Bayesian Affect Control Theory of Self

Jesse Hoey
David R. Cheriton School of Computer Science
University of Waterloo, 200 University Ave. West,
Waterloo, Ontario, N2L3G1, Canada
jhoey@cs.uwaterloo.ca

Tobias Schröder
Potsdam University of Applied Sciences,
Kiepenheuerallee 5,
14469 Potsdam, Germany
post@tobiasschroeder.de

Abstract

Notions of *identity* and of the *self* have long been studied in social psychology and sociology as key guiding elements of social interaction and coordination. In the AI of the future, these notions will also play a role in producing natural, socially appropriate artificially intelligent agents that encompass subtle and complex human social and affective skills. We propose here a Bayesian generalization of the sociological *affect control theory of self* as a theoretical foundation for socio-affectively skilled artificial agents. This theory posits that each human maintains an internal model of his or her deep sense of "self" that captures their emotional, psychological, and socio-cultural sense of being in the world. The "self" is then externalised as an identity within any given interpersonal and institutional situation, and this situational identity is the person's local (in space and time) representation of the self. Situational identities govern the actions of humans according to affect control theory. Humans will seek situations that allow them to enact identities consistent with their sense of self. This consistency is cumulative over time: if some parts of a person's self are not actualized regularly, the person will have a growing feeling of inauthenticity that they will seek to resolve. In our present generalisation, the self is represented as a probability distribution, allowing it to be multi-modal (a person can maintain multiple different identities), uncertain (a person can be unsure about who they really are), and learnable (agents can learn the identities and selves of other agents). We show how the Bayesian affect control theory of self can underpin artificial agents that are socially intelligent.

Introduction

Designers of intelligent agents are increasingly looking towards the social sciences for ideas on how to create more natural and aligned social interactions. This endeavour requires mathematical models of cognition, but also of emotion or affect, in ways that capture the subtle cultural rules underlying human coordination and cooperation. Affect control theory (ACT) (Heise 2007) is a mathematically formalized sociological theory of the interplays between cultural representations, interactants' identities[1], and affective experience (Heise 2007). ACT posits that humans will strive to achieve consistency in shared affective cultural sentiments about events, and will seek to increase *alignment* with other agents (including artificial ones).

While ACT models interactions of agents specific to given situations, the affect control theory of self (ACT-S) was developed to include more dispositional, cross-situational knowledge about agents (MacKinnon and Heise 2010). ACT-S describes the *self* as the totality of a person's sense of themselves, and may include many identities that a person can enact in different social situations. ACT-S describes how a person will select an identity to enact in a given situation as the one that will best lead to an interaction (according to ACT) that confirms their true sense of themselves (their *self*). This interaction makes the person feel *authentic*, but others may lead to feelings of inauthenticity, and to further pressure to seek self-confirming interactions.

Recently, a probabilistic and decision theoretic generalisation of ACT, called *BayesAct*, has been shown to be a suitable framework for developing socially and emotionally intelligent agents (Hoey, Schröder, and Alhothali 2013b; Lin et al. 2014). *BayesAct* is a partially-observable Markov decision process (POMDP) model, and thus generalises ACT by modeling affective states as probability distributions, and allowing decision-theoretic reasoning about affect and other application-specific "cognitive" aspects. In this paper, we present a similar generalisation for the affect control theory of self, called *BayesAct-S*. The introduction of probability distributions to ACT-S resolves some key problems with the sociological theory that we expose in this paper.

Background

Affect Control Theory

Affect Control Theory (ACT) arises from work on the psychology of human social interaction (Heise 2007). ACT proposes that social perceptions, behaviours, and emotions are guided by a psychological need to minimize the differences between culturally shared fundamental affective sentiments

[1]The meaning of the term *identity* differs considerably across scientific disciplines. Here, we adhere to the tradition in sociology where it essentially denotes a kind of person in a social situation.

about social situations and the transient impressions resulting from the interactions between elements within those situations. Fundamental sentiments, $\mathbf{f}$, are representations of social objects, such as interactants' identities and behaviours, as vectors in a 3D affective space, hypothesised to be a universal organising principle of human socio-affective experience (Osgood, May, and Miron 1975). The basis vectors of affective space are called Evaluation/valence, Potency/control, and Activity/arousal (EPA). EPA profiles of concepts can be measured with the *semantic differential*, a survey technique where respondents rate affective meanings of concepts on numerical scales with opposing adjectives at each end (e.g., good, nice vs. bad, awful for E, weak, little vs. strong, big for P, and calm, passive vs. exciting, active for A). Affect control theorists have compiled lexicons of a few thousand words along with average EPA ratings obtained from survey participants who are knowledgeable about their culture (Heise 2010). For example, most English speakers agree that professors are about as nice as students (E), more powerful (P) and less active (A). The corresponding EPAs are $[1.7, 1.8, 0.5]$ for professor and $[1.8, 0.7, 1.2]$ for student[2]. In Japan, professor has the same P (1.8) but students are seen as much less powerful (-0.2). As the EPA structure of social concepts has been shown to be highly consensual within cultures, affective lexicons can be regarded as rudimentary collective representations of human sociality (Ambrasat et al. 2014).

Social events can cause transient impressions, $\boldsymbol{\tau}$ (also three dimensional in EPA space) of identities and behaviours that may deviate from their corresponding fundamental sentiments, $\mathbf{f}$. ACT models this formation of impressions from events with a grammar of the form actor-behaviour-object. Consider for example a professor (actor) who yells (behaviour) at a student (object). Most would agree that this professor appears considerably less nice (E), a bit less potent (P), and certainly more aroused (A) than the cultural average of a professor. Such transient shifts in affective meaning caused by specific events are described with models of the form $\boldsymbol{\tau}' = M\mathscr{G}(\mathbf{f}', \boldsymbol{\tau})$, where M is a matrix of statistically estimated prediction coefficients from empirical impression-formation studies and $\mathscr{G}$ is a vector of polynomial features in $\mathbf{f}'$ and $\boldsymbol{\tau}$. In ACT, the weighted sum of squared Euclidean distances between fundamental sentiments and transient impressions is called *deflection*, and is hypothesised to correspond to an aversive state of mind that humans seek to avoid. This *affect control principle* allows ACT to compute *normative* actions for humans: those that minimize the deflection. Normative predictions of ACT have been shown to be highly accurate in explaining verbal behaviours of mock leaders in a computer-simulated business (Schröder and Scholl 2009), non-verbal displays in dyadic interactions (Schröder et al. 2013), and group dynamics (Heise 2013), among others (MacKinnon and Robinson 2014).

