Emotion Contagion in a Cognitive Architecture

Joseph Allen
Department of Cognitive Science
Rensselaer Polytechnic Institute
Troy, NY 12180, USA

Ron Sun, Ph.D, IEEE Fellow
Department of Cognitive Science
Rensselaer Polytechnic Institute
Troy, NY 12180, USA

Abstract— **The issue of emotion contagion in various settings has been gaining attention recently. Humans share emotions, for example, through gestures or speech or through online text via social media. There have been computational models in the past to try to capture emotion contagion. However, these models are limited as they tend to represent human agents mainly through simple equation-based or rule-based systems. We use a more psychologically realistic model --- the Clarion cognitive architecture --- as the basic building block to model emotion and emotion contagion. In particular, we use Clarion to capture and explain human data from a human experiment on emotion contagion.**

Keywords—emotion, cognitive architecture, contagion

I. Introduction

The issue of emotion contagion in various settings has been gaining attention in recent years. Evidently, humans do share emotions, and emotions spread from individuals to individuals. Emotion contagion occurs, for example, through gestures and speech or through online text via social media. Such phenomena have been documented in the literature, for example, by Hatfield, Cacioppo, and Rapson (1993), Kramer et al. (2014), and so on. Despite their differences, collectively, they are referred to as emotion contagion.

Thus far, there have been computational models proposed that may be used to capture emotion contagion. However, these models are limited as they tend to represent agents in an overly simplified way, for example, through simple equation-based or rule-based methods (Parunak et al., 2014; Deffuant, Weisbuch, Neau, and Amblard, 2000).

In this paper, we use a more sophisticated model --- the Clarion cognitive architecture --- as the basic building block to model emotion contagion. Each individual is modeled by a computational cognitive architecture, and they interact with each other so that emotion contagion occurs from individual to individual. We capture and explain computationally some human data from an experiment on human emotion contagion.

Below, we will first describe briefly the issues concerning emotion and emotion contagion, including empirical findings of emotion contagion. We will then describe the Clarion cognitive architecture, which serves as the basis for capturing emotion contagion. In the process, we will describe the different subsystems and their importance in modeling human behavior. We will then discuss how emotion is represented in Clarion and how Clarion can capture emotion contagion. A simulation will be described that illustrates the working of the model.

II. Emotion and Contagion

The term "emotion" has been used to indicate a variety of phenomena. Human emotion often includes a complex of experiential, behavioral, cognitive, motivational, psychological, and physiological characteristics, with many underlying mechanisms and processes. Emotion is, to say the least, a set of complex and multi-faceted phenomena.

We have argued (see, e.g., Sun and Mathews, 2012 and Sun, Wilson and Lynch, 2016) that human emotion should be computationally captured and explained by a broad computational cognitive architecture (Sun, 2002, 2016), using psychologically realistic mechanisms and processes. That is, emotion should be captured and explained based on adequate representation of basic action decision making, reasoning, motivation, metacognition, and other psychological processes, within a comprehensive, computational model of the mind. These mechanisms and processes can capture the interaction of internal needs and external situations in determining motivations and actions by individuals (and hence emotions).

Among other models of emotion, a detailed computational model of emotion was proposed based on the Clarion cognitive architecture, using its existing mechanisms and processes, thus integrating emotion fully into the operations of the cognitive architecture. See, for example, Sun et al. (2016), Sun and Mathews (2012), Wilson and Sun (2014), and Wilson (2012). To address emotion contagion, we need to extend the Clarion model of emotion to take into account results of social influences.

III. The Clarion Cognitive Architecture

A. Overview of Clarion

Clarion is a broad computational model of essential psychological mechanisms and processes. It has been described, justified, and psychologically validated extensively on the basis of psychological data; see Sun (2002, 2016); see also Sun et al. (2001, 2005) and Helie and Sun (2010). A quick sketch is provided below.

Clarion includes a number of subsystems: the action-centered subsystem (the ACS), the non-action-centered subsystem (the NACS), the motivational subsystem (the MS), and the metacognitive subsystem (the MCS). The role of the action-centered subsystem is to control actions (regardless of whether the actions are for external physical movements or for internal mental operations), utilizing procedural knowledge. The role of the non-action-centered subsystem is to utilize declarative knowledge for information and inferences of various kinds. The role of the motivational subsystem is to provide underlying motivations for perception, action, and cognition. The role of the metacognitive subsystem is to regulate the operations of the other subsystems dynamically. See Sun (2016) for detailed arguments regarding the division of subsystems.

