markov decision process models, which satisfy the specified design parameters.

A. Design Parameters

By investigating prior works in various research areas, we found three different perspectives on how an agent (robotic or human) produces nonverbal behaviors. The embodied conversational agents community has focused on producing nonverbal behaviors to accompany spoken text through predefined rules. We view ECA's approach as a function mapping from input text to nonverbal behaviors (as shown in Equation 1a). Alternatively, the social signal processing community uses data-driven methods to generate nonverbal behaviors for a

virtual agent in response to human behavior. We view SSP's approach as a function mapping from human behavior to agent behavior (Equation 1b). Lastly, according to theories of human nonverbal communication from the field of social psychology, nonverbal behaviors are produced from some given implicit or explicit mental state (Equation 1c).

$$f : \text{input text} \mapsto \text{nonverbal behaviors} \tag{1a}$$
$$f : \text{human behaviors} \mapsto \text{agent behaviors} \tag{1b}$$
$$f : \text{mental state} \mapsto \text{nonverbal behaviors} \tag{1c}$$

Prior approaches have been successful in producing appropriate nonverbal behaviors in one-shot interactions in well-defined contexts, which systems built using pre-defined rules and data-driven responses are capable of supporting. However, when considering long-term interactions in various contexts, these approaches are limited in their ability to generalize to new and different situations unless more rules are created or more data is collected. We therefore want to take a more human-centric approach that closely aligns with how humans are theorized to model nonverbal communication in anticipation of greater generalizability when a system's underlying representation is similar to that of people. This leads us to our first design parameter for our computational nonverbal communication model:

> **Design Parameter 1**: the model representation should follow that nonverbal behaviors are manifestations of mental states.

Since the encoding and decoding of nonverbal behaviors change depending on context, a system capable of interactively learning from its environment can more readily adapt to new people, situations, cultures, etc. As such, our second design parameter includes:

> **Design Parameter 2**: the model should be capable of interactively learning from everyday people

Research in nonverbal inference and production has traditionally been treated as two separate problems. Researchers working in domains such as social signal processing have focused on interpreting nonverbal behaviors, while researchers in domains such as embodied conversational agents have focused on developing behavioral planners. By addressing both inference and production in a unified model, we constrain the model to share the same mechanism when encoding and decoding nonverbal behaviors, which ensures that nonverbal behaviors retain their meaning in both processes. This leads to our final design parameter:

> **Design Parameter 3**: the model should support both inference and production of nonverbal behaviors

Given these three design parameters, we propose to develop a *unified* computational approach to *interactively learn* the connection between *nonverbal behaviors and mental states*.

B. Nonverbal Communication Framework

We propose a partially observable Markov decision process (POMDP) framework that is capable of inferring an interactant's mental state based on their nonverbal behaviors as well as produce appropriate nonverbal behaviors for a robot to communicate its internal state. We take a Bayesian reinforcement learning (RL) approach to interactively learn the probability distribution of observing particular nonverbal behaviors given a person's mental state (i.e., the observation probabilities). By relying on an interactant that provides optimal policy information, the system can gradually learn the observation probabilities through experiences of execution and learning. With more interaction episodes, we can eventually converge on an observation model that can generate policies that is similar to the interactant's, which will enable the robot to make decisions that demonstrate an accurate inference of the interactant's current mental state. Our system then reuses this POMDP model learned for inference to also solve the problem of nonverbal production, which we reframe as an "inversion" of the inference problem (as inspired by Smith and Lieberman's work in natural language generation [18]). This novel solution to nonverbal production requires minimal modification to the inference model and is also solved using inference techniques.

IV. Proposed Model

We describe our proposed nonverbal communication framework through a simplified problem scenario. We define this problem scenario as a POMDP model and detail the setup of the model parameters. Using pre-existing methods as well as our proposed novel method, we demonstrate how nonverbal inference and production can be addressed in a unified computational model.

A. Problem Illustration

We illustrate our research problem with the following simplified scenario (also see Figure 1a):

> *Sarah sees a toy and starts to smile. She tells a story of seeing something similar on TV in a happy tone of voice. She excitedly fidgets and leans forward to get closer to the toy. Sarah really wants this toy.*

B. POMDP Setup

Framing this problem as a POMDP model, Sarah's mental state of either wanting the toy or not wanting the toy S: $\{S_w, S_{\overline{w}}\}$ is partially observed through her nonverbal expressions such as smiling, happy tone of voice, leaning forward, and fidgeting. The robot maintains a belief state, which is characterized as a probability distribution over the world states, which includes the probability that Sarah wants the toy and the probability that Sarah does not want the toy.

