# A modeling framework for inter-cultural social interactions


Taranjeet Singh Bhatia, Saad Ahmad Khan and Ladislau Bölöni

Dept. of Electrical Engineering and Computer Science
University of Central Florida
4000 Central Florida Blvd, Orlando FL 32816
{tsbhatia,skhan,lboloni}@eecs.ucf.edu



**Abstract.** In this paper, we describe a modeling framework which allows us to reason about the social-cultural behavior of humans in interaction scenarios. The objective is to provide explanatory and predictive power to an agent or robot which either witnesses or participates in the interaction: why have the human agents behaved the way they are and what kind of behavior can I expect as an answer to certain actions? We describe a the scenario modeling framework, the culture-dependent Culture Sanctioned Social Metrics (CSSM) model and the scenario dependent Concrete Belief (CB) model. We illustrate the use of the framework using an inter-cultural scenario.


## 1 Introduction

For an autonomous robot or software agent to participate in the social life of humans, it must have a way to perform a *calculus of social behavior* [11], an operational model which would allow it to reason about the implications of it's own and the humans' actions.

To be useful, such a calculus must satisfy several requirements. First, it must have *explanatory power* (it must provide a coherent theory for why the humans act the way they do), and *predictive power* (it must provide some plausible scenarios about the future actions of the humans). The social calculus must also be *culture-specific*, it needs to consider the different social norms and requirements in various societies. As many social interactions take place in public, it also needs to have a *model of public perception.*

The social calculus does not directly determine the behavior of the agent (unless the only goal of the agent is to act in a socially acceptable way). However, the behavioral and planning component can use social calculus to weigh different plans - this is especially critical if the plans require the goodwill or active help of the human actors.

In this paper, we describe a set of social calculus techniques designed to satisfy these requirements. In Section 2 we describe a formal model for the representation of social interactions, with special focus on the separation of the progress state and the full state. The next two sections introduce specific models for the representation of the components of the full state: culture sanctioned social metrics in Section 3, and concrete beliefs of individual actors and perception of the crowd in Section 4. Section 5 describes the use of the model on an intercultural interaction example. We briefly survey related work in Section 6.

## 2 Scenario modeling

### 2.1 The scenario model

The modeling of arbitrary, free format interactions between humans is clearly out of reach for theoretical models. Instead, we model specific *scenarios of human interaction* which center around the resolution of a small number of issues, have a limited number of participants and a limited time span.

**Definition 1.** *We call a* scenario *a tuple* $\{\mathbf{A}, \boldsymbol{\alpha}, \boldsymbol{\tau}, \mathbf{S}, \mathcal{F}, \mathcal{P}\}$*, where:*

> $\mathbf{A} = \{A_1, A_2 \ldots\}$ *is a set of* actors*, who are usually humans, although they can also be autonomous robots or software agents.*
> $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2 \ldots\}$ *is a set of distinct* action types*. A concrete action a is characterized by the tuple* $\{\alpha, A, (x_1 \ldots x_n)\}$*, that is, by the action type, the performing actor and a list of parameters of arbitrary length. We denote with* $\mathbf{a}$ *the (not necessarily discrete) space of all posible actions.*
> $\boldsymbol{\tau} \subset \boldsymbol{\alpha}$ *is the collection of* terminal action types*. A terminal action, for any actor and parametrization, terminates the scenario.*
> $\mathbf{S}$ *is the (not necessarily discrete) collection of* full states of the scenario $S$*.*
> $\mathcal{F}$ *is the* action impact function $\mathcal{F} : \mathbf{A} \times \mathbf{S} \times \mathbf{a} \to \mathbf{S}$*. We interpret* $S' = \mathcal{F}(A, S, a)$ *as the new full state of the system if actor* $A$ *performs action* $a(\alpha, x_1 \ldots x_n)$ *in state* $S$*.*
> $\mathcal{P} : \mathbf{A} \times \mathbf{S} \to \boldsymbol{\alpha}^*$ *is the* progress function*. We interpret* $\mathcal{P}(A, S) = \{\alpha_{p1}, \ldots, \alpha_{pn}\}$ *as the set of action types available to actor* $A$ *in state* $S$*. If the actor can perform a certain action type, it is free to use an arbitrary parametrization of it.*

### 2.2 The progress model

The progress function $\mathcal{P} : \mathbf{A} \times \mathbf{S} \to \boldsymbol{\alpha}^*$ had been defined on the full state space of the scenario, which, in many cases is very large and not necessarily discrete. However, many human interaction scenarios are *progress-segmented*, that is, the full states can be grouped into equivalence classes with regards to the output of the progress function.