[2]Values range by convention from -4.3 to $+4.3$ (Heise 2010). This range has historical roots in Osgood's massive cross-cultural studies of affective sentiments (Osgood, May, and Miron 1975). The Bayesian version (*BayesAct*) operates on a real-valued scale and so is not restricted to only consider vectors within that range.

Partially Observable Markov Decision Processes

A partially observable Markov decision process (POMDP) (Åström 1965) is a stochastic control model that consists of a finite set $\boldsymbol{\mathcal{X}}$ of states; a finite set $\boldsymbol{\mathcal{A}}$ of actions; a stochastic transition model $\Pr : X \times A \to \Delta(X)$, with $\Pr(x'|x, a)$ denoting the probability of moving from state x to x' when action a is taken, and $\Delta(X)$ is a distribution over $\boldsymbol{\mathcal{X}}$; a finite observation set $\boldsymbol{\Omega_x}$; a stochastic observation model, $\Pr(\omega_x|x)$, denoting the probability of making observation $\omega_x \in \boldsymbol{\Omega_x}$ while the system is in state x; and a reward $R(a, x')$ for a transition to x' induced by action a.

A *policy* maps *belief states* (i.e., distributions over $\boldsymbol{\mathcal{X}}$) into choices of actions, such that the expected discounted sum of rewards is (approximately) maximised.

In this paper, we will be dealing with *factored* POMDPs in which the state is represented by the cross-product of a set of variables or features. POMDPs have been used as models for many human-interactive domains, including intelligent tutoring systems (Folsom-Kovarik, Sukthankar, and Schatz 2013), spoken dialogue systems (Williams and Young 2006), and assistive technologies (Hoey et al. 2012).

BayesAct

Recently, ACT was generalised and formulated as a POMDP (Hoey, Schröder, and Alhothali 2013b) for human-interactive artificially intelligent systems. This new model, called *BayesAct*, generalises the original theory in three ways. First, sentiments are viewed as probability distributions over latent variables rather than points in the EPA space, allowing for multimodal, uncertain and dynamic sentiments to be modeled and learned. Second, affective interactions are augmented with *propositional* state (e.g. the usual state space considered in typical AI applications). Third, an explicit reward function allows for goals that go beyond simple deflection minimization.

A *BayesAct* POMDP models an interaction between two agents (human or machine) denoted *agent* and *client*. The state is the product of six 3-dimensional continuous random variables corresponding to fundamental and transient sentiments about the *agent*'s identity $(\mathbf{F}_a, \mathbf{T}_a)$, behaviour $(\mathbf{F}_b, \mathbf{T}_b)$ and *client*'s identity $(\mathbf{F}_c, \mathbf{T}_c)$. We use $\mathbf{F} = \{\mathbf{F}_a, \mathbf{F}_b, \mathbf{F}_c\}$ and $\mathbf{T} = \{\mathbf{T}_a, \mathbf{T}_b, \mathbf{T}_c\}$. Actions in the *BayesAct* POMDP are factored in two parts: $\mathbf{b_a}$ denotes the *affective* component or interpretation, and a represents the *propositional* part. For example, if a tutor gives a hard exercise to do, the manner in which it is presented, and the difficulty of the exercise, combine to form an affective impression $\mathbf{b_a}$ that is communicated. The actual exercise itself (content, difficulty level, etc) is the *propositional* part of the action, a. The fundamental behaviour, $\mathbf{F}_b$, denotes either observed *client* or taken *agent* affective action. That is, when *agent* acts, there is a deterministic mapping from the affective component of his action (denoted $\mathbf{b_a}$ in *BayesAct*) to the *agent*'s behaviour $\mathbf{F}_b$. When *client* acts, *agent* observes evidence of $\mathbf{F}_b$ (as the affective behaviour of the other agent).

The transient impressions, $\mathbf{T}$, evolve according to the deterministic impression-formation operator in ACT ($M\mathscr{G}$). Fundamental sentiments are expected to stay approximately constant over time, but are subject to random drift (with noise Σ_f) and are expected to also remain close to the transient impressions because of the *affect control principle*. Thus, the dynamics of $\mathbf{F}$ can be written as:

$$Pr(\mathbf{f}'|\mathbf{f}, \boldsymbol{\tau}) \propto e^{-\psi(\mathbf{f}', \boldsymbol{\tau}) - \xi(\mathbf{f}', \mathbf{f})} \tag{1}$$

where $\psi \equiv (\mathbf{f}' - \boldsymbol{M}\mathscr{G}(\mathbf{f}', \boldsymbol{\tau}))^T \boldsymbol{\Sigma}^{-1} (\mathbf{f}' - \boldsymbol{M}\mathscr{G}(\mathbf{f}', \boldsymbol{\tau}))$ combines the *affect control principle* with the impression formation equations and $\xi \equiv (\mathbf{f}' - \mathbf{f})^T \boldsymbol{\Sigma}_f^{-1} (\mathbf{f}' - \mathbf{f})$ represents the inertia of fundamental sentiments. The two terms can be combined into a single Gaussian that is non-linearly dependent (due to the polynomial features in $\mathscr{G}(\tau, \mathbf{x})$) on the previous state. *BayesAct* thus uses sample-based representation for beliefs, and a *bootstrap filter* for belief updates.

The probability distribution in (1) gives the normative (expected) action as one of the components of $\mathbf{f}'$: $\mathbf{f}'_b$. Thus, by integrating over $\mathbf{f}_a$ and $\mathbf{f}_c$ and the previous state, we obtain a distribution over $\mathbf{f}_b$ that acts as a *normative action prediction*: it tells the agent what to expect from other agents, and what action is expected from it.