Each of these interacting subsystems consists of two "levels" of representations, that is, a dual-representational structure, as theoretically posited in Sun (2002). Generally speaking, in each subsystem, the "top level" encodes *explicit* knowledge with associated explicit processes (using symbolic/localist representations), while the "bottom level" encodes *implicit* knowledge with associated implicit processes (using connectionist distributed representations; Rumelhart et al., 1983). [1] The two levels interact, for example, by cooperating in action decision making through integration of the action recommendations from the two levels of the action-centered subsystem respectively. See Figure 1 for a sketch of Clarion**.**

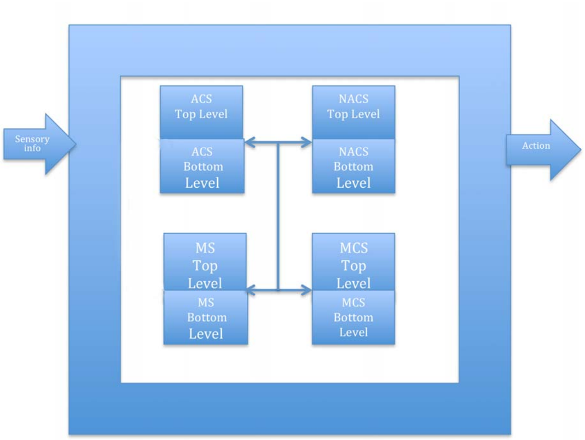

Figure 1: The Clarion cognitive architecture.

As has been pointed out before, existing theories tend to confuse implicit and explicit processes; hence the "perplexing complexity" (Smillie et al., 2006). In contrast, Clarion generally separates implicit and explicit processes in each of its subsystems. With this framework, Clarion can provide explanations of empirical findings in a wide range of domains (see, e.g., Sun et al., 2001, 2005; Helie and Sun, 2010; Wilson and Sun, 2014). Another important characteristic of Clarion is its focus on the cognition-motivation-environment interaction (Sun, 2009, 2016), as opposed to dealing only with cognition in its narrow sense.

B. The action-centered subsystem

The action-centered subsystem (the ACS) captures the action decision making of an individual when interacting with the world, involving procedural knowledge (Sun, 2002).

In the action-centered subsystem, the process for action selection is essentially the following: Observing the current (observable) state of the world, the two levels within the action-centered subsystem (implicit or explicit) make their action decisions in accordance with their respective (procedural) knowledge, and their outcomes are "integrated". Thus, a final selection of an action is made and the action is then performed. The action changes the world in some way. Comparing the changed state of the world with the previous state somehow, the person learns. The cycle then repeats itself. See Sun (2002, 2016) for algorithmic details.

In this subsystem, the bottom (implicit) level is implemented using neural networks (Rumelhart et al., 1986), and the top level is implemented using symbolic/localist representations.

For the bottom level, the input state (x) consists of the sensory input (environmental or internal), the current goal, and the working memory. All that information is important in deciding on an action. The input state is represented as a set of microfeatures. The output of the bottom level is the action choice, also represented as a set of microfeatures.

At the top level, "chunk" nodes are used for denoting concepts. A chunk node connects to its corresponding microfeatures at the bottom level (represented by a set of separate nodes, constituting a distributed representation in the bottom level). At the top level, action rules connect chunk nodes representing conditions to chunk nodes representing actions. If the condition of an action rule is met, then the corresponding action is recommended. See Sun (2002, 2016) for implementation details.

At the bottom level of the action-centered subsystem, with neural networks encoding implicit procedural knowledge, actions are selected based on their Q values, which are the outputs of the neural networks. A Q value is an evaluation of the "quality" of an action in a given input state: $Q(x, a)$ indicates how desirable action a is in state x. At each step, given input state x, the Q values of all the actions (i.e., $Q(x, a)$ for all a's) are computed in parallel. Then the Q values are used to decide stochastically on an action to be performed, through a Boltzmann distribution of Q values (i.e., a softmax function; Sun, 2002).