**Definition 2.** *We define* $\mathbf{P} = \{P_1, P_2 \ldots P_n\}$ *the collection of a finite number of* progress states*. A progress state* $P$ *is a (not necessarily discrete) collection of full states, such that if* $S \in P \wedge S' \in P \Rightarrow \forall A \;\; \mathcal{P}(A, S) = \mathcal{P}(A, S')$*. The* progress state discretization *function* $PSD : \mathbf{S} \to \mathbf{P}$ *maps states to progress states.*

**Definition 3.** *We will call the function* $\mathcal{P}_R : \mathbf{P} \times \mathbf{A} \to \boldsymbol{\alpha}^*$ *the* reduced progress function *and define it as* $\mathcal{P}(A, S) = \mathcal{P}_R(A, PSD(S))$.

In contrast to $\mathcal{P}$, the reduced progress function $\mathcal{P}_R$ is defined on a discrete and (usually) small space. We will also consider an even more specific class of scenarios where for every progress state only one actor is allowed to take actions.

**Definition 4.** *We define a* turn taking scenario *a progress-segmented scenario where for any progress state $P$ the reduced progress function $\mathcal{P}_R(A, P)$ is non-empty for only one specific actor $A_t$. We say that $A_t$* has the turn *in progress state $P$.*

### 2.3 A simple example: Human Bargaining

To illustrate the model, let us consider a simple example. In the Human Bargaining scenario two humans, a seller A and a buyer B are arguing over the price of a good. The action type set contains three action types: $\boldsymbol{\alpha} = \{\alpha_O, \alpha_a, \alpha_w\}$ with the following interpretation:

$\alpha_O$ making an offer
$\alpha_a$ accepting the latest offer
$\alpha_w$ withdrawing from the bargaining

The choice of the parameters is a function of the action type, the social context and the goals and limitations of the model. If we assume that the actors are simple software agents, a single parameter (the value of the offer) is sufficient. If, however, the actors are humans many other parameters can be considered: the verbal phrasing of the offer, the politeness of the addressing form used, the tone of the voice, the body language and facial expressions which accompany the offer and so on.

The scenario can be modeled using only four progress states $\mathbf{P} = \{OA, OB, TN, TP\}$:

$OA$ turn of A to take an action
$OB$ turn of B to take an action
$TN$ the bargaining had been broken with no-deal
$TP$ deal accepted

Note that this is a turn taking scenario: in progress state $OA$ only actor $A$ can take an action, in progress state $OB$ only $B$, while $TN$ and $TP$ are terminal states.

In this case the reduced progress function $\mathcal{P}_R$ can be visualized as a *progress graph*, a directed graph where the nodes are progress states and the edges are labeled with the pair of an actor and an action type (see Figure 1).

The progress graph is a helpful modeling tool for the knowledge engineer, but it should not be mistaken for a *full state-action graph* of the scenario.

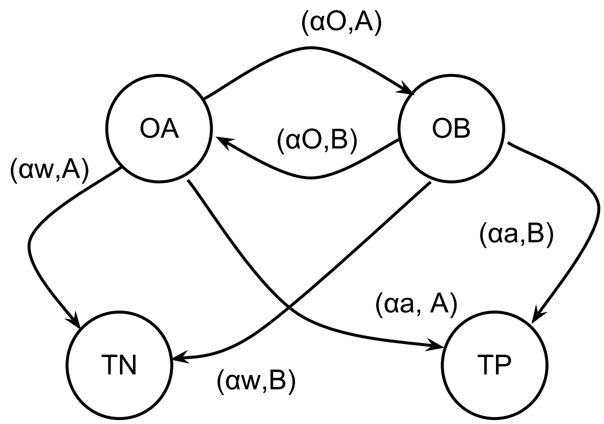


Fig. 1: The progress graph of the Human Bargaining scenario

Such a graph would have full states $S$ as nodes and fully parametrized actions $a(\alpha, A, x_1 \ldots)$ as edges. The full state-action graph is a suitable model for decision theoretic analysis - for instance, it can form the basis of a Markov Decision Process. The progress graph is not sufficient for this.

Normally, the full state includes orders of magnitude more information than the progress state. Even if, A and B are software agents, the full state would have to include the pending offers, the internal valuations of the good by the actors A and B, their negotiation strategies, and possibly other factors such as their models of each other. If A and B are humans, the full state is even more complex: it might include factors such as the level of annoyance of the actors, judgment of personal dignity, the feelings of friendship or animosity towards the negotiating partner, and so on.

We conclude that the full state-action graph is too large for human visual analysis, even for the most simple scenarios such as Human Bargaining. In contrast, the progress graph remains small enough for human intuition for turn taking scenarios, and even for some scenarios which do not verify the turn taking criteria at every progress state.