BayesAct includes an application-specific set of random variables $\mathbf{X}$ that are interpreted as the *non-affective* elements of the domain (e.g. whose turn it is, behaviours of other agents, game states, student knowledge representations for intelligent tutors, or plan-steps representing progress through a task). The dynamics of $\mathbf{X}$ are application specific, but depend in general on the *deflection*, and on the propositional component of the action, a (which complements the affective component, $\mathbf{b_a}$). Thus, we require a definition for $Pr(\mathbf{x}'|\mathbf{f}, \boldsymbol{\tau}, \mathbf{x}, a)$. We refer to this distribution as the *social coordination* rule: it defines what agents are expected to do in a situation with sentiments $\mathbf{f}$ and $\boldsymbol{\tau}$. For example, we would expect faster learning from a student if deflection is low, as they do not have to use valuable cognitive resources to deal with any mis-alignment with the tutor. $\mathbf{X}$ also may condition the affective dynamics of $\mathbf{F}$ and $\mathbf{T}$ (e.g. a professor will act differently with students of differing skill levels).

The *BayesAct* POMDP assumes that the state is not observable, but that observations are obtained for some (possibly empty) subset of $\mathbf{X}$ and for the affective behaviour $\mathbf{F}_b$. As affective identities are latent variables, they are learned (as inference) in the POMDP. Thus, if behaving normatively, an agent will perform affective actions ($\mathbf{F}_b$) that allow other agents to infer what his (true) identity is. The normative action is thought to be used by humans as an emotional "fast thinking" heuristic: roughly corresponding to "System 1" thinking (Kahneman 2011). If agents are fully cooperative and aligned, then no further planning is required to ensure goal achievement. Agents do what is expected of them (which may involve planning over $\mathbf{X}$, but not over $\mathbf{F}$ and $\mathbf{T}$), and expect others to do so as well. However, when alignment breaks down, or in non-cooperative situations, then slower, more deliberative "System 2" thinking is required.

To accomplish this deliberative thinking, *BayesAct* uses a Monte-Carlo tree search (MCTS) method called POMCP-C (Asghar and Hoey 2014). This method builds a search tree by sampling trajectories from the POMDP as a simulator. It requires an *action bias* from which samples are drawn in the action space, which here is the normative action prediction as above.

Affect Control Theory of Self

The affect control theory of self (ACT-S) describes a higher-order level of socio-affective control than ACT (MacKinnon and Heise 2010). ACT governs the choice of behaviours in a given situation given an identity. In contrast, ACT-S is used to select identities to enact in a particular setting and institution, respecting situational and dispositional constraints simultaneously. ACT-S describes a person's sense of *self* using the *fundamental self-sentiment*, an EPA vector that describes a composite feeling for the overall self. The fundamental self sentiment is thought to be relatively stable over time, providing a sense of consistency and giving protection against disruptive events. However, it can change drastically in short periods in exceptional circumstances.

ACT-S describes the recent history of self-feelings with a *situational self-sentiment* (again, an EPA vector). The situational self-sentiment is the ephemeral feeling about the *self* that a person has within a given setting. This feeling is a combination of the recently experienced interactions with others, with the situation, and with the institution in which the experiences occurred. The affect control principle's primary proposition is that *people construct situational self-sentiments that are consistent with their fundamental self-sentiment*. The difference between these two is called the *inauthenticity*: if one cannot enact a series of identities that create a situational self-sentiment that is consistent with one's fundamental self-sentiment, one feels a sense of *inauthenticity* in the situation (MacKinnon and Heise 2010).

We will denote the fundamental self sentiment as $\mathbf{S_f} = \{\mathrm{S}_{\mathrm{f}j}\}$ and the situational self-sentiment as $\mathbf{S_s} = \{\mathrm{S}_{\mathrm{s}j}\}$ where $j \in \{e, p, a\}$ and both are three-dimensional vectors. The *inauthenticity* is a three dimensional vector $\mathbf{I} = \{\mathrm{I}_j\}$, and is computed as $\mathbf{I} = \mathbf{F}_a - \mathbf{S_f}$ (a 3-vector difference): it gives the difference between the currently felt identity $\mathbf{F}_a$ and the fundamental self sentiment, $\mathbf{S_f}$.

The situational self-sentiment is defined as a weighted accumulation of identities enacted over time (Heise 2007):

$$\mathbf{s_s}^T = \sum_{t=0}^{T} w(t, T) \mathbf{f}_a^t \tag{2}$$

where superscripts denote time indices and $w(t, T)$ is a decay function, e.g. $w(t, T) = \eta^{T-t}$, with $0 < \eta \leq 1$ being a decay constant. The inauthenticity also accumulates over time, such that, at time T we have the accumulated inauthenticity, $\mathbf{I_a}$:

$$\mathbf{i_a}^T = \sum_{t=0}^{T} w(t, T) \left(\mathbf{f}_a^t - \mathbf{s_f}^t \right) = \mathbf{s_s}^T - \sum_{t=0}^{T} w(t, T) \mathbf{s_f}^t \tag{3}$$

We can make the further assumption that the fundamental self sentiment does not change very quickly, so that $\mathbf{s_f}^t = \mathbf{s_f}$, and we understand that $\mathbf{s_f}$ means the fundamental self-sentiment at the current time, T. Thus, we can write:

$$\mathbf{i_a}^T = \mathbf{s_s}^T - \mathbf{s_f} \sum_{t=0}^{T} w(t, T)$$

which, if we use the geometric decay constant η as above, and let $T \to \infty$, becomes

$$\mathbf{i_a} = \mathbf{s_s} - \mathbf{s_f}\frac{1}{1-\eta} \tag{4}$$

Where $\mathbf{i_a}$, $\mathbf{s_s}$ and $\mathbf{s_f}$ represent the accumulated inauthenticity, situational self-sentiment, and fundamental self sentiment at the current time.