For learning implicit knowledge at the bottom level (i.e., the Q values), the Q-learning algorithm (Watkins, 1989), which is a reinforcement learning algorithm, may be used (Sun et al., 2001; Sun, 2002, 2016). At a result of such learning, the Q values come to represent, roughly, the maximum cumulative reinforcement that can be received from the current point on (Watkins, 1989), where reinforcement represents the fulfillment of needs and achievement of goals (as decided by

[1] Roughly speaking, explicit knowledge is directly consciously accessible (i.e., conscious or potentially conscious), while implicit knowledge is consciously inaccessible directly. Explicit processes involve explicit knowledge, while implicit processes involve implicit knowledge. See Sun (2002, 2016) for details.

the motivational subsystem and the metacognitive subsystem; more later).

For learning explicit action rules at the top level, a variety of learning algorithms may be used. See Sun (2002, 2016) for details.

C. The non-action-centered subsystem

The non-action-centered subsystem (the NACS) is for dealing with declarative, or non-action-centered, knowledge (Sun, 2016). It stores such knowledge in a dual representational form (the same as in the action-centered subsystem): that is, in the form of explicit "associative rules" (at the top level), and in the form of implicit "associative memory" (at the bottom level).

With the interaction of the two levels, the non-action-centered subsystem carries out rule-based, similarity-based, and constraint-satisfaction-based reasoning. Details of the non-action-centered subsystem will not be covered here because they are not needed in this paper (see, e.g., Sun, 2016 or Helie and Sun, 2010).

D. The motivational subsystem

The motivational subsystem (the MS) is concerned with why an individual does what he/she does (Sun, 2009). The relevance of the motivational subsystem to the action-centered subsystem lies primarily in the fact that it provides the context in which goals and reinforcements for the action-centered subsystem are determined. It thereby influences the working of the action-centered subsystem.

The motivational subsystem involves a dual motivational representation. The explicit *goals* (such as "*find food*"), which are essential to the working of the action-centered subsystem as explained before, may be generated based on implicit *drives* (e.g., "*hunger for food*"). For psychological justifications, see Sun (2009, 2016); see also Murray (1938) and Maslow (1987).

Among drives, *primary drives* are essential, most likely built-in (hard-wired) to a significant extent to begin with. Some sample low-level primary drives include *food, water,* and *reproduction*. Beyond those concerning mostly physiological needs, there are also high-level primary drives: for example, *dominance, fairness*, and *deference*, (Murray, 1938; Reiss, 2004; Sun, 2009). The primary drives (both low-level and high-level) are listed in Table 1. This set of primary drives has been extensively justified before (Sun, 2009; see also see Murray, 1938; Maslow, 1987; Reiss, 2004).

Note that some of these primary drives are approach-oriented, while others are avoidance-oriented. This distinction has been argued by many before (e.g., Smillie, Pickering, and Jackson, 2006; Sun, 2016).

Table 1. Primary drives, including both approach-oriented and avoidance-oriented primary drives.

Approach Drives	**Avoidance Drives**	**Both**
Food	*Sleep*	*Affiliation & Belongingness*
Water	*Avoiding Danger*	*Similance*
Reproduction	*Avoiding Unpleasant Stimuli*	*Deference*
Nurturance	*Honor*	*Autonomy*
Curiosity	*Conservation*	*Fairness*
Dominance & Power		
Recognition & Achievement		

The processing of the drives involves determining the strength of a drive, which is done using a neural network based on the product of the input *stimulus* and the internal "*deficit*" (which represents an individual's intrinsic sensitivity and inclination toward activating a drive). The justifications for this may be found in a variety of literatures (e.g., Sun, 2009).

E. Metacognitive Subsystem

In Clarion, the metacognitive subsystem (the MCS) is closely tied to the motivational subsystem (the MS). The metacognitive subsystem monitors, controls, and regulates cognitive processes. Control and regulation may be in the forms of setting goals (which are then used by the action-centered subsystem) on the basis of drives, generating reinforcement signals for the action-centered subsystem for learning (on the basis of drives and goals), interrupting and changing on-going processes in the action-centered subsystem and the non-action-centered subsystem, setting essential parameters of the action-centered subsystem and the non-action-centered subsystem, and so on (Sun, 2016).

Structurally, this subsystem may be divided into a number of functional modules corresponding to these functions above.