$$b(s) = \begin{Bmatrix} p_1 = b(S_w) \\ p_2 = b(S_{\overline{w}}) \end{Bmatrix}$$

The robot can choose among 3 different actions. The robot can decide to $give$ the toy to Sarah, $take$ the toy away, or $watch$ to gain more information about Sarah's mental state by observing her nonverbal behaviors A: $\{watch, give, take\}$. The robot's goal in this scenario is encoded by its reward function. A positive reward (e.g., +100) is given when either the following goal configurations are obtained a) Sarah has a mental state of wanting the toy and the robot $gives$ the toy

$r : S \times A \rightarrow \Re$	
$r(S_w, give) = +100$	$r(S_{\overline{w}}, give) = -100$
$r(S_w, take) = -100$	$r(S_{\overline{w}}, take) = +100$
$r(S_w, watch) = -1$	$r(S_{\overline{w}}, watch) = -1$

TABLE I: Reward probabilities

$T : S \times A \rightarrow \Pi(S)$		
$p(S_w \| S_w, give) = 0.5$	$p(S_w \| S_w, take) = 0.5$	$p(S_w \| S_w, watch) = 1.0$
$p(S_{\overline{w}} \| S_w, give) = 0.5$	$p(S_{\overline{w}} \| S_w, take) = 0.5$	$p(S_{\overline{w}} \| S_w, watch) = 0.0$
$p(S_w \| S_{\overline{w}}, give) = 0.5$	$p(S_w \| S_{\overline{w}}, take) = 0.5$	$p(S_w \| S_{\overline{w}}, watch) = 0.0$
$p(S_{\overline{w}} \| S_{\overline{w}}, give) = 0.5$	$p(S_{\overline{w}} \| S_{\overline{w}}, take) = 0.5$	$p(S_{\overline{w}} \| S_{\overline{w}}, watch) = 1.0$

TABLE II: Transition probabilities

$O : S \times A \rightarrow \Pi(\Omega)$		
$p(smile \| S_w, give) = 0.5$	$p(smile \| S_w, take) = 0.5$	$p(smile \| S_w, watch) = 0.9$
$p(frown \| S_w, give) = 0.5$	$p(frown \| S_w, take) = 0.5$	$p(frown \| S_w, watch) = 0.1$
$p(smile \| S_{\overline{w}}, give) = 0.5$	$p(smile \| S_{\overline{w}}, take) = 0.5$	$p(smile \| S_{\overline{w}}, watch) = 0.1$
$p(frown \| S_{\overline{w}}, give) = 0.5$	$p(frown \| S_{\overline{w}}, take) = 0.5$	$p(frown \| S_{\overline{w}}, watch) = 0.9$

TABLE III: Observation probabilities

b) Sarah has a mental state of not wanting the toy and the robot $takes$ the toy. A negative reward (e.g., -100) is given when the opposite goal configuration is obtained. A small cost (e.g., -1) is given for a $watch$ action (see Table I).

The world's state dynamics are encoded through the state-transition function, which defines the probability of ending in state s' given that the agent starts in state s and takes action a. For our particular scenario, the $watch$ action does not change the state of the world in that Sarah's wanting of a toy does not change due to the robot's $watch$ action. The $give$ and $take$ terminating actions cause a transition to the world state of S_w with probability 0.5 and state $S_{\overline{w}}$ with probability 0.5, which essentially resets the problem (see Table II).

To simplify our problem, let us assume that there are only two possible observations O:$\{smile, frown\}$. Through a $watch$ action, the robot can observe the nonverbal behaviors exhibited by Sarah to infer her underlying mental state. Given that the robot took action a and Sarah's mental state is s, the observation function defines the probability of making observation o. Under the assumption that our POMDP model parameters are fully known, let us set the observation probabilities to be "correct estimates," or ones that closely resemble Sarah's. That is, if Sarah's mental state is S_w (i.e., wanting the toy), then there is a higher probability of observing her $smile$ over a $frown$. Similarly, if Sarah has a state $S_{\overline{w}}$ (i.e., not wanting the toy), then there is a higher probability of observing her $frown$. These observations are only gained from an information seeking action like the $watch$ action; other actions like $give$ and $take$ do not yield any helpful observations (see Table III).

In our scenario, the observation function defines the probability distribution of observing various nonverbal behaviors given a person's mental state of either wanting or not wanting the toy. But how does this function become defined? For this scenario, how does a robot come to understand that observing Sarah smile has the underlying meaning of her wanting the toy? For nonverbal inference, the primary goal is to learn the observation probabilities for this POMDP model.