We can see from Equation 2 that the enacted identities are summed to create the situational self sentiment. What this implies is that a person can create a situational self-sentiment that is neutral by acting very positively (E), and then acting negatively to offset the positivity. For example, one could enact the identity of a *saint* ($E = 2.81$), followed by the identity of a *sinner* ($E = -1.68$), and if $w = 0.6$, the resulting self sentiment would be close to $E = 0$, and the actor would feel very neutral. As a more realistic example, consider the fictional case of Sara Sim, the CEO of a large and prosperous technology firm. Sara has both the *female* identity $EPA = \{1.60, 0.22, 0.42\}$, and the *executive* identity $EPA = \{1.26, 1.93, 1.44\}$. These two identities are substantially different on the potency and activity axes. If we assume that Sara primarily enacts her *executive* identity during the day while at work (call this $t = 2-5$), and her *female* identity in the evenings ($t = 6-7$) and mornings ($t = 0-1$) while at home with her family, then, taking $\eta = 0.8$, we have that her self sentiment at the end of the day according to Equation 2 would be $\{6.02, 4.15, 3.68\}$, which if rescaled by $1 - \eta$ according to Equation 4, gives $\{1.20, 0.83, 0.74\}$, with a closest label of *businesswoman* in EPA lexicons.

Bayesian Affect Control Theory of Self

In *BayesAct*, identities are no longer points in the EPA space, but probability distributions. Here, we hypothesise that both fundamental self-sentiments and situational self-sentiments are also probability distributions. The idea of treating sentiments as distributions removes a drawback of affect control theory, that a single identity must be selected at each time, and that each behaviour can only correspond with a single sentiment. One frequently encounters situations where multiple identities may be mixed, and multiple sentiments may occur simultaneously (Smith-Lovin 2007).

If we follow (Heise 2007), we would model the situational self-sentiment is as a sum of random variables given by Equation 2. We describe this model in the next section. Alternatively, in the subsequent section, we consider the situational self sentiment as a random variable that evolves as a function of the current enacted identity and the previous situational self-sentiment in a probabilistic mixture or "noisy-or" model.

Situational Self Sentiment: Summed Identities

Assume $w(T, T) = 1$ and re-write Equation 2 as:

$$\mathbf{s_s}^T = \mathbf{f}_a^T + \eta \sum_{t=0}^{T-1} w(t, T-1)\mathbf{f}_a^t = \mathbf{f}_a^T + \eta \mathbf{s_s}^{T-1}$$

or, alternatively:

$$\mathbf{s_s}' = \eta \mathbf{s_s} + \mathbf{f}_a \tag{5}$$

and thus, to compute $Pr(\mathbf{S_s}^T)$, we will need the convolution

$$Pr(\mathbf{s_s}^T) = Pr(\mathbf{f}_a^T) * Pr(\eta \mathbf{s_s}^{T-1}) \tag{6}$$

This convolution is not obvious to compute given distributions as sample sets of three dimensional vectors.

Situational Self Sentiment: Mixed Distributions

Rather than computing a probability distribution over the sum of fundamentals, we can sum the probability distributions by considering them as a probabilistic mixture model. The idea is that the situational self-sentiment at the current time step, $\mathbf{S_s}'$, is *caused* by *either* the previous situational self-sentiment, $\mathbf{S_s}$ *or* the current fundamental identity, $\mathbf{F}'_a$. However, the person is *uncertain* about which of these elements is the true cause of current situational self sentiment. As we are representing the person's subjective belief, the situational self-sentiment is really a probabilistic mixture of both. The person knows they used to feel one way, and knows they have just felt another way (which may be consistent), and knows that the way they feel is something of a combination of both of these feelings.

Mathematically, we consider that there is a binary random variable C that is drawn from a Bernoulli distribution with parameter η (i.e. a weighted coin flip with probability η of turning up heads), and that $C = c$ ("true") means that $\mathbf{S_s}'$ is caused by $\mathbf{S_s}$, and $C = \overline{c}$ ("false") means that it is caused by $\mathbf{F}'_a$. Thus:

$$c \sim Bernoulli(\eta, 1-\eta) \tag{7}$$

$$\mathbf{s_s}' = c\mathbf{s_s} + (1-c)\mathbf{f}'_a \tag{8}$$

This means that to compute $Pr(\mathbf{S_s}')$ we can sum over $C, \mathbf{F}'_a, \mathbf{S_s}$:

$$\begin{aligned} Pr(\mathbf{s_s}') &= \int_{\mathbf{s_s},\mathbf{f}'_a} \big[Pr(\mathbf{s_s}', c|\mathbf{s_s}, \mathbf{f}'_a) + \\ &\qquad Pr(\mathbf{s_s}', \overline{c}|\mathbf{s_s}, \mathbf{f}'_a)\big] Pr(\mathbf{s_s}, \mathbf{f}'_a) d\mathbf{s_s} d\mathbf{f}'_a \\ &= \int_{\mathbf{s_s},\mathbf{f}'_a} \big[Pr(\mathbf{s_s}'|c, \mathbf{s_s}) Pr(c) Pr(\mathbf{s_s}, \mathbf{f}'_a) + \\ &\qquad Pr(\mathbf{s_s}'|\overline{c}, \mathbf{f}'_a) Pr(\overline{c}) Pr(\mathbf{s_s}, \mathbf{f}'_a)\big]\, d\mathbf{s_s} d\mathbf{f}'_a \\ &= Pr(c)Pr(\mathbf{s_s}) + Pr(\overline{c})Pr(\mathbf{f}_a) \\ &= \eta Pr(\mathbf{s_s}) + (1-\eta)Pr(\mathbf{f}_a) \end{aligned} \tag{9}$$

We may also want C to be conditioned on other elements of the state, such as the institution (part of $\mathbf{X}$ in *BayesAct*), in which case we work with $Pr(C|\mathbf{X})$, or we could write the parameter as a function of the state, $\eta(\mathbf{x})$. We also may consider that some additional noise is added at each time step, so that

$$\mathbf{s_s}' = c\mathbf{s_s} + (1-c)\mathbf{f}'_a + \rho \tag{10}$$

where $\rho \sim \mathcal{N}(\mathbf{s_s}; 0, \mathbf{\Sigma_s})$ is normally distributed random noise (may be different noises for c and $\overline{c}$). In this case, we have

$$\begin{aligned} Pr(\mathbf{s_s}') = \eta \int_{\mathbf{s_s}} \mathcal{N}(\mathbf{s_s}'; \mathbf{s_s}, \mathbf{\Sigma_s}) Pr(\mathbf{s_s}) + \\ (1-\eta)\int_{\mathbf{f}'_a} \mathcal{N}(\mathbf{s_s}'; \mathbf{f}'_a, \mathbf{\Sigma_s}) Pr(\mathbf{f}'_a) \end{aligned} \tag{11}$$

As we are representing these distributions with sample sets, it is easy to apply Equation 9 in practice, by choosing a random sample of ηN samples from $Pr(\mathbf{s_s}^{T-1})$ and $(1-\eta)N$ samples from $Pr(\mathbf{f}_a)$. Equation 11 is also straightforward, as we simply do the same but add some random normally distributed noise to each sample.

