

Modeling Social Emotions in Intelligent Agents Based on the Mental State Formalism

Alexei V. Samsonovich

Krasnow Institute for Advanced Study, George Mason University
4400 University Drive MS 2A1, Fairfax, Virginia 22030-4444
asamsono@gmu.edu

Abstract

Emotional intelligence is the key for acceptance of intelligent agents by humans as equal partners, e.g., in ad hoc teams. At the same time, its existing implementations in intelligent agents are mostly limited to basic affects. Currently, there is no consensus in the understanding of complex and social emotions at the level of functional and computational models. The approach of this work is based on the mental state formalism, originally developed as a part of the cognitive architecture GMU BICA. In the present work, complex social emotions like humor, jealousy, compassion, shame, pride, etc. are identified as emergent patterns of appraisals represented by schemas, that capture the cognitive nature of these emotions and enable their modeling. A general model of complex emotions and emotional relationships is constructed that can be validated by simulations of emotionally biased interactions and emergent relationships in small groups of agents. The framework will be useful in cognitive architectures for designing human-like-intelligent social agents possessing a sense of humor and other human-like emotionally intelligent capabilities.

Keywords: emotional intelligence; biologically inspired cognitive architectures; humor; higher-order cognition; affective computing; human-level AI.

Introduction

Computational models of agents with elements of emotional intelligence become popular nowadays (e.g., Laird, 2012). Yet, there is no general consensus on the approach to representation of complex or social emotions, feelings, appraisals, etc. in a cognitive architecture. The state of the art is represented by logic-based formalizations (e.g., Steunebrink, Dastani & Meyer, 2007) of phenomenological models like the OCC model (Ortony, Clore & Collins, 1988), which in effect defines a case-

based reasoning approach to cognitive appraisal of emotional values associated with a given situation. One may wonder whether there is a possibility of a more elegant theoretical model that may have explanatory and predictive powers.

An attempt to develop this sort of an approach was undertaken in the previous work (Samsonovich, 2012), where a general model was introduced as a parsimonious extension of the mental state framework of GMU BICA (Samsonovich, De Jong & Kitsantas, 2009). Here this model is elaborated further and applied to analysis and interpretation of selected examples.

The Extended Framework

The present study is based on the formalism of schemas and mental states (Samsonovich, De Jong & Kitsantas, 2009) developed as a basis for the cognitive architecture known as GMU BICA (Samsonovich & De Jong, 2005) or Constructor (Figure 1 C; Samsonovich, 2009). The essence of the mental state formalism is based on the observation that cognitive representations in human working memory are usually attributed to some mental perspective of a subject, while there are typically several mental perspectives simultaneously represented in human working memory. According to Samsonovich and Nadel, this attribution results in contents of awareness in working memory clustered into **mental states** that can be labeled in an intuitive self-explanatory manner: “I-Now”, “I-Previous”, “He-Now”, and so on, accordingly to the represented mental perspectives (Samsonovich & Nadel, 2005). The contents of mental states include instances of schemas that represent elements of awareness attributed to the mental perspectives of the represented subjects.

Emotional cognitive elements of awareness can be parsimoniously added to this base framework as three new categories: an **emotional state** (understood as an attribute

of a mental state based on appraisal of the current situation of which the agent is aware), an **appraisal** (understood as an attribute of an instance of a schema), and a higher-order appraisal schema, or **moral schema**, enabling appraisals of

specific patterns of appraisals and emotional states (Samsonovich, 2012). The new building blocks are shown in Figure 1 A, B.

