Computational Approaches to Emotional Modeling

Sinem Şemsioğlu
KUAR and KARMA Lab
Koç University
ssemsioglu16@ku.edu.tr

Fuat Balcı
Biological Sciences & Computer Science
University of Manitoba
fuat.balci@umanitoba.ca

Asım Evren Yantaç
KUAR and KARMA Lab
Koç University
eyantac@ku.edu.tr

***Abstract*—Emotion is a concept that is very familiar to all of us but has been proven to be hard to define, not only to other people but also through artificial systems. There are many approaches to modeling emotions, some of them theoretical/qualitative and some of them computational. It is hard to integrate qualitative models into empirical studies, as the concepts they define are hard to measure. On the other hand, computational models are usually developed for practical purposes and do not comprehensively describe emotions; which makes them suitable in applied fields but not when probing the nature of emotions. Based on earlier work, we argue that it is possible to develop an emotion model that is both quantitative and comprehensive. Here, we talk about existing connections between emotions and mathematical theories: information theory, game theory, set theory, dynamical systems and hope to inspire future research on mathematics based comprehensive emotion models.**

***Index Terms*—emotion modeling, affective computing, mathematical theories, computational emotion modeling, affective computing, emotion theories, theory formalization**

I. Introduction

Although emotions are ubiquitous components of our daily experience, there is not a single clear definition of what emotions are. There are many approaches to characterizing emotions in philosophy and affective sciences both to understand and to be able to work with them computationally in artificial systems; the multitude of approaches causes a lack of standardization [1].

Apart from the ongoing discourse on defining and modeling emotions, work in some of the fields such as HCI and AI usually stick to more traditional models, such as Ekman's 6 basic emotions [2] and Russell's circumplex model of affect [3]. The recent developments in affective sciences and philosophy are not reflected in these areas. As work in these fields have their own focus, it is understandable that researchers continue working with models that have been accepted in their fields; however, (a) this sometimes translates into missed opportunities for validating newly proposed models, (b) limits the depth of discussion on emotions and the breadth of the implication of the presented insights. Thus, unifying emotional approaches and knowledge sharing between disciplines can be identified as two main challenges that remain to be resolved. In this paper we review the use of mathematical theories and tools in emotion modeling, and speculate on a way various approaches can be combined to reach a more unified state.

II. Emotion modeling

A. Challenges

Stanford Encyclopedia of Philosophy provides an interdisciplinary and comprehensive overview of approaches to defining emotions, covering views among history - from Greek Philosophers to current scholars (including published work up to 2018) [4]. In this entry, Scarantino and de Sousa categorize traditional approaches into three, treating emotions as: feelings, evaluations or motivations; and elaborate on how they have evolved into current theories [4]. Looking at theories commonly referred in HCI, aforementioned Basic Emotions [2] and Russell's circumplex model [3] fall into the motivational category, as they define/classify emotions based on the (linked/resulting) physiological changes and behavior; whereas, appraisal theories like Scherer's [5] are categorized as evaluative since they focus on how stimuli are processed to cause emotional experiences [4]. Example of a theory that joins the feeling based approach with the evaluative one is proposed by Goldie, who treats emotions as intrinsically intentional, thinking of feelings as directed [6].

Scarantino and de Sousa identify four core challenges that concern all of the approaches [4]:

- Differentiation: How are emotions different from one another, and from things that are not emotions?
- *Motivation: Do emotions motivate behavior, and if so how?*
- *Intentionality: Do emotions have object-directedness, and if so can they be appropriate or inappropriate to their objects?*
- *Phenomenology: Do emotions always involve subjective experiences, and if so of what kind?*

In addition to the challenges on defining emotions, another set of challenges can be defined for representing emotions within interactive systems. Bucci et al. discuss shortcomings of existing models in interactive affective scenarios and propose four properties for affective representations to be used in such interactive applications: (1) allowing multiple emotions to be represented, (2) modeling uncertainty estimates, (3) time-variance, (4) non-linearity [7]. Among these, time-variance and non-linearity have theoretical roots. Variance of emotions over time is part of our common experience of emotions, yet is not included in many models; as Bucci et al. state, models should be able to account for real-life transitions we experience (ie. going from a very unpleasant state to a pleasant one) [7]. Moreover, the experience of and the effort required for different emotional transitions can be compared and questioned in terms of linearity. Taking Russell's circumplex model affect [3] as an example, we can ask whether the transition from one point to any equidistant point yields similar experiences, and requires the same effort? In a broader sense, while developing emotion models we think it is important to consider *how different emotional states relate to each other, and whether one can transform to another?*. While working on a quantitative model these questions can become problems of defining distance and transformation functions within the affective space.