### 2.4 Group actors and overlapping scenarios

The scenario framework allows us to segment the social-cultural behavior of human actors into manageable entities. An argument can be made, however, that the flow of events in social life is not normally segmented into such clearly defined scenarios. We will further increase the representational power of the model in two ways:

- We will allow *group actors* to model coordinated groups or unorganized crowds of humans. Naturally, the state and the actions taken by the group actors must be compatible with their internal organization.

– We will allow more than one scenario to be executed simultaneously, with some participants (for instance, the crowd) participating in more than one scenario. As the states of the shared participant are part of the full state of both scenarios, this will allow state information to *leak* from one scenario to another.

### 2.5 The next steps

The progress model, as defined above, allows us to describe the general flow of the scenarios without having to consider the full richness of the state characteristic for human interactions. The progress states and the reduced progress function describe the various options actors have at certain moments in the scenario - but they do not have explanatory and predictive (E&P) power. The actors can take any action which is permitted by the reduced progress function, without the ability to predict what they will do and how, nor in retrospect why they did it.

To add E&P power to our model, in the following sections we take a closer look at the full state space $\mathbf{S}$ and the action impact functions $\mathcal{F}$.

In particular we will make the claim that the large majority of human interaction scenarios can be modeled with E&P power while restricting the state to only three types of values: *culture sanctioned social metrics* maintained by individual actors, a small set of *beliefs about concrete facts* (maintained by individual actors) and *perceptions* (maintained by group actors). We will show that these values can be given rigorous definitions, and can be reliably measured and estimated.

## 3 Culture-sanctioned social metrics

### 3.1 Definitions

Human interaction scenarios have complex states - involving the physical, social, cultural, psychological and even physiological aspects of the actors. Relatively obscure circumstances have been shown to have decisively affect the outcome of certain human interaction scenarios (see for instance the work of Kahneman and Tversky about judgement under uncertainty [23]).

We argue, however that in most human scenarios, the cultural expectations reduce the number of choices the participating humans can take, and such scenarios can be explained and reasonable predictions made based on a set of explicit metrics of the state, which are well specified for a given culture and can be readily estimated by the actors. We will call these *culture sanctioned social metrics* (CSSMs). We say that a culture sanctions a metric if it:

– provides a *name* for it
– provides an (informal) *algorithm* for its evaluation
– expects its members to *continuously evaluate* the metrics for themselves and salient persons in their environment

– provides *rules of conduct* which depend on these parameters

The CSSMs can be either tangible or intangible. Tangible metrics such as financial worth or spent time can be measured by physical means (although many times they are only estimated). Intangible metrics, such as politeness, dignity, "face" or friendliness are socially constructed, not directly measurable and depend on the specific culture.

### 3.2 Culture specificity and cross-cultural considerations

CSSMs are always defined by a specific culture, and the name of the metric is given in the language of the culture. Knowing the name of a metric is insufficient: it order to be educated in a culture an individual must know the evaluation algorithm and the rules of conduct associated. It is not required, however, for the individual to *follow* the rules of conduct - however, he or she will be *aware* of the rules and their transgression.

The same name might define different metrics in different cultures. For instance, the word "dignity" has different evaluations and rules of conduct in different English speaking cultures. The dictionary translation of the word in other languages, such as *'azmat'* in Urdu and *'pratistha'* in Hindi, can yield an even more divergent CSSM.

This being said, there are many CSSMs which appear in several cultures in identical or near identical form. There are groups of cultures with closely related metrics - for instance the cultures aligned with the Western European model, the culture of China and nations influenced by Chinese culture and the cultures of the Near East and North Africa. In addition, certain CSSMs are cross-cutting geographical, language and religious boundaries, such as the striking similarities between "cultures of honour" in places as far away as the Scottish highlands, the Bedouins of the Sahara or the Southern USA [18].

It is beyond the scope of this paper to establish specific measures of similarity between different cultures. A number of established metrics in sociology can be used as a starting point for such classification (for instance, Hofstade's cultural dimensions [10]).

### 3.3 The problem of perspective

Many rules of conduct associated with CSSMs consider not only the actor's own perspective, but the perspective of other actors in the scenario. For instance, in the Japanese definition of politeness, it is not sufficient for the actor to act such that its own evaluation of politeness is correct. Instead, the perception of the interaction partner and of external observers is just as (or more) important.

Taking this into consideration, a CSSM is identified by five parameters: *CSSM(C, M, SA, PA, EA)*, where:

– *C* is the culture which defines the CSSM and specifies its rules.

- $M$ is the name of the metric, which is unique in the given culture (but different cultures might mean different metrics under the same name).
- $SA$ is the *subject actor*, the actor who owns the metric.
- $PA$ is the *perspective actor*, from whose perspective the metric is evaluated.
- $EA$ is the *estimator actor*, who estimates the CSSM and who is the owner of the knowledge.