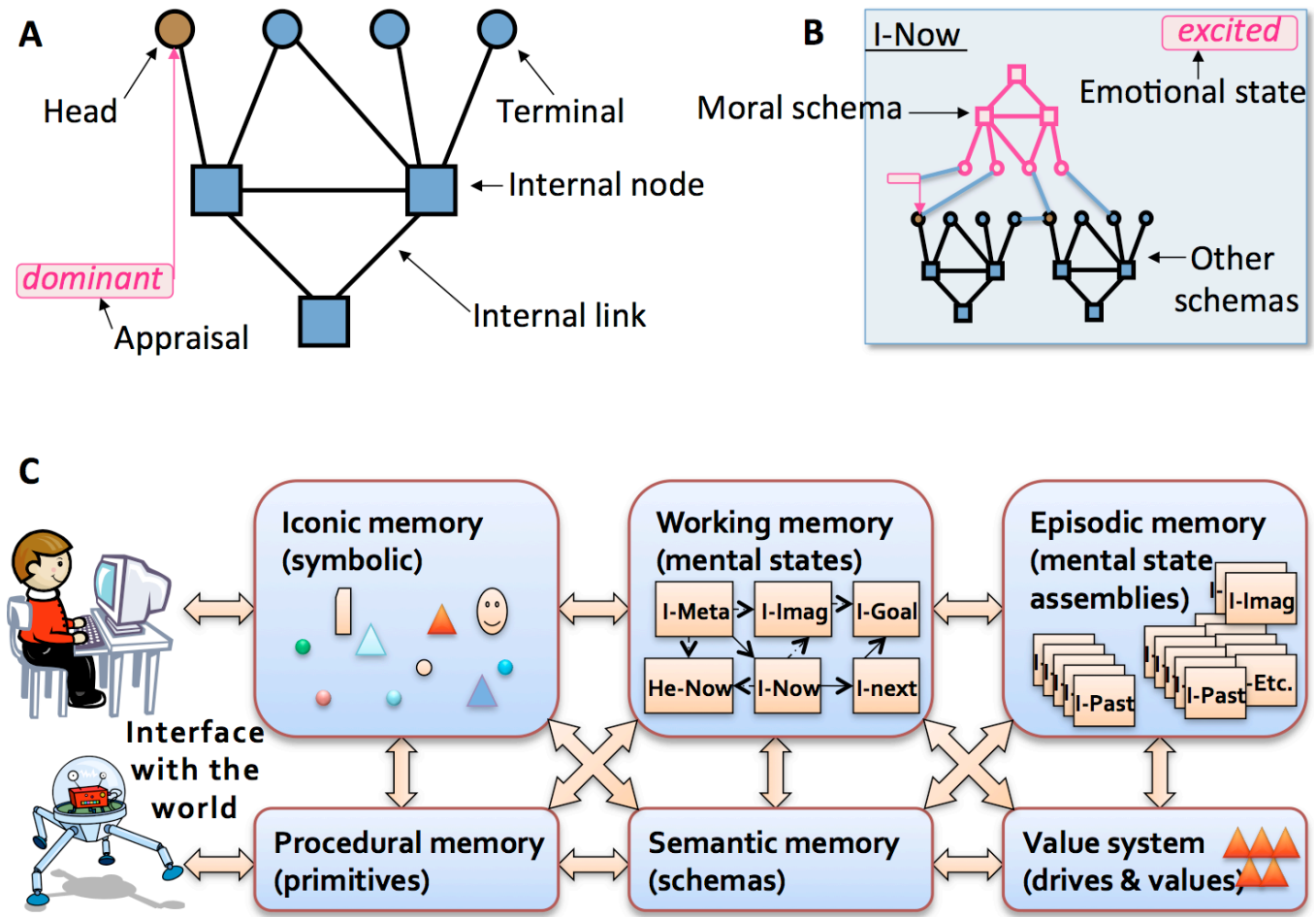


Figure 1. Extended framework of GMU BICA and Constructor. New elements are shown in pink in A and in B. **A:** A schema with appraisal added as a new attribute; **B:** A mental state, with emotional state added as a new attribute, populated by instances of schemas, including a new kind: moral schemas. Bindings are shown by light-blue lines. **C:** A bird's-eye view of the Constructor architecture (from Samsonovich, 2009). Here “iconic memory” is an augmented virtual reality simulator.

All representations in working, semantic and episodic memories of GMU BICA are at symbolic level, but not necessarily logic-based. There are two main building blocks of which these representations are constructed: a schema and a mental state (formally speaking, a mental state is also a schema, which is treated differently from other schemas). Both terms are highly overloaded in the literature. Here they refer to precisely defined computational objects.

A **schema**, as this term is understood here and in related works (Samsonovich & De Jong, 2005; Samsonovich, 2009) can be used to represent any concept or category, including entities, properties, events, relations, etc. of which the agent may be aware. All available schemas are stored in semantic memory. They can be invoked and instantiated in working memory to create new representations (including decisions and action commands). Instead of using logical inference, new representations in working memory are generated by

matching, binding and processing of schemas. Schemas are essentially different from predicates, as well as from production rules and operators used in Soar (Laird, 2012).

A schema can be represented as a graph (Figure 1 A) or, equivalently, as a table (not shown in Figure 1). There are three kinds of nodes of the graph: a head that represents the schema itself as a category; terminals (including the head) that can be bound to other schemas; and internal nodes that can be used to specify various conditions and effects of instantiation and processing of the schema. A minimal schema has only the head, while there is no upper limit on complexity of a schema. Each node of a schema has a standard set of domain-independent attributes that, in particular, determine how schemas can be bound to each other.