C. Unified Model

We propose a unified model that shares the same underlying mechanism when decoding and encoding nonverbal behaviors by using a pre-existing method called *Bayesian reinforcement learning* as well as our proposed novel method we call *belief-shaping through perspective-taking*. By learning the observation probabilities, the POMDP model is capable of creating policies that will enable the robot to make appropriate decisions and take actions that demonstrate an accurate inference of the interactant's current mental state. Our system reuses this inference model to also solve the problem of nonverbal production, which we reframe as an "inversion" of the inference problem. Although our approach is unified, we will describe the details in two phases:

1) **Nonverbal Inference** (behaviors $\rightarrow$ mental state)
2) **Nonverbal Production** (mental state $\rightarrow$ behaviors)

D. Nonverbal Inference

We take a Bayesian reinforcement learning (RL) approach to interactively learn the observation probabilities of our POMDP model. Following the framework described by Atrash and Pineau [20], a robot can gradually learn the observation model through experiences with an interactant (i.e., an oracle) that provides optimal policy information. The Bayesian RL approach begins with some prior probability distribution over all possible observation models. This prior can have some helpful initial estimates based on domain knowledge, or if no information is given then it begins as an uniform prior. From this prior probability distribution (generally in Dirichlet form), we sample a set of observation models to create a variety of different POMDP models. Each model's optimal policy is computed using a solver such as value iteration or policy iteration. The robot executes actions either based on a randomly selected policy or by selecting actions with the greatest expected value over all computed policies. If an incorrect action is executed, the oracle is expected to correct the robot's actions by demonstrating the action that should have been executed given the world state. For each of the

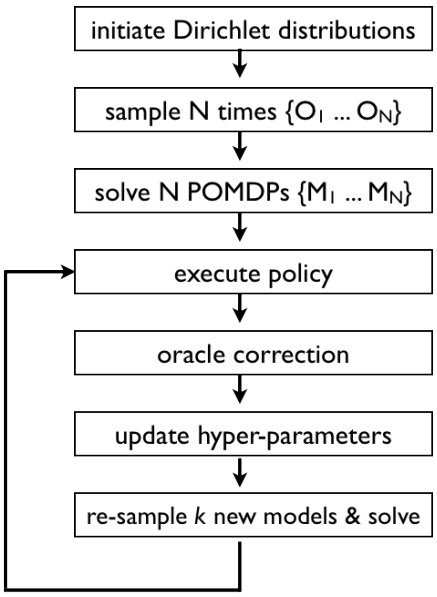


Fig. 3: High-level step-by-step process of the Bayesian reinforcement learning method described by [19], [20].

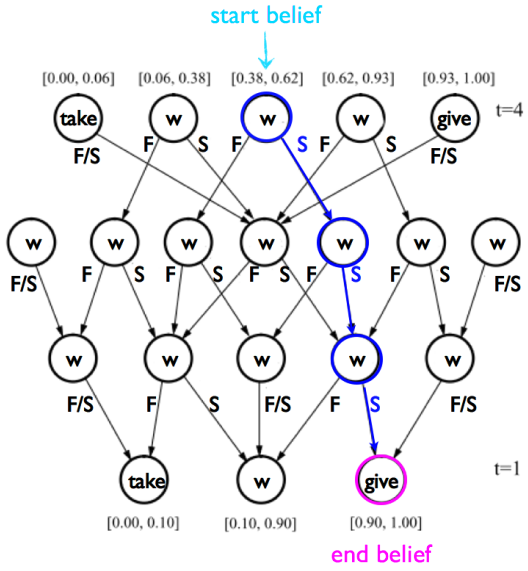


Fig. 4: An example of an optimal policy diagram for our problem scenario (original image courtesy of [21]). The highlighted path is the shortest path from a particular starting belief to a target belief given that three watch actions were taken by the interactant. At each step, the robot should produce a *smile* (S) behavior for the interactant to observe in order to shape or manipulate their belief toward a target end belief.

policies that agreed with the oracle's demonstrated action, their observation models are used to update the hyperparameters of the Dirichlet distribution. With more and more examples, this maintained posterior distribution over the possible observation models gets tightened up and moves in the direction of models that can generate similar policies to that of the oracle. With this updated posterior distribution, we resample some new models and remove some old models. Through iterations of execution and learning, we can eventually converge on an observation model that enables the POMDP framework to produce policies that is similar to the interactant's (see Figure 3 for a diagram of the Bayesian RL process).