The subject, perspective and estimator actors can be the same. For a CSSM to play a role in a scenario, we need that the $EA$ to be cognizant of culture $C$. In addition, it is necessary for $EA$ to believe that $PA$ is cognizant of culture $C$ (although this belief might be incorrect). It is not necessary for $SA$ to be cognizant of the culture (although whether he is or not might be a factor in the behavior of other actors).

In the following we will present several examples of CSSMs. Our goal will be to illustrate that all the five parameters of the CSSM model are necessary for building a model with E&P power.

**Self perspective** the *CSSM(Western, dignity, John, John, John)* represents John's estimate of its own dignity, in the Western cultural model.

**Peer perspective** the *CSSM(Western, politeness, John, Mary, John)* represents John's estimate about how Mary sees his politeness. If John cares about Mary's opinion, he will adjust its behavior in such a way that Mary's perspective will improve. Note that this value might not be identical to *CSSM(Western, politeness, John, Mary, Mary)*, that is, Mary's own opinion about John's politeness.

**Cross-cultural perspective** Let us consider the case of János, a Hungarian businessman in China, who publicly admits to a business partner Chen a mistake in formulating a purchase order. This will affect *CSSM(Chinese, Face, János, János, Chen)* that is, Chen's estimate of János's own estimate of loosing face. In this context, Chen does not understand why János would do such a thing. What happens here, is that Chen is evaluating a CSSM which János does not: János is not educated in Chinese culture, and the concept of "face" as a metric does not exist in Hungarian. Nevertheless, this particular CSSM can be an important consideration in action - for instance, Chen might act to prevent János from loosing face, even if János is unaware of this.

### 3.4 Intra-cultural uniformity

The multiplication of possible perspectives increases the complexity of the CSSM evaluation. However, this increase is mitigated, in part, by the *intra-cultural uniformity conjecture*.

*Conjecture 1.* Let us consider two human actors A and B, educated in the same culture which defines a CSSM $x$. Let us now consider a scenario $S$ and a series of actions $a_1, \ldots, a_n$, of which both A and B are aware. The *intra-cultural uniformity conjecture* asserts that the evaluation of A and B of the CSSM $x$ will be similar, irregardless of the position of A and B with regards of the scenario (they can be active participants, real-time passive observers or post-facto judges).

This conjecture is supported by the definition of the CSSM: the two actors have the same information and they use the same algorithm for the evaluation provided by the shared culture.

An example of what the conjecture says is as follows: let us consider two Japanese persons, one of them a participant in a social situation with involves interacting with a Westerner while the other one an outside observer. Let us now consider that the Westerner unknowingly commits an action which is considered impolite in the Japanese culture. The intra-cultural uniformity conjecture states that the two Japanese participants will evaluate the level of impoliteness similarly. This fact will not be changed by the fact that the Japanese might also be familiar with the Western culture.

In conclusion, between two participants educated in the same culture, any difference in the evaluation of the CSSM is reduced to the perspective actor's knowledge of specific events. (Naturally, if the estimator actor itself is unaware of an event, it will automatically mean that he or she cannot assign it to the perspective actor either).

### 3.5 The problem of cognitive load

The evaluation of CSSMs is a significant *cognitive load*. Although the culture requires every actor to continuously evaluate all CSSMs for *every* salient person in the environment, in complex situations with many actors present, many actors will not be able to evaluate every possible action impact function.

Different CSSMs, actions and actors will be differently affected by the problem of cognitive load. The more complex a CSSM, the more likely that it will not be estimated. Self perspective CSSMs will more likely be evaluated than cross-cultural peer perspectives. The actors involved in the CSSM also affect their priority. CSSMs where the subject and perspective actors are random members of a crowd will have much lower priority than CSSMs where the subject actor and/or perspective actor is the self or close peers.

Finally, the salience of the action also affects the priority. Striking actions, such as large gestures, loud voice, strong verbal expressions will raise the action's evaluation priority.

### 3.6 Numerical values of CSSMs

For the purpose of our analysis, we will map the CSSMs to a numerical value in the range of $[0.0, 1.0]$. In order to achieve modeling fidelity, this value must be *calibrated*, that is, different numerical values must be associated with the internal perception of the CSSMs (such as degrees of politeness) by the human actors.

The simplest way to perform calibration is by directly requesting numerical values from the human informants. For instance, we can present a human informant with a social situation and ask her to rate the politeness of actor X on a scale from 0 to 100. However, this approach would only work with informants who are comfortable expressing social values on numerical scales.

However, we conjecture that the cultural education with respect to CSSMs is not transmitted through numerical or degree forms, but through narratives containing key words, which can be assigned to various points. It is the task of the knowledge engineer to create a mapping from keywords to numerical values.