For example, the goal module, in order to select a new goal, first determines goal strengths for relevant goals, based on information from the motivational subsystem (e.g., the drive strengths) and the current sensory input. For instance, a simple case is the calculation of the strength of a goal based on the sum of the products of the strength of an activated drive and the *relevance* of the drive to the goal. Once calculated, the goal strengths are turned into a Boltzmann distribution and the new goal is chosen stochastically from that distribution. Arguments in support of goal setting on the basis of implicit motives (i.e., drives) may be found in the literature (e.g., as described in Sun, 2016)

IV. MODELING EMOTION CONTAGION IN CLARION

A. Emotion in Clarion

Emotion may roughly be considered in terms of three different aspects: reactive affect, deliberative appraisal, and coping/action (Wilson, 2012; Wilson and Sun, 2014; Sun et al, 2016).

First, reactive affect is a fast, reactive component that precedes other, slower components of emotion (Zajonc, 1980; Sun and Mathews, 2012). Emotion-evoking stimuli may impact an individual before (or without) conscious awareness (e.g., Winkielman, Berridge, & Wilbarger, 2005), separate from slower, more deliberative processes. It is fast, reactive, and implicit (as is consistent with the Clarion framework described earlier). Reactive affect may be either positive or negative in valence, and to different degrees of intensity (i.e., arousal; Wilson, 2012).

Deliberate appraisal is the evaluation of the significance of an event or a situation (Lazarus, 2001) on the basis of various inputs including reactive affect. Emotions are shaped by such cognitive evaluation of situations and events. Deliberative appraisal also determines how emotions are consciously registered and reported (i.e., emotion terms/labels), and recommends possible goals and actions to pursue.

Coping is cognitive and behavioral efforts at managing specific external and/or internal demands. Coping often follows appraisal (i.e., after ascertaining the meanings of the situation; Lazarus, 2001).

Note that, according to the CLARION framework, which is motivationally based, emotion is rooted in basic human motives or needs (i.e., drives) and their possible fulfillment (Sun and Mathews, 2012). Smillie et al. (2006) and Ortony et al. (1988) also stressed the importance of motivation and expectation in emotion.

B. *Details of Emotion in Clarion*

Let us explore further these three aspects within the Clarion cognitive architecture.

First, reactive affect is determined by the reactive affect module of the metacognitive subsystem (Wilson, 2012) based on motivation as well as potentials for actions. According to the literature, it may have a lot to do with the relationship between situations/events and an individual's desires and intentions (i.e., drives and goals). The motivational subsystem captures the role of motivation in reactive affect, with both drives (at the bottom level) and goals (at the top level).

On the other hand, within Clarion, the bottom (implicit) level of the action-centered subsystem contains neural networks (Sun, 2002, 2016). These networks propagate activation from input nodes representing input state information to output nodes representing actions. Before the final action decision is made, the activations of the output nodes represent potentials to act, or "action potentials". These activations capture the degree and likelihood that the actions will lead to a desirable outcome (see section 3). Fundamentally, actions are chosen based on their potentials to satisfy needs (Maslow, 1987). Thus, the "action potential" represents the expected degree to which a set of actions (that can be started in the current state) will be successful in attending to the needs of an individual.

Positive and negative reactive affect can result from a combination of drive activation and action potential, as follows:

- When some drives are highly activated (i.e., some needs must be attended to), but action potentials are also high (i.e., those needs are likely to be met), then affect could not be very negative --- it could range from highly positive to slightly negative.
- When some drives are highly activated (i.e., some needs are high) and action potentials are low (i.e., those needs will likely not be met), affect could range from slightly negative to very negative.
- When drive activations are low (i.e., needs have been met) and action potentials are high (i.e., needs will likely continue to be met), then affect should be in the positive range.
- When drive activations are low (i.e., needs have been met) and action potentials are low (i.e., needs may not continue to be met), then affect should not be very positive but should not be very negative either.

The reactive affect module of the metacognitive subsystem determines reactive affect through a function of a drive activation and its resulting action potentials.

In Clarion, the non-action-centered subsystem is for storing and utilizing declarative knowledge in various forms (Sun, 2016; Helie & Sun, 2010). The non-action-centered subsystem of Clarion provides the mechanistic means by which various aspects of deliberative appraisal may be actualized. See Wilson (2012) for details.