Differences between the two representations

The two representations of the situational self-sentiment described in the last two sections are fundamentally different ways of looking at how the self-sentiments are updated sequentially. In the first case, identities are summed over time as vectors, thus averaging out to create the situational self-sentiment. In the following, we refer to this as "averaging". In the second case, identities are mixed over time. We will refer to this as "noisy-or".

Consider again the case of Sara Sim. If using averaging, Sara's situational self-sentiment will be a unimodal distribution centered around "businesswoman" (EPA=$\{1.4, 1.1, 0.9\}$). If using noisy-or, the situational self-sentiment will be bimodal, with one mode at "female" and the other at "executive". Now, which of these two self-sentiments match her fundamental self-sentiment better? If she could enact the identity of "businesswoman" exactly at home and at work, would she feel more or less authentic? If the former, then we should consider the averaging model as more accurate. If the latter, it would be the noisy-or model. We believe the latter is more accurate: Sara wants to feel like a female when out on a romantic dinner with her partner, not like a businesswoman. She wants to feel like a powerful executive while at work, not a menial businesswoman. If one were to ask Sara at the end of her day how she felt using a single word, she might average and report *businesswoman*. On the other hand, if given more flexibilty in her answer, she might report that she feels a bit of both *female* and *executive*.

Females in powerful roles will often be assigned the less powerful of the two identities by others, especially if there are no direct personal or institutional indications of the more powerful roles. For example, consider the case of Merribel G. Krosby, a female professor with a feminine and youthful outward appearance. The identity of professor has an EPA of $\{1.61, 1.58, 0.35\}$, while female is as before $\{1.6, 0.22, 0.42\}$. The average identity for Merribel is $EPA = \{1.6, 0.86, 0.36\}$, and is closest to that of a "graduate student" ($EPA = \{1.6, 1.58, 0.35\}$) if considering the academic setting. People across the academic setting, from directors and deans to undergraduate students, often treat Merribel as a grad student, or simply a student ($EPA = \{1.49, 0.31, 0.75\}$), or as a female, because there is no direct visual evidence when meeting her that she is a professor. However, this assignment of role is not satisfying to Merribel, as it minimises her self-sentiments of powerfulness within the academic setting. In contrast, if she is going out with a romantic interest who treats her only as a professor, she may feel frustrated as she wishes to be seen outside of that role. In these situations, Merribel will feel like she has to play the role assigned to her by the other person, as not doing so will create significant deflection that is difficult to avoid. Merribel's feelings of inauthenticity when treated as the average of her two identities seems to be a good indication that the noisy-or model may be a more accurate representation of what is happening in this situation at least[3].

Comparing Situational and Fundamental Selves

Once we have the distribution $Pr(\mathbf{s_s})$, we want to compare to an existing distribution describing the fundamental self-sentiment, $Pr(\mathbf{s_f})$ in order to compute the accumulated inauthenticity. Although it may seem intuitive to use Equation 4, when considering distributions, the resulting probabilistic difference will be incorrect [4]. Instead, we want to measure inauthenticity as a function over the sentiment space which is a ratio of the two:

$$\mathbf{i_a}(\mathbf{s}) = ln\left(\frac{Pr(\mathbf{s_s})}{Pr(\mathbf{s_f})}\right) \tag{12}$$

Equation 12 can be used to compute the expected inauthenticity as the integral over the EPA space, $\mathbf{s}$:

$$\mathbb{E}[\mathbf{i_a}] = \int_{\mathbf{s}} \mathbf{i_a}(\mathbf{s}) Pr(\mathbf{s_s}) d\mathbf{s} \tag{13}$$

We recognize this as the Kullback-Leibler (KL) divergence between $\mathbf{s_f}$ and $\mathbf{s_s}$. This can be used to select an person to interact with in the following way. Suppose our agent already interacted with M different persons, and had learned a situational self-sentiment, $\mathbf{s_s}_j$ for each $j \in \{1 \ldots M\}$. Assuming the agent has a fixed fundamental self sentiment, $\mathbf{s_f}$, it would select j^* as the next person to interact with, where

$$j^* = \arg\min_j \left(\int_{\mathbf{s}} ln\left(\frac{Pr(\mathbf{s_s}_j)}{Pr(\mathbf{s_f})}\right) Pr(\mathbf{s_s}_j) d\mathbf{s} \right) \tag{14}$$

Updating fundamental self-sentiments

Finally, we may want to be able to update the fundamental self-sentiment as a function of the history of situational self-sentiments. This would enable agents to learn through social interaction about other agents' more enduring sense of self or personality. We can do this using the following analogue of Equation 8:

$$c_f \sim Bernoulli(\eta_f, 1-\eta_f) \tag{15}$$

$$\mathbf{s_f}' = c_f \mathbf{s_f} + (1-c_f)\mathbf{s_s}' \tag{16}$$

As we expect fundamental self-sentiments to change slowly, we would expect that $\eta_f > \eta$ and is very close to 1.0.

We may find some parts of an individual's self-sentiment decaying faster than others, and this may have a profound impact on interpersonal relations. Take for example, a father-daughter relationship which supports the self-sentiment of

[3] While we have received significant support in personal communication from sociologists for the "noisy-or" model, a social-psychological experiment will need to be carried out to test these models. For example, people could be queried for sentiments of different pairs of identities, and then for the average.

[4] e.g. consider a convolution of two identical distributions, which gives a third distribution with all possible ways of subtracting elements from each.

father for the father and *daughter* for the daughter. Upon leaving home, the daughter's *daughter* identity rapidly diminishes as she engages in fewer interactions with her father, and she has a small η_f value (perhaps brought on by her change of situation). The father's η_f value remains the same though, so his self-sentiment of *father* decreases much less quickly. Thus, a discrepancy is created. The father attempts to resolve the discrepancy (the daughter is not communicating with him enough to support his sense of *father*) by attempting to interact with his daughter. This creates inauthenticity for both people, due to a difference in η_f values.