E. Nonverbal Production

We take inspiration from Gray and Breazeal [22], [23] and view nonverbal communication as a process of *mental state manipulation* through physical actions. Given the constraints of the world, we cannot directly communicate our mental state to another person (i.e., telepathy as illustrated in Figure 5a). Consequently, we communicate indirectly to others using the world as our medium. To attempt to cause a mental state change in another agent, we act in the physical world through mechanisms such as speech and nonverbal behaviors. Since mental states are only changed by the agent itself, we have to act on and alter the world so that the agent's perceptual and mental processes will bring about a targeted change (see Figure 5b). As such, we pose the problem of nonverbal production as "acting" the appropriate behaviors such that an observer can correctly infer the actor's mental state. Our approach is to use the learned inference model to also solve the problem of nonverbal production by reframing it an inference problem:

> **nonverbal production** is to act nonverbal behaviors such that the observer forms the desired mental state inference.

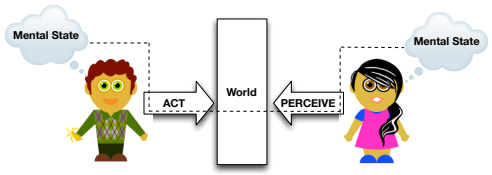
For our toy scenario, the question then becomes "what observations (i.e., nonverbal behaviors) should a robot produce for Sarah so that she correctly infers the robot's mental state?" Assuming that the robot has learned a similar policy to that of Sarah, the robot in essence asks "what would I think if I were Sarah?" The robot maintains a distribution of Sarah's beliefs about the robot's belief. This distribution ranges from Sarah believing that the robot does not want the toy to Sarah believing that the robot wants the toy. If the robot has an internal state of wanting the toy, then this is the target belief the robot wants Sarah to have. When the robot sees that Sarah is taking a *watch* action, the question now becomes, "what behavior does Sarah need to observe in order to change her belief towards the target belief?" Based on the learned inference model, the robot follows its optimal policy to act/produce observations that would guide Sarah's belief to the target belief (see Figure 4 for an illustrated example).

We describe this novel process as *belief-shaping through perspective-taking*, where the learned inference model is projected to Sarah's perspective to determine what Sarah needs to observe in order to shape or manipulate her beliefs.

V. Related Works

Similar to our proposed problem space and approach, prior work by Hoey and Little [24], [25] focused on understanding the meaning, or the affordances, of nonverbal behaviors using POMDPs. By learning the relationship between observed human nonverbal behaviors and robot actions, their system can learn that a person's forward hand-gesture captured in video images means that the robot should move forward. Much

(a) direct communication



(b) indirect communication

Fig. 5: Telepathy allows for direct communication from one's mental state to another's. Given our world's constraints, we act in the physical world through mechanisms such as speech and nonverbal behaviors, which are perceived by other agents.

of the complexity of this work comes from automatically discovering and learning the salient nonverbal behaviors from camera images. They trade off learning more everyday and rich interaction behaviors for obvious nonverbal behaviors, which are more perceptually recognizable. Furthermore, our approach differs in its explicit representation of mental states as the mediating factor between observed human nonverbal behaviors and robot actions. Atrash and Pineau [20] similarly assume an one-to-one mapping from human behavior to robot action. As mentioned previously, we use their framework to interactively learn the observation probabilities of the POMDP model. Their framework gradually learns the relationship between a user's spoken intentions and robot actions. In their demonstration, an autonomous wheelchair robot interactively learns that when a user says "move forward" the robot should drive straight for one meter.

Both of these related works focus on the inference problem of interpreting human behavior. Our more holistic approach addresses both inference and production in a single model to ensure that nonverbal behaviors retain their underlying meaning in both processes.

Acknowledgment

We acknowledge the support of National Science Foundation Research Grants CCF-1138986.

References

[1] T. Kanda, T. Hirano, D. Eaton, and H. Ishiguro, "Interactive Robots as Social Partners and Peer Tutors for Children: A Field Trial," *Human-Computer Interaction*, vol. 19, no. 1, pp. 61–84, Jun. 2004.

[2] D. M. Wilkes, S. Franklin, E. Erdemir, S. Gordon, S. Strain, K. Miller, and K. Kawamura, "Heterogeneous Artificial Agents for Triage Nurse Assistance," in *International Conference on Humanoid Robots*, 2010, pp. 130–137.

[3] M. F. Jung, J. J. Lee, N. Depalma, S. O. Adalgeirsson, P. J. Hinds, and C. Breazeal, "Engaging Robots: Easing Complex Human-Robot Teamwork using Backchanneling," in *Proceedings of the Conference of Computer Supported Cooperative Work*, 2013, pp. 1555–1566.