Coping/action occurs on the basis of regulation by the metacognitive subsystem. As described before, the metacognitive subsystem captures many regulatory functions. Goals may be recommended by deliberative appraisal from processes within the non-action-centered subsystem. In the metacognitive subsystem, in order to take account of deliberative appraisal in setting goals, goal strengths are determined by both drive activations and the outcomes of deliberative appraisal. In other words, goals are set based not only on their relevance (as described in section 3), but also based on appraisal regarding their feasibility and benefits (see Wilson, 2012 for the extended equation).

Once goals are selected, those behaviors that best facilitate goal achievement are likely to be selected by the action-centered subsystem. The action selection mechanisms are well defined in Clarion (as briefly reviewed in section 3).

C. *Details of Emotion Contagion in Clarion*

Based on the model of emotion above, there are two routes for emotion contagion in Clarion. One is through input to drive activation (and then to goals and actions), with other individuals' emotional states entering through input stimulation to drives (within the motivational subsystem) and reactive affect being determined based on drive activation and action potential (by the metacognitive subsystem). The other route is directly through the metacognitive subsystem, which assesses affect based on drive activation and action potential but also based on simply incorporating other individuals' emotional states (which constitute part of the input state).

With the first route above, other individuals' behaviors (which represent their emotional states) constitute part of the input state to an individual. Thus, such emotional states influences the activation of drives within the motivational subsystem. In turn, these activated drives, along with action potentials, determine the reactive affect that the individual experiences (determined by the metacognitive subsystem), as well as his/her cognitive appraisal (by the non-action-centered subsystem), action/behavior (within the action-centered subsystem), and so on.

V. Simulation of Human Data

A. A Human Experiment on Emotion Contagion

In the experiment of Barsade (2002), a managerial exercise took place in the form of a leaderless group discussion. Each trial consisted of a group of 5 subjects. Each subject was to act as a manager participating in a meeting of the salary committee. The committee needed to hand out a limited sum of bonus money. Each manager was supposed to represent a candidate from his/her own department. Each subject gave a two-to-three minute presentation on behalf of his/her candidate. Each subject was given two goals: (1) to give out the resources in a way that would benefit the company as a whole, and (2) to get as many resources as possible for their own candidate.

Before the exercise started, the subjects were handed written materials that gave context to the discussion. In each group, there was a confederate actor who would display one of four different emotional states throughout the trial: cheerful, serene, depressed, and hostile. That is, an emotional state was selected from a two-by-two table of valence (positivity/ pleasantness) by intensity, with each dimension having two values (low vs. high). The confederate made a presentation first, followed by presentations by subjects.

Before and after the experiment, subjects were given a seven-point Lichert scale test about their current mood to see if there was any group wide contagion towards the confederate's mood.

In the work of Barsade (2002), there were three important variables: (1) the confederate's affect (pleasantness); (2) a subject's self-reported affect (pleasantness) before trial, denoted as "T1"; (3) a subject's self-reported affect (pleasantness) after trial, denoted as "T2".

From these, the original findings were: (1) the difference between the affect before trial and after trial was correlated with the confederate's affect; in other words, subjects' affect was indeed influenced by that of the confederate; (2) the difference between the affect before trial and after trial was negatively correlated with the affect before trial; in other words, it was not influenced by the affect beforehand; (3) the difference between the affect before trial and after trial was correlated with the affect after trial. See Table 2. Together, the results demonstrated emotion contagion.

Table 2. The human data from Barsade (2002).

Human Results:	
T1-T2 correlation with Confederate Affect	0.31
T1-T2 correlation with T1	-0.56
T1-T2 correlation with T2	0.51
($p < 0.05$, 0.05, 0.05 respectively)	

B. Simulation Setup

The simulation of this experiment focused on reactive affect and contagion through drive activation (as described earlier). Deliberative appraisal was not involved.

For this simulation, six drives were included in the motivational subsystem, which were relevant based on their definitions in Clarion (Sun, 2009):

- *Dominance and Power* – To influence others. It is common that people want to dominate a conversation much like when people consider a debate more about winning than about truth seeking.
- *Affiliation and belongingness* – To form bonds and relationships. One may give a presentation to please other people.
- *Autonomy* – To resist influence and coercion. One may prefer to provide a different opinion.
- *Similance* – To identify oneself with others; try to behave like others. They may try to do well for the company as a whole.
- *Fairness* – To behave appropriately, often reciprocally. This leads to trying to gain a proportional amount of resources for one's own candidate.
- *Honor* – To obey social norms and codes. There could be "company" social norms as well as wider cultural norms.