Again, here a **mental state** (Samsonovich et al., 2009) means a “box” that represents one particular mental perspective of some agent (Figure 1 B). The mental perspective (indicated by the mental state label, e.g., “I-Now”) may specify the context, the viewpoint, the identity and the status of the subject. The content of the “box” consists of bound instances of schemas that are interpreted as current elements of awareness of the agent whose mental perspective it is. The set of all logical possibilities of mental perspectives – e.g., “I-Now”, “I-Next”, “He-Past” – determines a mental state lattice. Mental states change their labels and functional roles with time, and eventually move from working memory to episodic memory.

The extension of the framework outlined above enables emotional cognition (Samsonovich, 2012), but does not affect the main principles of the architecture and its dynamics, that were described previously (Samsonovich & De Jong, 2005; Samsonovich et al., 2006). This framework will be applied below to specific examples of social emotions.

Understanding Social Emotions

The main part of the present study is focused on several specific examples of social emotions and emotional relationships, including shame, pride, trust, guilt, jealousy, humor, compassion. As pointed by Lazarus (2001), social emotions are better understood when grouped into clusters. This statement becomes even more meaningful when a cluster of emotions can be described by one moral schema. The following three examples of clusters of social emotions will be considered to illustrate this point.

Shame and Pride

According to the global structure of emotion types in OCC, shame and pride occur as valenced reactions to actions attributed to self, perceived as an agent (Clore et al., 1988,

p. 19). On the other hand, Tangney et al. (2001) found shame-proneness to be correlated with suspiciousness, resentment, and tendency to blame others. If shame and pride do not need to involve representations of other minds, then, how their social correlates can be understood?

The answer can be given based on the mental state formalism, if we agree to understand social emotions as emotions involving multiple mental states. For comparison, basic emotions like joy or fear do not need to rely on other mental states, when their source is present in the current situation and does not depend on other agents. In contrast, shame and pride arguably involve appraisal of self from another mental perspective. As we shall see, this condition can be described in terms of two interacting mental states of ‘I’: e.g., one at present, and one in the past.

Trust, Guilt and Jealousy

As pointed by Baumeister, Stillwell and Heatherton (2001), in contrast with shame, guilt occurs primarily in the context of ongoing relationships, e.g., involving trust and responsibility (e.g., love or close friendship), and may be triggered by a failure to meet another’s expectations. Furthermore, Baumeister et al. (2001) show that guilt has the ability to heal broken relationships.

In analogy with the emotions of shame and guilt, jealousy is differentiated from envy by Parrott (2001):

In envy, one’s own appraisal leads to dissatisfaction with oneself. In jealousy, the reflected appraisal of another leads to a lack of security and confidence.

These observations are reflected in model interpretations of specific examples considered below: e.g., assumptions about the moral schemas of shame and trust.

Humor, Compassion, Resentment...

The nature of humor is one of the topics that still evades scientific analysis. Of all social emotions, the sense of humor is probably most poorly understood. This could be partially due to a large number of humor subtypes. While this feature is typical for many social emotions (e.g., Parrott, 2001, points to a number of subtypes of both envy and jealousy), the sense of humor is probably unique in this regard. Not surprisingly, there is no precise scientific definition of the sense of humor: the term is used as an umbrella for a broad spectrum of phenomena (Lefcourt, 2001; Hurley, Dennett & Adams, 2011).

While the sense of humor is frequently taken for granted as an exclusive human characteristic, little is known about humor-like emotional states in other animals. Do animals other than humans have a sense of humor? Do they laugh? These are two very different questions. E.g., human laughter can be triggered by tickling rather than by an emotional experience. It was recently found that rats

respond to tickling with frequency-modulated 50 kHz vocalization (Panksepp, 2007, Burgdorf et al., 2011). Furthermore, these studies suggest that this kind of vocalization in rats signifies a positive affective state and has functional similarities with human laughter. Yet, there is no documented evidence suggesting that rats, or any nonhuman animals in general, may have a sense of humor.

Among cases when laughter is triggered by emotions, the sensation of humor caused by a joke may be an exception. Bering (2012) points that laughter is typical for many other emotional contexts, e.g., joy, affection, amusement, cheerfulness, surprise, nervousness, sadness, fear, shame, aggression, triumph, taunt, schadenfreude (pleasure in another's misfortune). E.g., types of laughter studied by Szameitat et al. (2011) include: joy, tickling, taunting (gloating), schadenfreude. On the other hand, sensation of humor triggered by a joke may not be expressed as laughter.

The picture from the theoretical perspective does not look better. A large number of theories of humor were developed over the centuries. Many of them correctly capture some aspect or type of humor, while missing others.

The present study focuses on the cognitive aspect of humor, as well as other considered here social emotions. More specifically, the focus is on the type of humor that is called “higher-order humor” by Hurley et al. (2011). This is the type of humor that involves the Theory-of-Mind¹: i.e., the human ability to simulate and understand other minds (Nichols & Stich, 2003). The reason is obvious: a simulationist computational model of a Theory-of-Mind lies at the foundation of the mental state framework used here.