## 4 Concrete beliefs

In the previous section we discussed about the CSSMs which are components of the full state given by the culture. These extend across multiple scenarios, and span the life of the actors even outside the specific scenarios. However, in order to model the scenario accurately, we also need to consider beliefs of the agents which are directly relevant to the specific scenario. We find that for achieving E&P power, we don't need to model large sets of belief structures. Rather, we are only interested in the agent's belief about a small number of *concrete questions* which are important for the ongoing scenario. Examples of concrete questions are "Actor X is holding a gun" or "Actors A and B are engaged in a commercial transaction". We will only consider questions which can be answered unequivocally by an omniscient external observer of the scenario. The participants of the scenario, however, need to work with incomplete knowledge and limited rationality, thus they maintain an uncertain answer.

We will call *concrete beliefs* (CBs) the beliefs maintained by the actors in a scenario with regards to the answers of concrete questions. We say that a scenario implies a CB if:

- there is an algorithm which an omniscient external observer of the scenario could use to unequivocally answer the question underlying the CB
- the scenario expects at least one actor to *continuously evaluate* the CB for themselves and salient persons in their environment
- the scenario provides *rules of conduct* which depend on the CB *or* the CB affects the calculation of CSSMs of the actors involved in the scenario

Concrete beliefs can influence a scenario even when they are evaluated from another player's perspective. Taken this into consideration, we will identify a concrete belief with four parameters: *CB(SC, BD, PA, EA)*, where:

- *SC* is the scenario which specifies the question
- *BD* a description of the belief (normally, through the associated question)
- *PA* is the *perspective actor*, from whose perspective the belief is evaluated.

– *EA* is the *estimator actor*, which performs the estimate and who is the owner of the knowledge.

The definition of the CBs parallel the definition of CSSMs, with several important differences. While CSSMs are defined by the culture, the CBs are tied to a specific scenario. The evaluation of the CBs might be incomplete or wrong due to the lack of information or cognitive overload of the estimator actors. The rules associated with CBs might be broken - but the agents estimating the CB will be aware of the fact that they broke a rule. Another different is that the CB does not provide information for the owner actor. While CSSMs are always an internal value represented by the agent, CBs can represent concrete physical values.

### 4.1 Representation formats of CBs

CBs can represent any statement which can be made about the scenario, and thus can take various formats. For instance, for situations involving negotiating about goods, the CB represents beliefs about the private values of different actors. In such situations, the statement can have a scalar or, in the case of multi-issue negotiations, a vector value, and the corresponding CBs will take the form of probability distributions of these values.

Another major type of CBs are cases when the question has a boolean answer, where the CB can be a simple probability number. However, for many applications, it is desirable to use a method which keeps track not only the incidents, while at the same time also maintains a representation of the uncertainty. This is especially important if we do not want to continuously maintain all the evidence for the belief.

Our current approach is based on the Dempster-Shaffer model of evidence with the following assumptions:

- the agent's current beliefs are fully encoded in the mass function
- no previous evidences are remembered
- incoming evidence can be weighted by significance
- at every incoming evidence, the belief is updated using the standard Dempster's rule of combination (conjunctive merge).
- the value for the positive belief is used as the indicator of the belief

## 5 An example of a multi-cultural scenario

Let us now illustrate the way in which the model described in the previous sections can be used to model a simple scenario involving multi-cultural interaction. The Give Way scenario involves two agents A and B approaching simultaneously a door. We assume that the agents are humans, potentially of different cultures, who can have various ages, gender and social status. The scenario also generalizes to situations where one of the agents is a robot.

For each of the agents, there are three different resolutions (1) enter the door first, (2) open and hold the door to the other agent and (3) give way to the other agent to enter first. We assume that the agents do not know each other, they might act under different cultural assumptions, that is they have different CSSMs, with different update rules and associated social requirements. A further complexity can be considered if the scenario happens in the view of the public, in which case the agents also need to consider their estimates of the beliefs of the crowd, in forms of CBs.

Let us now proceed to model the Give Way scenario using the framework developed in the previous sections. The scenario can be modeled with three action types as shown in Table 1. The progress graph, where the nodes are progress states and the edges are labeled with the pair of an actor and an action type is shown in Figure 2). The scenario begins with the start state SS and continues until one of the agents perform the action $\alpha 2$ to reach the terminal state TN.

Table 1: The action types of the Give Way scenario.

| Action type | Description |
|---|---|
| $\alpha 1$ | open the door |
| $\alpha 2$ | enter the door |
| $\alpha 3$ | give way to other agent |

Let us now consider the CSSMs which determine the behavior of the agents in the scenario. Depending on the cultural background of the agents, different set of CSSMs are evaluated. We will consider agent's of two different cultures, Western and Indian. For the purpose of this paper, we will assume that the two participants are of the same gender and they do not have a significant difference of social rank. By and large, politeness considerations in Western culture require the agents to give way to the peer (although this requirement is frequently ignored). In Indian culture, giving way is considered an ineffectual, wimpish behavior.