These drives had activation levels between 0 and 1.

Before each trial, each agent's drive activations were randomized (to represent somewhat what happened after the materials handed out in the original experiment). To start the trial, the confederate selected its emotional state and then gave its presentation. The other agents followed with their presentations. Every time a presentation was given, the other agents' drives were stimulated and activated accordingly.

In determining the drive strength, the *deficit* parameter represented individual differences and so was chosen randomly between 0 and 1. Different kinds of presentations had different effects on drive activations. The mapping from perceived presentations to the *stimulus* parameter (to the drive) was as shown in Table 3.

Table 3. The mapping from perceived presentation to drive stimulus

Forceful presentation –	
Dominance and Power	.7
Autonomy	.3
Honor	.3
Pleasant presentation –	
Affiliation and Belongingness	.7
Similance	.3
Honor	.3
Fairness	.1

Accurate presentation –	
Fairness	.7
Autonomy	.1
Similar Presentation –	
Affiliation and Belongingness	.3
Similance	.7
Fairness	.3
Honor	.5
Normative Presentation –	
Fairness	.5
Honor	.7
Autonomy	.3
Different Presentation –	
Dominance and Power	.5
Autonomy	.7

Once the drives were activated (using these parameter settings specified above), an agent selected a goal through the goal module of the metacognitive subsystem, determined by how relevant a goal was to the activated drives. The *relevance* parameter (see section 3) was set between 0 and 1 (as indicated by Table 4).

Table 4. Relevance of a drive to a goal.

Win argument –	
Relevance with Dominance and Power	-8
Relevance with Fairness	.3
Relevance with Honor	.3
Be accurate –	
Relevance with Fairness	8
Relevance with Similance	.2
Relevance with Honor	.4
Be polite –	
Relevance with Affiliation and Belongingness	8
Be similar –	
Relevance with Similance	.8
Relevance with Affiliation and Belongingness	.2
Be different –	
Relevance with Autonomy	.8
Relevance with Dominance and Power	.3
Follow norms –	
Relevance with Honor	.8
Relevance with Fairness	.3

Once a goal was selected, the action-centered subsystem chose an action in accordance with that goal. The action-centered subsystem was trained through reinforcement learning to choose the right actions most of the time. The pairings that were reinforced during training included: if the goal was *win argument*, then the action should be *forceful presentation*; if *be polite,* then *pleasant presentation;* if *be accurate,* then *accurate presentation;* if *be similar,* then *similar presentation;* if *follow norms*, then *normative presentation*; if *be different*, then *different presentation*. Once the action was selected, the corresponding presentation was then made to the other agents, which in turn affected their drives, goals, and actions.

To determine reactive affect, the reactive affect module of the metacognitive subsystem took each of the drives and determined its affect value through a six-by-three-by-three table. The drive strength was low (<0.33), medium (<0.66), or high (>0.66), for each of the six drives, and the action potential was also low, medium, or high. For example, if the agent has low *Affiliaton and Belongingness* drive value and its action potentials are also low, we would expect a low affect magnitude. See Table 5 for further details. Note that for drives such as *Dominance and Power*, *Affiliation and Belongingness*, and *Similance* (i.e., mostly approach-oriented drives), action potentials were trained to be medium to high; for drives such as *Honor, Autonomy*, and *Fairness* (mostly avoidance-oriented drives), action potentials were trained to be low to medium. Once the affect value for each drive was obtained by the metacognitive subsystem, the minimum negative value and the maximum positive value were summed. The resulting reactive affect for each agent was then reported.

Table 5. The reactive affect table, in the form of "Drive Name, Drive Value, Action Potential Value: Reactive Affect Value".