It would be nice if the sense of humor had a concise definition in terms of Theory of Mind. For example, Hurley et al. (2011) have a proposal of this sort (p. 145):

We may find things funny either if they are invalidated mental spaces in our own knowledge representations or if we recognize that they are invalidated mental spaces for another entity's knowledge representation.¹

In other words, according to Hurley et al. (2011), mental perspectives that are either internally inconsistent or based on false beliefs in our belief system appear funny to us.

Obviously, this formula does not work as a generally acceptable definition: e.g., if your friend is in danger of which (s)he is not aware, then you may be more likely to experience anxiety or compassion than humor. Similarly, a person who is an object of a practical joke may experience resentment. But these observations only suggest a

¹ Hurley, Dennett and Adams (2011) use the term “intentional stance” instead of the term “Theory-of-Mind”, and the term “mental space” instead of the term “mental perspective”.

clustering of emotions, that may be possible to describe with one and the same moral schema.

In the next section, specific examples of emotional interactions among agents are analyzed and interpreted in terms of mental states, appraisals, and moral schemas in a framework that is suitable for modeling.

Interpretation of Selected Examples

The following consideration is based on three moral schemas corresponding to the three clusters of social emotions considered above (Figure 2):

- the schema of shame and pride;
- the schema of trust, guilt and jealousy;
- the schema of humor.

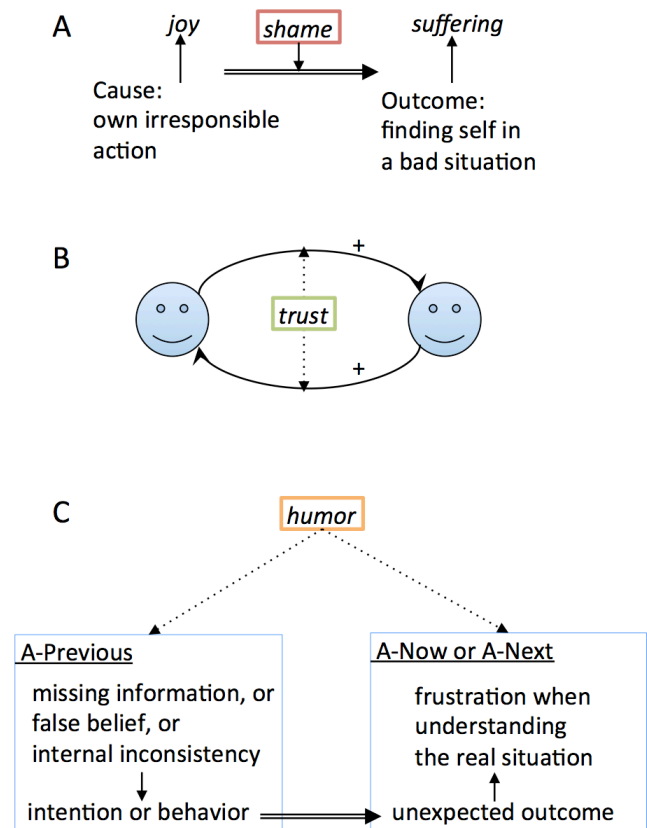


Figure 2. Examples of moral schemas corresponding to the three considered clusters of social emotions. **A:** A specific schema of shame and pride (cf. Figure 3). **B:** A schema of trust, guilt and jealousy (cf. Figure 4). **C:** One of possibly many schemas of humor (cf. Figure 5), that may be related to compassion and other emotions.

Each of these schemas binds to certain elements of two interacting mental states. The schemas themselves define specific conditions for the binding. The appraisal develops inside one of the two mental states, except for humor (in this interpretation), in which case the appraisal of the schema develops in a separate (third) mental state. These examples of schemas do not account for all cases of the corresponding emotions, and represent limited cases.

Shame and Pride

Consider an abstract virtual world paradigm, in which the agent's life cycle includes two alternating phases: A and B (Figure 3). The agent is in good shape as long as it keeps consuming an abstract resource, which is replenished every time at the beginning of Phase A. Consumption at a high rate gives pleasure to the agent, no consumption causes pain. High-rate consumption during phase A results in the depletion of resource in phase B.