Simulations

To implement *BayesAct-S* in practice, we make use of the *BayesAct* function $\texttt{SIMULATE}(\mathbf{f}_a, \mathbf{f}_c) \rightarrow (\mathbf{f}'_a, \mathbf{f}'_c)$[5]. This function is used to map a distribution over agent and client identities, $\mathbf{f}_a$ and $\mathbf{f}_c$, respectively, to another such distribution after a single interaction (agent acts, client acts). There are two cases to be considered: known or unknown *client* identity. If the client identity is known, `SIMULATE` creates two *BayesAct* Agent instances, one for *agent* and one for the known *client* identity. Two updates then occur: one to update an *agent* simulator, and one to update a (known) *client* simulator. In the second case, the client identity is unknown (a stranger, so $\mathbf{f}_c$ is set to $\mathbf{s}_\circ$, e.g. the mean of all identities in the lexicon). In this case, `SIMULATE` only create a single *BayesAct* Agent instance for *agent*, and does a *BayesAct* simulation where the observations on *client* turn are null, so the weights computed are the same for all samples. This is the same as assuming the *client* does the deflection minimizing predictions of the *agent* (with added noise).

A simulation proceeds for T steps, starting with a situational self-sentiment $\mathbf{s_s}$ equal to the fundamental self-sentiment, $\mathbf{s_f}$. After each call to `SIMULATE`, we update the situational self-sentiment using the resulting $\mathbf{f}_a$ as $\mathbf{s_s}' \leftarrow \eta\mathbf{s_s} + (1-\eta)\mathbf{f}_a$. This is shorthand to mean that $\mathbf{s_s}'$ is distributed according to a probabilistic mixture of $\mathbf{s_s}$ and $\mathbf{f}_a$, with probabilities η and $1-\eta$. In practice, we accomplish this by drawing ηN samples from $Pr(\mathbf{s_s})$ and $(1-\eta)N$ samples from $Pr(\mathbf{f}_a)$ to create the distribution $Pr(\mathbf{s_s}')$ represented with N samples. At the end of T steps, we can then look at the situational self-sentiment and see how this has changed based on the interaction.

We will consider the case of a female *agent* with a mixture of two identities, *daughter* (EPA=$[2.73, 1.13, 1.28]$) and *employer* (EPA=$[1.48, 1.93, 0.74]$), with equal proportions (0.5). There are two known client identities: *mother* (EPA=$[3.12, 2.98, 1.44]$ from female or agent perspective) and *employee* (a female, with EPA=$[1.88, 0.05, 0.84]$). We will expect that, when interacting with the *mother*, the *female/employer agent* will end up with a self sentiment focused on *daughter*. In contrast, when interacting with the *employee*, the *female/employer agent* will end up with a self-sentiment focused on *employer*.

Figure 1 shows the results of a simulation with $T = 20$, $\eta = 0.95$, $\eta_f = 1.0$ and with $N = 2000$ samples for the

[5]*BayesAct* code and videos of simulations are available at `bayesact.ca`.

	employee	stranger	mother
employee	3.09	2.46	**2.17**
stranger	2.37	2.96	**2.27**
mother	**2.18**	2.38	2.80

Table 1: KL-divergences after $T = 20$ steps. If the daughter/employer *agent* has interacted for 20 steps with the (row) identity, then interacting with the (column) identity will cause the divergence between situational self-sentiments and fundamental self-sentiments as shown. Smaller divergences means the identity is a preferable interactant.

identities in *BayesAct*. The *BayesAct* parameters are: Σ is diagonal with elements 0.1, and Σ_f is diagonal with elements 0.001 for client identity, and 0.01 for agent identity. Thus, we have made the *agent* more *susceptible* to the actions of the *client* than the other way around. The initial variances on agent and client identities are 0.01.

The simulated *client* in `SIMULATE` uses the heuristic policy of *BayesAct*, since this will be approximately correct for it in any case. This heuristic policy does a one-step lookahead using the normative predictions of the affect control principle (see (Hoey, Schröder, and Alhothali 2013a) for details). The simulated *agent*, on the other hand, *must* use the POMCP-C policy, as the heuristic policy uses an *average* identity to compute the best behaviour. For a person with multiple, simultaneous identities, this will cause problems as the reactions will be for neither one or the other identity.

We see in Figure 1 that interactions with a *mother* identity accentuates the *daughter* self-sentiment, while interactions with a *employee* identity accentuates the *employer* self-sentiment. Interactions with a stranger (unknown identity)creates a sort of middle ground identity.

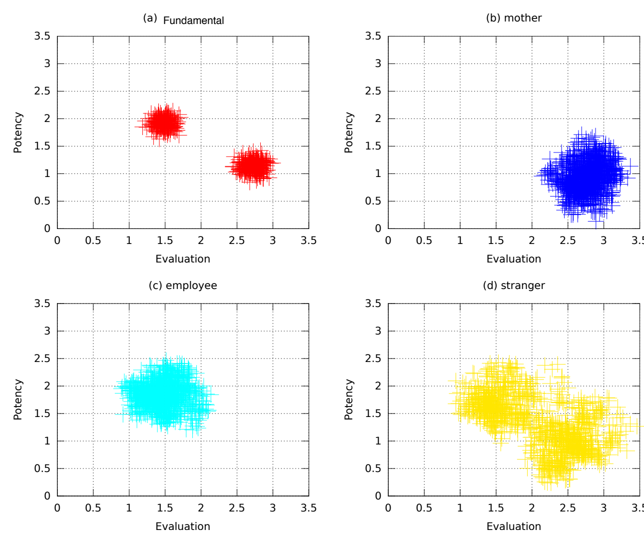

Figure 1: *agent* final self-sentiment after interacting with different *clients*. *agent* starts with the self-sentiment that is mixed between *daughter* (P=1.13) and *employer* (P=1.93). After $T = 20$: (a) fundamental self-sentiment; and after interacting with (b) *mother*; (c) *employee*; (d) a stranger.