Affiliation and Belongingness:	
AandB, low, low	.05;
AandB, low, med	.1;
AandB, low, high	1.5;
AandB, med, low	-.05;
AandB, med, med	.05;
AandB, med, high	.3;
AandB, high, low	-.1;
AandB, high, med	.3;
AandB, high, high	.7;
Autonomy:	
auto, low, low	.05;
auto, low, med	.1;
auto, low, high	1.5;
auto, med, low	-.1;
auto, med, med	.3;
auto, med, high	.5;
auto, high, low	-.3;
auto, high, med	.3;
auto, high, high	.7;
Similance:	
sim, low, low	.05;
sim, low, med	.1;
sim, low, high	1.5;
sim, med, low	-.05;
sim, med, med	.05;
sim, med, high	.3;
sim, high, low	-.1;
sim, high, med	.3;
sim, high, high	.7;
Fairness:	
fair, low, low	.05;
fair, low, med	.1;
fair, low, high	1.5;
fair, med, low	-.2;
fair, med, med	.4;
fair, med, high	.6;
fair, high, low	-.5;
fair, high, med	.4;
fair, high, high	.8;
Dominance and Power:	

DandP, low, low	.05;
DandP, low, med	.1;
DandP, low, high	1.5;
DandP, med, low	-.3;
DandP, med, med	.4;
DandP, med, high	.5;
DandP, high, low	-.8;
DandP, high, med	-.1;
DandP, high, high	.8;
Honor:	
hon, low, low	.05;
hon, low, med	.1;
hon, low, high	1.5;
hon, med, low	-.3;
hon, med, med	.4;
hon, med, high	.5;
hon, high, low	-.8;
hon, high, med	-.1;
hon, high, high	.8;

Other details were as described in sections 3 and 4 before. Note that in Clarion, there are, of course, many more variables and mechanisms, but they are not relevant to this particular simulation.

C. *Simulation Results*

In the original human experiment of Barsade (2002), there were three important variables: the confederate's affect (pleasantness); a subject's affect before trial---the "T1" affect; a subject's affect after trial---the "T2" affect. From these, three of the original findings were: the difference between the affect before trial and after trial was correlated with the confederate's affect; the difference between the affect before trial and after trial was negatively correlated with the affect before trial; the difference between the affect before trial and after trial was correlated with the affect after trial.

We were able to simulate these measures: the "pleasantness" measure used in the experiment was assessed through the positivity of the reactive affect described earlier. Thus we captured these findings regarding correlations through our simulation.

Specifically, after running our simulation, the following results were found: (1) the difference between the affect before trial and after trial was correlated with the confederate's affect; therefore, simulated subjects' affect was indeed influenced by that of the confederate; (2) the difference between the affect before trial and after trial was negatively correlated with the affect before trial; in other words, it was not influenced by the affect beforehand; (3) the difference between the affect before trial and after trial was correlated with the affect after trial. See Table 6 for a comparison of human and simulation results.

Table 6. Simulation results, compared with the human data from Barsade (2002).

Simulation Results:	
T1-T2 correlation with Confederate Affect	0.41
T1-T2 correlation with T1	-0.67
T1-T2 correlation with T2	0.74
($p < 0.01$, 0.001, 0.001 respectively)	

Human Results:	
T1-T2 correlation with Confederate Affect	0.31
T1-T2 correlation with T1	-0.56
T1-T2 correlation with T2	0.51
($p < 0.05$, 0.05, 0.05 respectively)	

The simulation results seem to capture the human results well. Based on this result, it is reasonable to claim that emotion contagion was indeed taking place in the simulation and the affect of the simulated subjects moved toward the confederate's affect.

VI. Concluding Remarks

In this paper, a model of emotion and emotion contagion based on a comprehensive computational cognitive architecture, namely Clarion, was described. To demonstrate the model, an existing experiment was utilized, and the Clarion-based model simulated and captured this experiment. Clarion provided a realistically complex model for the agents involved in this experiment.

Through conducting the simulation, we specified in concrete terms many details of the agent, including various subsystems of the agent, which together gave a more thorough explanation than what a simple equation-based or rule-based model could provide. Overall, our simulation was able to produce behavior highly similar to what was found in the original human experiment.

Acknowledgment

We thank Nick Wilson for his work on emotion modeling, which we relied on in this work.

References

[1] S. G. Barsade, (2002). The ripple effect: Emotional contagion and its influence on group behavior. *Administrative Science Quarterly,* 47 (4), 644-675.

[2] A. Chmiel and J. A. Hołyst. (2010). Flow of emotional messages in artificial social networks. *International Journal of Modern Physics C,* 21 (5), 593-602.

[3] G. Deffuant, D. Neau, F. Amblard, and G. Weisbuch, (2000). Mixing beliefs among interacting agents. *Advances in Complex Systems*, 3(1-4), 87-98.

[4] J. J. Gross, (ed.) (2007). *Handbook of Emotion Regulation*. Guilford Press, New York.