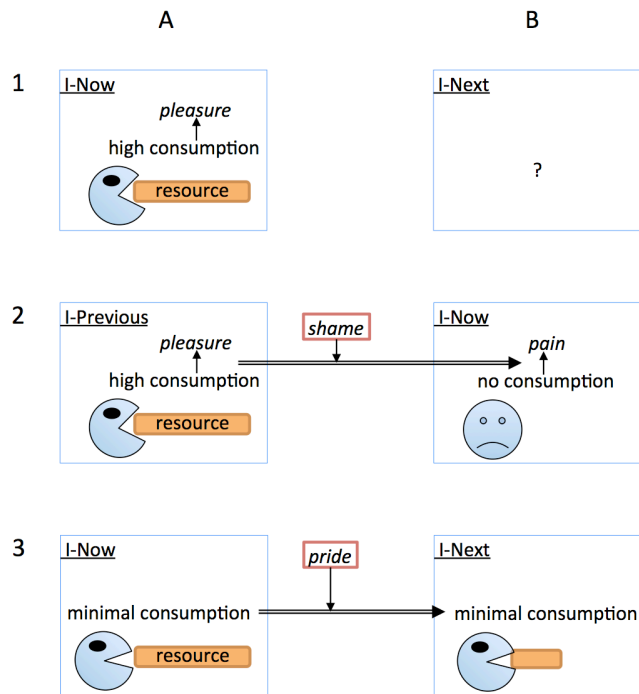


Figure 3. Example illustrating the shame-pride schema in action. The mental state in 1B is not simulated by the architecture. The schema of shame is invoked when conditions match its definition.

The agent is controlled by a cognitive architecture based on the formalism of mental states and schemas. Two of the mental states in working memory are shown in each row in Figure 3. Mental states are shown as boxes; mental state labels (e.g. “I-Now”) are underlined. The agent’s action selection is driven by appraisal values that work as goals and rewards for reasoning and learning mechanisms.

Initially being naïve (Figure 3-1A), the agent starts consuming the resource at a high rate in order to maximize immediate pleasure, without simulating the next situation. At the next phase (Figure 3-2B), the agent suffers from pain, because there is no available resource left. Performing metacognitive analysis of its recent memories, the agent attributes the cause of the unwanted outcome – pain in phase B – to the pleasure-driven high consumption of resource in phase A. This pattern (irresponsible pleasure-driven action causing a bad outcome) invokes the moral schema of shame and pride, thereby instantiating a feeling of shame, that produces effect on action selection.

Next, entering the phase A again, the agent retrieves the relevant recent episodic memory of a bad experience (Figure 3 row 2 stored as episodic memory), which prompts the agent to simulate the next situation (Figure 3 row 3). Now, by minimizing the negative feelings in all mental states, including pain and shame, the agent develops a strategy that reverts the sign of the appraised shame, resulting in a feeling of pride. As a result, the agent avoids suffering from pain.

In the outlined above scenario, deleting or disabling the shame-pride schema would make the agent unable to find an optimal strategy, unless an alternative approach is used: e.g., based on logical reasoning and planning, given a specified goal. The strength of the emotional approach is that it does not require an a priori defined goal and in general can work in unexpected situations.

Trust, Guilt and Jealousy

Consider a group of three agents, A, B and C, that interact with each other in some unspecified environment. Suppose A and B develop mutually positive relationships that allow them to trust each other. The pattern of mutually positive relationships (Figure 4, row 1) invokes the schema of trust, two instances of which become bound to the two relationships in mental states of A and B. Like in Figure 3, each set of mental states (e.g., Figure 4, row 1) describes the state of the mind in one agent. In this case, all 3 agents are assumed to have mutually consistent states of mind. The instantiated schema of trust stabilizes relationships.

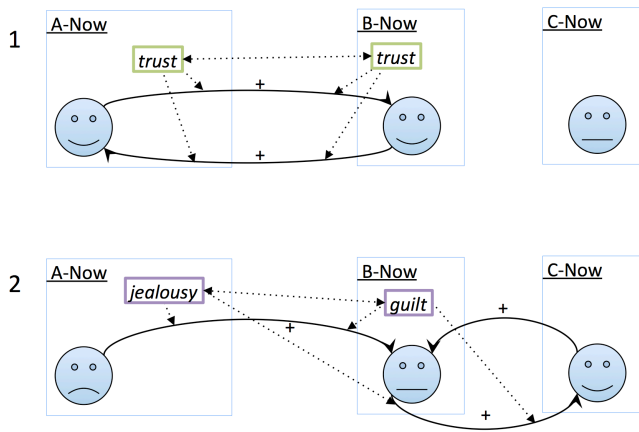


Figure 4. Example illustrating the moral schema of trust, guilt and jealousy. The initially formed instances of trust in A-Now and in B-Now turn into jealousy and guilt, respectively.

It is assumed further that the nature of this sort of positive relationship (solid lines in Figure 4 marked by a “+”) is such that it allows for only one recipient: e.g., B cannot have the same relationships of this kind simultaneously with A and with C. If nevertheless for some reason B should flip its appraisal of A to C, then the bound instance of the moral schema that represented trust will induce a feeling of guilt in the mental state B-Now and a feeling of jealousy in a mental state A-Now. This new configuration (Figure 4 row 2) has a potential to restore the original configuration (Figure 4 row 1) through interactions of the two instances of the moral schema.