We can now investigate how the *agent* selects a new interactant based on the divergence between the situational and fundamental self-sentiments. Given the situational self-sentiment $\mathbf{s_s}$ for each interactant (e.g. as shown in Figure 1), we create a *predicted* situational self-sentiment (that will be achieved by future interactions) by simply creating a mixture of every pair of situational self-sentiments. We then compute the divergence from the fundamental self-sentiment of this mixture, using Equation 14. To compare two belief distributions in practice, we use a nearest-neighbours approach (Perez-Cruz 2008). Table 1 shows the KL-divergences (closer to zero means a more compatible interactant, shown in **boldface**). We see that, after interacting with the *employee* for 20 steps (first row), the *mother* is the most likely interactant. After interacting with a *mother*, however, the best interactant is *employee*. After interacting with a *stranger*, the *agent* will prefer interacting with either *mother* or *employee*, with slight preference for *mother*.

Discussion

The socio-affective model described in this paper is used for fast, heuristic, everyday human interaction: it is what agents can use to "get along" with other agents or humans in a social world. While *BayesAct* (Hoey, Schröder, and Alhothali 2013b; 2013a) integrates this heuristic into a cognitive reasoning engine (a POMDP), *BayesAct-S* (this paper) expands this notion to a societal level, giving agents an intrinsic guide to selecting goals, settings, institutions, and individuals to engage with. In this hierarchy, the high level is reasoning about the self and longer-term goals, while the lower level is considering individual interactions with other agents, guided in part by the affect control principle. Although this could be formalised as a hierarchical POMDP (Theocharous, Rohanimanesh, and Mahadevan 2001), here the higher level simply uses a heuristic policy based on the inauthenticity.

Thus, *BayesAct-S* agents use a novel unification of cognitive (individual) and affective (social) reasoning, but are free to use the cognitive component to model other agents at any level of detail, including as full POMDPs (Doshi and Gmytrasiewicz 2009). Linking the social and individual has relatively recently come into focus in AI research (Castelfranchi 1998), and recent work has looked at unifying appraisal theories of emotions with cognitive reasoning and reinforcement learning. For example, (Hogewoning et al. 2007) links valence with an exploration bonus, and in (Marinier III and Laird 2008), the SOAR cognitive architecture is augmented with a reinforcement learning agent that uses emotional appraisals as intrinsic reward signals.

BayesAct-S is defined with observations and actions in the 3-D EPA space. Therefore, to implement a *BayesAct-S* agent, we require the definition of input and output mappings to translate EPA vectors ($\mathbf{b_a}$) into physical actions and to translate observations of the world into EPA vectors ($\mathbf{F}_b$). These I/O mappings are application dependent. For example, in text-based applications such as the tutoring system in (Hoey, Schröder, and Alhothali 2013b), one may use a semantic analysis of the messages (Pang and Lee 2008). In an embodied agent, one might use facial expressions, gestures or body language measures (Schröder et al. 2013; Lin et al. 2014).

As discussed in (Rogers, Schröder, and von Scheve 2014), affect control theory is conceptually compatible with both dimensional (Russell and Mehrabian 1977) and appraisal (Scherer, Banziger, and Roesch 2010) theories of emotions. The EPA dimensions of affective space can be understood as very basic appraisal rules related to the goal congruence of an event (E), the agent's competence/coping potential (P), and the urgency implied by the situation (A) (Rogers, Schröder, and von Scheve 2014; Rogers, Schröder, and Scholl 2013; Scherer, Dan, and Flykt 2006). However, ACT is also more general and more parsimonious than many appraisal theories, since deflection minimisation is the only prescribed mechanism, while the more specific goals tied to types of agents and situations are assumed to emerge from the semantic knowledge base of the model. Although it is clear that a 3-D vector cannot explain *all* the facets of emotions and identities, this simple scheme has been repeatedly shown to explain roughly 50% of the semantic variability in word association data (Osgood, Suci, and Tannenbaum 1957; Osgood, May, and Miron 1975). The EPA dimensions are also thought to be related directly to intrinsic reward (Fennell and Baddeley 2013). That is, reward is assessed by humans along the same three dimensions: Evaluation roughly corresponds with expected value, Potency with risk (e.g. powerful things are more risky to deal with, because they do what they want (Scholl 2013)), and Activity corresponds roughly with uncertainty, increased risk, and decreased values (e.g. faster more exited things are more risky and less likely to result in reward).

ACT has intellectual origins in symbolic interactionism, a paradigm that emphasizes the importance of semiotic references for the regulation of social interaction (MacKinnon 1994). Recent work has pointed out how the statistical EPA model is compatible with symbolic approaches to social cognition at the societal level as well as notions of embodied "deep meanings" at the level of individual brains. In many ways, the EPA model can be regarded as a "translation" between societal symbols and signals of the body (Schröder and Thagard 2013; Thagard and Schröder 2014).

Conclusion

We have described a probabilistic and decision theoretic generalisation of the *affect control theory of self*, a coherent and well-grounded sociological theory of human social interactions that uses affective descriptions of identities, selves and behaviours. We have shown how this theory can be used to construct socially and emotionally aware agents that can reason about the affective identities of others and use this reasoning to guide their choices of whom to interact with and what to do in each interaction. In future, we plan to apply these concepts in social networking applications, assistive technologies, and to further investigate the usage of external rewards in guiding agent behaviours. We plan to investigate large-scale social simulations based on these concepts. *BayesAct-S* could also substantiate recent developments of a "social physics" (Pentland 2014) by enhancing it with a well-grounded theory of the behaviours of individuals.

Acknowledgments

Thanks to Neil J. MacKinnon, Lynn Smith-Lovin, David R. Heise, and Kimberly B. Rogers for helpful discussions. This work was funded in part by a grant from the Natural Sciences and Engineering Research Council of Canada.

References

Åström, K. J. 1965. Optimal control of Markov decision processes with incomplete state estimation. *J. Math. Anal. App.* 10:174–205.

Ambrasat, J.; von Scheve, C.; Conrad, M.; Schauenburg, G.; and Schröder, T. 2014. Consensus and stratification in the affective meaning of human sociality. *Proceedings of the National Academy of Sciences* 111(22):8001–8006.