Thus our model will consider three CSSMs, one concrete (time), and two intangible (politeness and wimpiness). The time T is the amount of time spent on the current scene measured in seconds. Every time taking action $\alpha 3$ by agent imposes a penalty of 5 seconds. In general, agents avoid wasting time.

The politeness is the conformance to the perceived social norms of speech and gestures. Both Western and Indian cultures have the definition of politeness, but there are different definitions associated with them, which translate into different action impact functions. Giving way in the Western culture is considered polite behavior. In the Indian culture however, giving way to a stranger does not impact the perception of politeness. We will consider the private, peer and public politeness aspects: *CSSM(Western, politeness, A, A, A)*, *CSSM(Western, politeness, A, B, A)* and *CSSM(Western, politeness, A, Crowd, A)*.

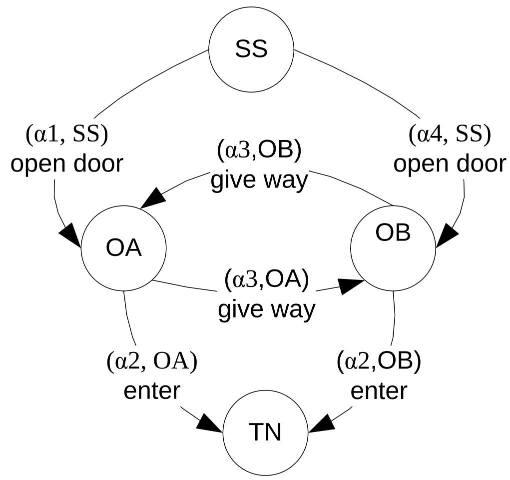


Fig. 2: The progress graph of the Give Way scenario

Wimpiness is the degree of lack of confidence and courage in a person to take initiative. Again, both cultures have definitions for this metric. However, giving way to a person of equal rank does not impact perception of wimpiness in the Western culture, however, it does in the Indian one. We will consider the private, peer and public wimpiness aspects *CSSM(Indian, wimpiness, A, A, A)*, *CSSM(Indian, wimpiness, A, B, A)* and *CSSM(Indian, wimpiness, A, Crowd, A)*.

To achieve E&P power, the analysis of the scenario needs to consider two concrete beliefs, concerning the culture of the two agents, and we need to consider this from the perspective of each other and, potentially, of the crowd. Naturally, *CB(GiveWay, Is-A-an-Westerner, A, A)* is a fixed value, because normally A would know whether he is a Westerner or not. On the other hand, *CB(GiveWay, Is-A-a-Westerner, B, B)*, representing B's belief whether A is an Indian, and *CB(GiveWay, Is-A-a-Westerner, Crowd, Crowd)*, representing the crowd's belief whether A is an Indian are values whose calculation contributes to the E&P power of the model.

Let us now illustrate through several examples the way in which the model traces the evolution of the CSSMs for agents in different cultures. We have modeled the Give Way scenario using our framework, and we traced the evolution of CSSMs for four different sequences of events, each of them representing a different path through the scenario:

$$\text{SS} \xrightarrow{\alpha 1} \text{OA} \xrightarrow{\alpha 3} \text{OB} \xrightarrow{\alpha 2} \text{TN}$$

$$\mathrm{SS} \xrightarrow{\alpha 4} \mathrm{OB} \xrightarrow{\alpha 2} \mathrm{TN}$$

$$\mathrm{SS} \xrightarrow{\alpha 4} \mathrm{OB} \xrightarrow{\alpha 2} \mathrm{TN}$$

$$\mathrm{SS} \xrightarrow{\alpha 1} \mathrm{OA} \xrightarrow{\alpha 3} \mathrm{OB} \xrightarrow{\alpha 2} \mathrm{TN}$$

These scenarios, however, lead to different perceptions and social metrics depending of the culture of the participating humans. We will describe two experiments with different outcomes and cultural backgrounds of the participants.

**Experiment 1:** we consider that both agents A and B belong to the Indian culture. In this case neither of them consider the politeness as a CSSM. As both parties have the same culture, their peer perceptions are a good approximation of the opponents self perception, *i.e. CSSM(Indian, wimpiness, A, A, A)* $\approx$ *CSSM(Indian, wimpiness, A, A, B)*. This allows the agents to make a reasonable prediction of the opponent's actions.

Let us assume that agent B arrives at the door before A, and simply moves on without being polite and giving way to A. The *CSSM(Indian, wimpiness, B, B, B)* will be lowered, while *CSSM(Indian, wimpiness, A, A, A)* will not be affected. In colloquial terms, A can feel himself as a efficient and non-wimpy person, and this can explain its behavior, and can be used to predict it.