It is remarkable that in this framework the inference about an emergent feeling of guilt or jealousy can be made immediately based on the same instance of the schema of trust, without the need to analyze the entire situation anew, which would be necessary in a traditional approach, e.g., based on OCC (Ortony et al. 1988).

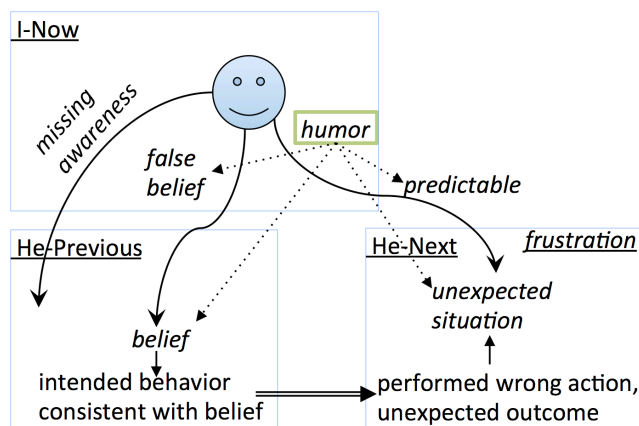


Figure 5. An abstract pattern of mental states and appraisals in which the schema of humor (Figure 2 C) can be bound.

A Model of a Sense of Humor

Consider now an agent labeled ‘I’ who is aware of the mind of another, possibly abstract agent labeled ‘He’ (or ‘She’) at two moments of time (Figure 5). It is assumed that conditions of the two mental states of ‘He’ match the schema of humor (Figure 2 C). At the first moment (Figure 5, He-Previous), the agent He is concerned with some belief, and acts based on it. However, when appraised from the ‘I-Now’ perspective, the belief appears to be false, and at the same time awareness of another important detail of the current situation appears to be missing.

When He behaves consistently with his beliefs and awareness, the outcome turns unexpectedly bad or frustrating for him, and completely changes his state of mind in the state ‘He-Next’, when He becomes aware of the real situation. The entire episode is finally clear to ‘I’.

This pattern of appraisals and mental state changes corresponds to many (but obviously not all) conditions under which a sensation of humor occurs. Therefore, it can be captured by a moral schema associated with the sense of humor (Figure 2 C). The same schema could also produce a feeling of compassion or resentment, depending on the role of ‘He’ with respect to ‘I’.

Analysis of Joke Examples

In the following examples, the humor schema represented in Figure 2 C is used for interpretation of specific jokes. Comparison is made to interpretation of the same jokes given by Hurley et al. (2011).

Of course, there are many jokes that would fail to match the schema. They may correspond to a different type of humor.

Example 1: The World’s Funniest Joke

(<http://en.wikipedia.org/wiki/LaughLab>; also in Hurley et al., 2011, p.164)

Two hunters are out in the woods when one of them collapses. He doesn't seem to be breathing and his eyes are glazed. The other guy whips out his phone and calls the emergency services. He gasps, "My friend is dead! What can I do?" The operator says "Calm down. I can help. First, let's make sure he's dead." There is a silence, then a gun shot is heard. Back on the phone, the guy says "OK, now what?"

Interpretation based on Figures 2 C and 5 goes like this. ‘He’ is the awake hunter, who is concerned with a possible claim of his responsibility for the death of his friend (a possibly false belief) more than he is concerned with saving the life of his friend (missing awareness of the necessity to help, if the friend is alive). He therefore commits to following exactly every word of authorities without thinking, in order to exclude any possibility of wrongdoing on his part. Having misunderstood the advice,

he acts contrary to the common sense (internal inconsistency). The outcome turns out to be precisely what he was trying to avoid (unexpected situation) He-Next is implied by the joke.

For comparison, the interpretation offered by Hurley et al. (2011) can be in its essence summarized as follows (it also maps onto the same schema). The ambiguity of the operator's advice leads the hunter to mistake the meaning of the advice in a dumb way. He believes that he follows the advice of somebody with experience and authority who wants to help, and commits to a stupid action without noticing its stupidity.

Example 2

(<http://en.wikipedia.org/wiki/LaughLab>)

Sherlock Holmes and Dr. Watson were going camping. They pitched their tent under the stars and went to sleep. Sometime in the middle of the night Holmes woke Watson up and said: "Watson, look up at the stars, and tell me what you see." Watson replied: "I see millions and millions of stars." Holmes said: "And what do you deduce from that?" Watson replied: "Well, if there are millions of stars, and if even a few of those have planets, it's quite likely there are some planets like Earth out there. And if there are a few planets like Earth out there, there might also be life." And Holmes said: "Watson, you idiot, it means that somebody stole our tent."