Asghar, N., and Hoey, J. 2014. Monte-carlo planning for socially aligned agents using Bayesian affect control theory. Tech. Rep. CS-2014-21, U. of Waterloo School of Computer Science.

Castelfranchi, C. 1998. Modeling social action for AI agents. *Artificial Intelligence* 103:157–182.

Doshi, P., and Gmytrasiewicz, P. 2009. Monte-Carlo sampling methods for approximating interactive POMDPs. *Journal of Artificial Intelligence Research* 34:297–337.

Fennell, J. G., and Baddeley, R. J. 2013. Reward is assessed in three dimensions that correspond to the semantic differential. *PLoS One* 8(2: e55588).

Folsom-Kovarik, J. T.; Sukthankar, G.; and Schatz, S. 2013. Tractable POMDP representations for intelligent tutoring systems. *ACM Trans. Intell. Syst. Technol.* 4(2):29:1–29:22.

Heise, D. R. 2007. *Expressive Order: Confirming Sentiments in Social Actions*. Springer.

Heise, D. R. 2010. *Surveying Cultures: Discovering Shared Conceptions and Sentiments*. Wiley.

Heise, D. R. 2013. Modeling interactions in small groups. *Social Psychology Quarterly* 76:52–72.

Hoey, J.; Boutilier, C.; Poupart, P.; Olivier, P.; Monk, A.; and Mihailidis, A. 2012. People, sensors, decisions: Customizable and adaptive technologies for assistance in healthcare. *ACM Trans. Interact. Intell. Syst.* 2(4):20:1–20:36.

Hoey, J.; Schröder, T.; and Alhothali, A. 2013a. Affect control processes: Intelligent affective interaction using a partially observable Markov decision process. http://arxiv.org/abs/1306.5279.

Hoey, J.; Schröder, T.; and Alhothali, A. 2013b. Bayesian affect control theory. In *Proc. of Intl. Conf. on Affective Computing and Intelligent Interaction*.

Hogewoning, E.; Broekens, J.; Eggermont, J.; ; and Bovenkamp, E. G. 2007. Strategies for affect-controlled action-selection in SOAR-RL. In Mira, J., and Àlvarez, J., eds., *IWINAC*, volume 4528 Part II of *LNCS*, 501–510.

Kahneman, D. 2011. *Thinking, Fast and Slow*. Doubleday.

Lin, L.; Czarnuch, S.; Malhotra, A.; Yu, L.; Schröder, T.; and Hoey, J. 2014. Affectively aligned cognitive assistance using Bayesian affect control theory. In *Proc. of IWAAL*.

MacKinnnon, N. J., and Robinson, D. T. 2014. 25 years of research in affect control theory. *Adv. in Group Processing* 31.

MacKinnon, N. J., and Heise, D. R. 2010. *Self, identity and social institutions*. New York, NY: Palgrave and Macmillan.

MacKinnon, N. J. 1994. *Symbolic Interactionism as Affect Control*. Albany: State University of New York Press.

Marinier III, R. P., and Laird, J. E. 2008. Emotion-driven reinforcement learning. In *Proc. of 30th Annual Meeting of the Cognitive Science Society*, 115–120.

Osgood, C. E.; May, W. H.; and Miron, M. S. 1975. *Cross-Cultural Universals of Affective Meaning*. U. of Illinois Press.

Osgood, C. E.; Suci, G. J.; and Tannenbaum, P. H. 1957. *The Measurement of Meaning*. Urbana: University of Illinois Press.

Pang, B., and Lee, L. 2008. Opinion mining and sentiment analysis. *Foundations and trends in information retrieval* 2.1(2):1–135.

Pentland, A. 2014. *Social Physics*. Penguin Press.

Perez-Cruz, F. 2008. Kullback-Leibler divergence estimation of continuous distributions. In *IEEE International Symposium on Information Theory (ISIT)*, 1666–1670.

Rogers, K. B.; Schröder, T.; and Scholl, W. 2013. The affective structure of stereotype content: Behavior and emotion in intergroup context. *Social Psychology Quarterly* 76:125–150.

Rogers, K. B.; Schröder, T.; and von Scheve, C. 2014. Dissecting the sociality of emotion: A multi-level approach. *Emotion Review* 6(2):124–133.

Russell, J. A., and Mehrabian, A. 1977. Evidence for a three-factor theory of emotions. *Journal of research in Personality* 11(3):273–294.

Scherer, K. R.; Banziger, T.; and Roesch, E. 2010. *A Blueprint for Affective Computing*. Oxford University Press.

Scherer, K. R.; Dan, E. S.; and Flykt, A. 2006. What determines a feeling's position in affective space: A case for appraisal. *Cognition and Emotion* 20(1):92–113.

Scholl, W. 2013. The socio-emotional basis of human interaction and communication: How we construct our social world. *Social Science Information* 52:3 – 33.

Schröder, T., and Scholl, W. 2009. Affective dynamics of leadership: An experimental test of affect control theory. *Social Psychology Quarterly* 72:180–197.

Schröder, T., and Thagard, P. 2013. The affective meanings of automatic social behaviors: Three mechanisms that explain priming. *Psychological Review* 120:255–280.

Schröder, T.; Netzel, J.; Schermuly, C.; and Scholl, W. 2013. Culture-constrained affective consistency of interpersonal behavior: A test of affect control theory with nonverbal expressions. *Social Psychology* 44:47–58.

Smith-Lovin, L. 2007. The strength of weak identities: Social structural sources of self, situation and emotional experience. *Social Psychology Quarterly* 70(2):106–124.

Thagard, P., and Schröder, T. 2014. Emotions as semantic pointers: Constructive neural mechanisms. In Barrett, L. F., and Russell, J., eds., *The psychological construction of emotion*. New York: Guilford.

Theocharous, G.; Rohanimanesh, K.; and Mahadevan, S. 2001. Learning hierarchical partially observed Markov decision process models for robot navigation. In *IEEE Intl. Conf. on Robotics and Automation*.

Williams, J. D., and Young, S. 2006. Partially observable Markov decision processes for spoken dialog systems. *Computer Speech and Language* 21(2):393–422.