**Experiment 2:** let us now consider an experiment in which we don't know the culture of agents A and B, however, we know that the culture of the onlookers (modeled as the Crowd agent) is Indian. If we don't know what culture the agent's belong to, we can simply not trace any CSSMs and CBs for it. We can, however, trace the crowd's belief. Let us consider a scenario where A approaches the door and opens it to agent B. Let us now see how the crowd can reason about this. If A is a westerner, than his politeness level will increase *CSSM(Western, politeness, A, A, Crowd)*. On the other hand, if A is an Indian, his wimpiness will increase *CSSM(Indian, wimpiness, A, A, Crowd)*. As the same action appears rational for a Westerner, but irrational for an Indian, the crowd will treat this occurrence as an evidence which increases *CB(GiveWay, Is-A-a-Westerner, Crowd, Crowd)*.

## 6 Related work

The study of social behavior of humans had been extensively studied in sociology, psychology and anthropology. The models developed in social sciences frequently rely on the understanding of a human observer. Thus, they are usually more qualitative rather than quantitative in nature, although examples of quantitative models exist (such as Hofstede's cultural factors [10]).

The work described in this paper is aligned with an ongoing effort of the autonomous agents and artificial intelligence communities to develop *operational models* of human social behavior. These research efforts are directed both towards a better understanding of social behavior in general, and towards the practical goal of improving the ability of software agents and autonomous robots to act in the presence of humans.

These efforts can be divided in two large categories. One category involves the study of large societies of humans, either being physical crowds or large organizations Bonabeau[3]. This type of study deals with emergent patterns over large number of interactions, and in recent years, had made significant progress through the data mining of information acquired from social networks, such as Twitter, Facebook, Sina Weibo and others.

The literature being very large, we can only consider several representative examples. Kottonau and Pahl-Wostl [12] studied the evolution of political attitudes in response to political campaigns - while in earlier work they studied the problem of new product diffusion. C. Motani et al. [17] implemented a virtual wireless social network based on the information spread in real social network such as a marketplace. Gruhl et al. [7] and Adar et al. [1] analyzed person-to-person information flow over blog space topic sharing. Recent analysis of Twitter followers by Cha et al. [5] had shown that the influence of user on the topic can be gained by a concerted effort over a long period of time and a large number of followers are not an assurance to fame.

A significant amount of research had been directed towards the epidemic propagation of information in social networks [19, 13, 14]. In these papers, the information spread is modeled as virus infection in computer networks.

The second category of social modeling involves singular interactions, but a higher detail model. An example of argument towards such models is the KIDS (Keep it Descriptive Stupid) approach advocated by Edmonds and Moss [6]. The work we describe here deals with this type of interactions.

An example of similar work involves the work of Miller et al. In [15] propose to operationalize the Brown and Levinson politeness model [4]. The implementation, the Etiquette Engine, is used to assess the politeness of a number of custom crafted social-interaction vignettes involving common culture but different rank (the interaction between a corporal and a mayor). The values were compared against the evaluation by human observers (unfamiliar with the Brown and Levinson model). In a follow-up work [16] they create a more complex model which investigates the relationship between culture (as examplified by Hofstede's cultural factors [10]) as well as politeness levels affect directive compliance. Directive compliance - the way in which people react to instructions, commands or requests, represent a very large proportion of human interactions in work and military settings.

The HAIDM workshop had been the venue of the presentation of a number of contributions in this area. Ramchurn et al. [21] deals with the representation of social and ethical issues in applications such as agile teaming, incentive engineering and flexible autonomy. Haim et al. [8] develop models of human behavior in the setting of a negotiation game with participants from different backgrounds (US, Lebanon and Israel), and develop an agent behavior called PAL (personality adaptive learning) which predicts the human adversaries behavior in a probabilistic manner. Salvit and Sklar [22] model human personality traits using the Myers-Briggs model, and integrate it in a BDI agent architecture.

Harriott et al. [9] describes an empirical study of human-robot teams using performance moderator functions (HPMF) which predict performance under various factors such as fatigue, mental workload or temperature.

Aylett et al. [2] describes a believable agent-based educational application based on emotions, personality and culture, where an agent is provided with a description of the symbols and rituals of a certain society. The application implements virtual reality characters of a fictional alien culture, while the users must adapt to their cultural behavior.

## 7 Conclusions

In this paper we described a modeling framework for reasoning about the social-cultural behavior of humans. The framework is currently implemented as a Java software library, and is a basis of our work in designing behaviors for autonomous robots which need to act in social settings. Our ongoing work is directed towards the modeling of real-life scenarios of robot deployments in various social settings, such as BigDog [20] class robots assisting a peacekeeping team in a foreign country. We are considering both scenarios where the objective is for the robot must act in socially acceptable ways, and scenarios where the objective is to pursue specified goals while considering the predicted social behavior of human interaction partners and bystanders.