Interpretation based on Figure 2 C goes like this. 'He' is Dr. Watson, who is concerned with answering the question of his friend as completely as possible, and misses the obvious motivation of the Holmes' question. Unexpectedly, his answer turns stupid and irrelevant.

Example 3

(Hurley et al., 2011, p. 160).

"Hey, did you know you have bananas in your ears?"

"Speak up! I have bananas in my ears!"

A possible interpretation based on Figure 2 C goes like this. The two agents (call them A and B) are both intended to help each other by trying to alert each other to a fact of which they both are aware (bananas in B's ears). However, each of them initially believes that the other one is not aware of this fact. Moreover, A does not realize that he cannot be heard by B. Their actions are motivated by these false beliefs and missing awareness. The possibility of the next mental state of confusion is merely implied.

If both A and B are pretending being stupid, then the interpretation would involve the played characters instead of the real A and B.

Hurley et al. (2011) do not offer an elegant interpretation of this joke in terms of their higher-order humor model.

Example 4

(Hurley et al., 2011, p.164)

A senior citizen is driving on the highway. His wife calls him on his cell phone and in a worried voice says, "Herman, be careful! I just heard on the radio that there was a madman driving the wrong way on Route 280!" Herman says, "Not just one, there are hundreds!"

It is easy to identify in this example all key elements: a false belief of Herman, his wrong behavior based on this belief, and an implied, unexpected for Herman unpleasant outcome. As Hurley et al. point, the role of his wife in this case is negligible, and indeed is missing in another variant of this joke.

Hurley et al. (2011) offer a different interpretation of this joke: in terms of a covert insertion. Still, they conclude that the lion's share of the humor in this joke comes from the listener's recognition that the fellow is obtuse.

Conclusion

A variety of intelligible jokes can be robustly mapped onto one and the same schema depicted in Figure 2 C. Cases that match the same schema and are not funny may be possible to exclude with additional filtering conditions.

Discussion

While this study focused on only three clusters of social emotions, one can imagine its extension to other emotions, representations of which involve multiple mental states. A collective term "**higher-order emotions**" seems to be an appropriate name for this specific category.

It is possible that the sense of humor is an exclusive human ability. If so, then understanding it computationally could be a key to human-level artificial intelligence.

More generally, computational modeling of higher-order emotions is essential in our advance toward human-level general artificial intelligence, for two reasons. One is the necessity to understand the nature of human intelligence. Another is to make sure that future intelligent machines will be human-compatible and acceptable by the human society as intelligent minds rather than intelligent tools.

A bigger problem is that, despite the impressive progress in modern artificial intelligence, most state-of-the-art intelligent agents and artifacts lack general commonsense initiative and remain task-oriented. Modern artificial agents can communicate with people using natural language and gestures, yet we cannot imagine them developing long-lasting personal relationships with people, or making their way up in the human social hierarchy on their own. But this picture may change in the near future. Artificial emotional intelligence emerges as a key for acceptance of agents by humans as "equal minds". At the

same time, it still has a long way to go. The distance can be compared to the distance between the discovery of a nuclear chain reaction and the development of a nuclear reactor. Indeed, current emotional cognitive capabilities of intelligent agents are mostly limited to basic affects. Even at the level of a theory, there is no consensus in the understanding of even the most basic social emotions. Nevertheless, the problem appears to be solvable in the near future, as the present study indicates.

Conclusions

In the present work, complex social emotions like humor, jealousy, compassion, shame, pride, etc. were identified as emergent patterns of appraisals represented by schemas, that capture the cognitive nature of these emotions and enable their modeling. A general model of social emotions and emotional relationships was constructed that can be validated by simulations of emotional interactions and emergent relationships in small groups of agents.

The framework will be useful in cognitive architectures for designing human-like-intelligent social agents possessing a sense of humor and other human-like emotionally intelligent capabilities. Appraisals of actions necessary for this purpose can be computed automatically using available data (Heise, 2007).