[5] E. Hatfield, J. L. Cacioppo, & R. L. Rapson, (1993). Emotional contagion. *Current Directions in Psychological Sciences, 2,* 96-99.

[6] S. Helie and R. Sun, (2010). Incubation, insight, and creative problem solving: A unified theory and a connectionist model. *Psychological Review*, Vol.117, No.3, pp.994-1024.

[7] A. D. Kramer, J. E. Guillory, and J. T. Hancock. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences* 111(24), 8788-8790.

[8] R. S. Lazarus, (2001). Relational meaning and discrete emotions. In K. R. Scherer, A. Schorr, & T. Johnstone,

(eds.), *Appraisal Processes in Emotion.* Oxford, UK: Oxford University Press.

[9] A. Maslow, (1987). *Motivation and Personality*. 3rd Edition. Harper and Row, New York.

[10] H. Murray, (1938). *Explorations in Personality*. Oxford University Press, New York.

[11] A. Ortony, G. Clore, and A. Collins, (1988). *The Cognitive Structures of Emotions*. Cambridge University Press, New York.

[12] H. V. D. Parunak, S. H. Brooks, S. A. Brueckner, R Gupta, and L. Li, (2014). Dynamically Tracking the Real World in an Agent-Based Model. *Multi-Agent-Based Simulation XIV,* pp.3-16.

[13] R. L. Rapson, E. Hatfield, and J. T. Cacioppo. (1993). Emotional contagion. *Studies in emotion and social interaction*. Cambridge: Cambridge University Press.

[14] S. Reiss, S. (2004). Multifaceted nature of intrinsic motivation: The theory of 16 basic desires. *Review of General Psychology*, 8 (3), 179-193.

[15] D. Rumelhart, J. McClelland and the PDP Research Group, (1986). *Parallel Distributed Processing: Explorations in the Microstructures of Cognition*, MIT Press, Cambridge, MA.

[16] L. D. Smillie, A. D. Pickering, & C. J. Jackson, (2006). The new Reinforcement Sensitivity Theory: Implications for personality measurement. *Personality and Social Psychology Review*, 10, 320-335.

[17] R. Sun, (2002). *Duality of the Mind*. Lawrence Erlbaum Associates, Mahwah, NJ.

[18] R. Sun, (2009). Motivational representations within a computational cognitive architecture. *Cognitive Computation*, Vol.1, No.1, pp.91-103.

[19] R. Sun, (2016). *Anatomy of the Mind*. Oxford University Press, New York.

[20] R. Sun and R. Mathews, (2012). Implicit cognition, emotion, and meta-cognitive control. *Mind and Society*, Vol.11, No.1, pp.107-119.

[21] R. Sun, E. Merrill, and T. Peterson, (2001). From implicit skills to explicit knowledge: a bottom-up model of skill learning. *Cognitive Science*, Vol.25, No.2, 203-244.

[22] R. Sun, P. Slusarz, and C. Terry, (2005). The interaction of the explicit and the implicit in skill learning: A dual-process approach. *Psychological Review*, 112 (1), 159-192.

[23] R. Sun, N. Wilson, and M. Lynch, (2016). Emotion: A unified mechanistic interpretation from a cognitive architecture. *Cognitive Computation*, 8 (1), 1-14.

[24] C. Watkins, (1989). *Learning with Delayed Rewards*. Ph.D. Thesis, Cambridge University, Cambridge, UK.

[25] N. Wilson, (2012). *Towards a psychologically plausible comprehensive computational theory of emotion.* Ph.D Thesis, Cognitive Sciences Department, Rensselaer Polytechnic Institute, Troy, NY.

[26] N. Wilson and R. Sun, (2014). Coping with bullying: A computational emotion-theoretic account. *Proceedings of the Annual Conference of Cognitive Science Society*, Quebec City, Quebec, Canada. pp.3119-3124. Published by Cognitive Science Society, Austin, Texas.

[27] P. Winkielman, K. Berridge, & J. Wilbarger, (2005). Unconscious affective reactions to masked happy versus angry faces influence consumption behavior and judgments of value. *Personality and Social Psychology Bulletin*, *31* (1), 121-135.

[28] R. B. Zajonc, (1980). Feeling and thinking: Preferences need no inference. *American Psychologist*, *35* (2), 151.