[4] N. Ambady and M. Weisbuch, *Nonverbal Behavior. Handbook of Social Psychology.* John Wiley & Sons, Inc., 2010.

[5] J. Cassell, H. Vilhjálmsson, and T. Bickmore, "BEAT: the Behavior Expression Animation Toolkit," *Computer Graphics and Interactive Techniques*, pp. 477–486, Aug. 2001.

[6] S. Kopp, B. Krenn, S. Marsella, A. Marshall, C. Pelachaud, H. Pirker, K. Thórisson, and H. Vilhjálmsson, "Towards a Common Framework for Multimodal Generation: The Behavior Markup Language," *Intelligent Virtual Agents*, pp. 205–217, 2006.

[7] K. Thórisson and H. Vilhjalmsson, "Functional Description of Multimodal Acts: A Proposal," in *Proceedings of the 2nd Function Markup Language Workshop*, 2009.

[8] J. Lee and S. Marsella, "Nonverbal Behavior Generator for Embodied Conversational Agents," *Intelligent Virtual Agents*, pp. 243–255, 2006.

[9] L. Huang, L. Morency, and J. Gratch, "Virtual Rapport 2.0," *Intelligent Virtual Agents*, pp. 1–13, 2011.

[10] M. Schroder, E. Bevacqua, R. Cowie, F. Eyben, H. Gunes, D. Heylen, M. ter Maat, G. McKeown, S. Pammi, M. Pantic, C. Pelachaud, B. Schuller, E. de Sevin, M. Valstar, and M. Wollmer, "Building Autonomous Sensitive Artifical Listeners," *IEEE Transactions on Affective Computing*, vol. 3, no. 2, pp. 165–183, 2012.

[11] L.-P. Morency, "Modeling Human Communication Dynamics," *IEEE Signal Processing Magazine*, vol. 27, no. 5, pp. 112 –116, 2010.

[12] A. Vinciarelli, M. Pantic, D. Heylen, C. Pelachaud, I. Poggi, F. D'Errico, and M. Schroder, "Bridging the Gap Between Social Animal and Unsocial Machine: A Survey of Social Signal Processing," *IEEE Transactions on Affective Computing*, pp. 1–20, 2012.

[13] L.-P. Morency, I. Kok, and J. Gratch, "A probabilistic multimodal approach for predicting listener backchannels," *Autonomous Agents and Multi-Agent Systems*, vol. 20, no. 1, pp. 70–84, May 2010.

[14] M. L. Knapp and J. A. Hall, *Nonverbal Communication in Human Interaction*. Cengage Learning, 2010.

[15] A. J. Fridlund, *Human Facial Expression: An Evolutionary View*. Academic Press, 1994.

[16] L. F. Barrett, B. Mesquita, and M. Gendron, "Context in Emotion Perception," *Current Directions in Psychological Science*, vol. 20, no. 5, pp. 286–290, Oct. 2011.

[17] L. F. Barrett, "Emotions Are Real." *Emotion*, vol. 12, no. 3, pp. 413–29, Jun. 2012.

[18] D. A. Smith and H. Lieberman, "Generating and Interpreting Referring Expressions as Belief State Planning and Plan Recognition," *European Workshop on Natural Language Generation*, Aug. 2013.

[19] R. Jaulmes, J. Pineau, and D. Precup, "Active Learning in Partially Observable Markov Decision Processes," *European Conference on Machine Learning*, vol. 3720, pp. 601–608, Oct. 2005.

[20] A. Atrash and J. Pineau, "A Bayesian Method for Learning POMDP Observation Parameters for Robot Interaction Management Systems," *The POMDP Practitioners Workshop*, 2010.

[21] L. P. Kaelbling, M. Littman, and A. R. Cassandra, "Planning and Acting in Partially Observable Stochastic Domains," *Artificial Intelligence*, vol. 101, no. 1-2, pp. 99–134, 1998.

[22] J. Gray, "Goal and Action Inference for Helpful Robots Using Self as Simulator," Doctoral Thesis, Massachusetts Institute of Technology, 2004.

[23] J. Gray and C. Breazeal, "Manipulating Mental States through Physical Action," in *International Conference on Social Robotics*, vol. 7621, 2012, pp. 1–14.

[24] J. Hoey and J. J. Little, "Value-directed human behavior analysis from video using partially observable Markov decision processes." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 7, pp. 1118–1132, Jul. 2007.

[25] J. Hoey, "Decision Theoretic Learning of Human Facial Displays and Gestures," Doctoral Thesis, The University of British Columbia, 2003.