References

- Baumeister, R. F., Stillwell, A. M., & Heatherton, T. F. 2001. Interpersonal aspects of guilt: Evidence from narrative studies. In: Parrott, W. G. (Ed.). *Emotions in Social Psychology: Essential Reading*, pp. 295-305. Ann Arbor, MI: Taylor & Francis.
- Bering, J. 2012. The rat that laughed: Do animals other than humans have a sense of humor? Maybe so. *Scientific American* 307 (1): 74-77.
- Burgdorf, J., Panksepp, J., and Moskal, J. R. 2011. Frequency-modulated 50 kHz ultrasonic vocalizations: a tool for uncovering the molecular substrates of positive affect. *Neuroscience and Behavioral Reviews* 35 (9): 1831-1836.
- Heise, D. R. 2007. *Expressive Order: Confirming Sentiments in Social Actions*. New York: Springer.
- Hurley, M. M., Dennett, D. C., & Adams, R. B. 2011. *Inside Jokes: Using Humor to Reverse-Engineer the Mind*. Cambridge, MA: The MIT Press.
- Laird, J. E. (2012). *The Soar Cognitive Architecture*. Cambridge, MA: MIT Press.
- Lazarus, R. S. 2001. Relational meaning and discrete emotions. In: Scherer, K.R., Schorr, A., & Johnstone, T. (Eds.). *Appraisal Processes in Emotion: Theory, Methods, Research*, pp. 37-67. Oxford, UK: Oxford University Press.
- Lefcourt, H. M. 2001. *Humor: The Psychology of Living Buoyantly*. New York: Kluwer Academic.
- Nichols, S. and Stich, S. 2003. *Mindreading: An Intergrated Account of Pretence, Self-Awareness, and Understanding Other Minds*. Oxford: Oxford University Press.
- Ortony, A., Clore, G. L., and Collins, A. 1988. *The Cognitive Structure of Emotions*.
- Panksepp, J. 2007. Neuroevolutionary sources of laughter and social joy: Modeling primal human laughter in laboratory rats. *Behavioral Brain Research* 182 (2): 231-244.
- Parrott, W. G. 2001. The emotional experiences of envy and jealousy. In: Parrott, W. G. (Ed.). *Emotions in Social Psychology: Essential Reading*, pp. 306-320. Ann Arbor, MI: Taylor & Francis.
- Samsonovich, A. V. 2009. The Constructor metacognitive architecture. In Samsonovich, A. V. (Ed.). *Biologically Inspired Cognitive Architectures II: Papers from the AAAI Fall Symposium. AAAI Technical Report FS-09-01*, pp. 124-134. Menlo Park, CA: AAAI Press.
- Samsonovich, A. V. 2012. An approach to building emotional intelligence in artifacts. In Burgard, W., Konolige, K., Pagnucco, M., and Vassos, S. (Eds.). *Cognitive Robotics: AAAI Technical Report WS-12-06*, pp. 109-116. Menlo Park, CA: The AAAI Press.
- Samsonovich, A. V. and De Jong, K. A. 2005. Designing a self-aware neuromorphic hybrid. In K.R. Thorisson, H. Vilhjalmsson, and S. Marsela (Eds.). *AAAI-05 Workshop on Modular Construction of Human-Like Intelligence: AAAI Technical Report*, volume WS-05-08, pp. 71-78. Menlo Park, CA: AAAI Press.
- Samsonovich, A. V. and Nadel, L. 2005. Fundamental principles and mechanisms of the conscious self. *Cortex* 41 (5): 669-689.
- Samsonovich, A. V., Ascoli, G. A., De Jong, K. A., and Coletti, M. A. 2006. Integrated hybrid cognitive architecture for a virtual roboscout. In M. Beetz, K. Rajan, M. Thielscher, and R.B. Rusu (Eds.). *Cognitive Robotics: Papers from the AAAI Workshop, AAAI Technical Reports WS-06-03*, pp. 129-134. Menlo Park, CA: AAAI Press.
- Samsonovich, A. V., De Jong, K. A., and Kitsantas, A. 2009. The mental state formalism of GMU-BICA. *International Journal of Machine Consciousness* 1 (1): 111-130.
- Steunebrink, B. R., Dastani, M., and Meyer, J. - C. 2007. A logic of emotions for intelligent agents. *Proceedings of the 22nd Conference on Artificial Intelligence (AAAI 2007)*, pp. 142-147. Menlo Park, CA: AAAI Press.
- Szameitat, D.P., Alter, K., Szameitat, A. J., Darwin, C.J., Wildgruber, D., Dietrich, S., and Sterr, A. 2009. Differentiation of emotions in laughter at the behavioral level. *Emotion* 9 (3): 397-405.
- Szameitat, D.P., Darwin, C. J., Wildgruber, D., Alter, K., and Szameitat, A.J. 2011. Acoustic correlates of emotional dimensions in laughter: Arousal, dominance, and valence. *Cognition and Emotion* 25 (4): 599-611.
- Tangney, J. P., Wagner, P., Fletcher, C., & Gramzow, R. 2001. Shamed into anger? The relation of shame and guilt to anger and self-reported aggression. In: Parrott, W. G. (Ed.). *Emotions in Social Psychology: Essential Reading*, pp. 285-294. Ann Arbor, MI: Taylor & Francis.