**Acknowledgements:** The research reported in this document/presentation was performed in connection with Contract Number W911NF-10-2-0016 with the U.S. Army Research Laboratory. The views and conclusions contained in this document/presentation are those of the authors and should not be interpreted as presenting the official policies or position, either expressed or implied, of the U.S. Army Research Laboratory, or the U.S. Government, unless so designated by other authorized documents.

## References

1. E. Adar and L. Adamic. Tracking information epidemics in blogspace. In *Proc. of 2005 IEEE/WIC/ACM International Conference on Web Intelligence*, pages 207–214. IEEE, 2005.
2. R. Aylett, N. Vannini, E. Andre, A. Paiva, S. Enz, and L. Hall. But that was in another country: agents and intercultural empathy. In *Proc. of the 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS-2009)*, pages 329–336, 2009.
3. E. Bonabeau. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America*, 99(Suppl 3):7280, 2002.
4. P. Brown and S. Levinson. *Politeness: Some universals in language usage*, volume 4. Cambridge Univ Pr, 1987.
5. M. Cha, H. Haddadi, F. Benevenuto, and K. Gummad. Measuring user influence on twitter: The million follower fallacy. In *Proc. of the Fourth Int'l AAAI Conference on Weblogs and Social Media*, pages 10–17. ACM, 2010.

6. B. Edmonds and S. Moss. From KISS to KIDS–an 'anti-simplistic' modelling approach. *Multi-Agent and Multi-Agent-Based Simulation*, pages 130–144, 2005.
7. D. Gruhl, R. Guha, D. Liben-Nowell, and A. Tomkins. Information diffusion through blogspace. In *Proc. 13th Internat. World-Wide Web Conference*. ACM New York, 2004.
8. G. Haim, Y. Gal, M. Gelfand, and S. Kraus. A cultural sensitive agent for human-computer negotiation. In *Proc. of the Eleventh Int. Conf. on Autonomous Agents and Multiagent Systems(AAMAS-2012)*, pages 451–458, 2012.
9. C. Harriott, R. Zhuang, J. Adams, and S. DeLoach. Towards using human performance moderator functions in human-robot teams. In *Proc. of the Human-Agent Interaction Design and Models Workshop*, 2012.
10. G. Hofstede, G. Hofstede, and M. Minkov. *Cultures and organizations: software for the mind*. McGraw-Hill Professional, 2010.
11. S. A. Khan, T. Singh, and L. Bölöni. Learning social calculus with genetic programing. In *Twenty-Sixth Florida Artificial Intelligence Research Society Conference (FLAIRS-26)*, 2013.
12. J. Kottonau and C. Pahl-Wostl. Simulating political attitudes and voting behavior. *Journal of Artificial Societies and Social Simulation*, 7(4), 2004.
13. Z. Liu and B. Hu. Epidemic spreading in community networks. In *Europhys. Lett. 72 315*. epl, 2005.
14. R. May and A. L. Lloyd. Infection dynamics on scale-free networks. *PhysRevE*, 64(6):066112, 2001.
15. C. Miller, P. Wu, and H. Funk. A computational approach to etiquette: operationalizing Brown and Levinson's politeness model. *IEEE Intelligent Systems*, pages 28–35, 2008.
16. C. Miller, P. Wu, V. Vakili, T. Ott, and K. Smith. Culture, politeness and directive compliance. In *Proc. of the 5th International Conference on Universal Access in Human-Computer Interaction*, pages 568–577, July 2009.
17. M. Motani, V. Srinivasan, and P. Nuggehalli. PeopleNet: Engineering a wireless virtual social network. In *Proc. of ACM MobiCom 2005*. ACM, 2005.
18. R. Nisbett and D. Cohen. *Culture of honor: The psychology of violence in the South*. Westview Press, 1996.
19. R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scale-free networks. *PhysRevLett*, 86(14):3200–3203, 2001.
20. M. Raibert, K. Blankespoor, G. Nelson, R. Playter, et al. Bigdog, the rough–terrain quadruped robot. In *International Conference of Automatic Control*, pages 10823–10825, 2008.
21. S. Ramchurn, N. Jennings, L. Moreau, S. J. Roberts, T. Rodden, and A. Rogers. ORCHID: Developing the science of human-agent collectives. In *Proc. of the Human-Agent Interaction Design and Models Workshop*, 2012.
22. J. Salvit and E. Sklar. Modulating agent behavior using human personality type. In *Proc. of the Human-Agent Interaction Design and Models Workshop*, 2012.
23. A. Tversky and D. Kahneman. Judgment under uncertainty: Heuristics and biases. *science*, 185(4157):1124–1131, 1974.