B. modeling Strategies

In a 2013 paper Reisenzein et al. call out for inter and intradisciplinary coordination in developing computational emotion models, and outline two main strategies: (a) modularization, (b) unification and standardization [8]. The former implies a more applied and practical approach while the latter implies a more theoretical and formal approach. In his 2019 paper we find Reisenzein reflecting on the past 30 years of research on cognition and emotions, stating that the field is still fragmented and that this state might even be causing loss of knowledge and hindering progress; he calls again for integration of theories [9]. In a similar vein, Hudlicka points out the need for a more systematic approach in emotion modeling with consistent terminology, guidelines and validation methods [1]. While reviewing existing efforts in modeling, she makes a distinction between research and applied models of emotion [1]; these roughly correspond to the two strategies proposed in [8]. Among those, we find the development of research models, geared towards a formal, theoretical and comprehensive definition of emotions more interesting. As Reisenzein et al. suggest, such models need not to be constrained to a conceptual state but can be formulated as a computational model, or can provide a basis for computational implementations and have quantitative extensions [8].

III. Mathematical Connections

Here we review some of the existing work on tying mathematical theories to characterization of emotional experiences.

A. Formalization and Set Theory

Set theory studies collections of objects, which are sets, and the operations that can be performed on them. It produces a well-determined and abstracted base to represent concepts and allows their formalization, which can be thought of as abstractions with very strict rules. All mathematical objects can theoretically be represented as sets [10] and studied within set theory.

Reisenzein suggests that formalizing existing emotion theories can be a step in unifying existing work [8]; he expects this process to also contribute to the refinement and clarification of these theories. One important aspect of formalizations is that they do not need to be strongly connected to any specific

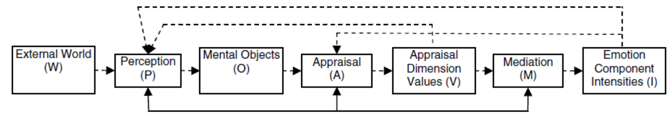

Fig. 1. The underlying appraisal structure used in Broekens et al.'s formalization [11]

theory but can act as a framework to explain multiple theories. For example Broekens et al. characterize our emotional processes in terms of sets of 1) events and objects in the external world, 2) our mental representations of these events and objects, 3) appraisal-dimensions that we use to evaluate our mental representations, 4) emotion-components [11]. These sets, especially appraisal dimensions and emotion components can be defined based on the implemented theory. Figure 1 contains a visual representation, and in Section 4 we build upon their theory. Reisenzein suggests that another less strict formalization approach can be to represent emotion theories in formal programming languages (logic languages); stating such formalizations have been used during implementation of emotions in autonomous agents [8].

B. Dynamical Systems

A dynamical system approach to emotions can account for the suggested temporality aspect. Dynamic systems examine the temporal changes of a system's state, defining potential transitions between different states. Dynamic systems can be linear or non-linear (chaos theory), deterministic or probabilistic. A common example of a dynamic system is the modeling of the swinging of a clock pendulum. Colombetti finds dynamical systems theory to be the best candidate to conceptualize emotional episodes [12]. Some important aspects of this view are: (1) thinking of emotional states as subject to change due to external as well as internal stimuli/parameters and being able to specify these, (2) thinking of some emotional states as attractors and some as repellors, (3) explaining why some states are easier or harder to reach and some are more stable or unstable, (4) thinking about emotional interaction scenarios as coupled dynamical systems, (5) the aspect of self-organization, suggesting that different affective components might be influencing each other to sustain order [12].

C. Signal and Information Theories

Both Broeken et al.'s set theoretical approach [11] and Colombetti's suggestion of using dynamical systems to model emotion episodes [12] recognize external factors as influential on the emotional state. But how should this effect be described? In his theory of feelings-as-emotion, Schwarz interprets feelings as an information source that informs us about the environment and the situation, which influences how we process external stimuli (ie. attention to details) and how we form judgments [13]. We do not only get informed by our own feelings but also by others'. In Emotions as Social Information (EASI) model, van Kleef et al. suggest that emotional expressions can be considered as social signals that carry information about both the person who broadcasts these signals and the environment around them [14].

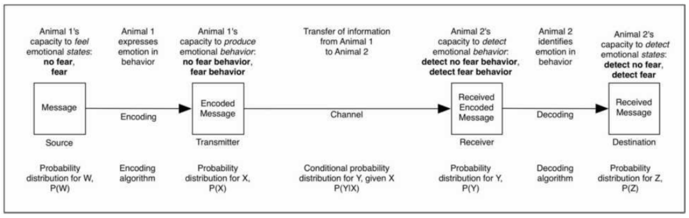

Fig. 2. Schema showing information theoretical approach to a dyadic emotion communication [15]

Kerr and Scharp build an information theoretical approach based on this view of emotions as social signals; suggesting that principles of information theory apply to emotion communication, considering emotion expressions as encoded emotion messages [15]. Information theory focuses on communication of messages through channels, originally proposed with telecommunication systems in mind [16]. It deals with concepts such as encoding and decoding of messages, transmission rate, error in transmission and error correction strategies. The simplest case is considering the case of a dyadic communication with the presence of one sender and one receiver (Figure 2). Communication networks including multiple senders and/or receivers are also studied. Both of these cases are applicable to communication of emotions.

D. Theory of Mind

A related theoretical approach is Theory of Mind; it suggests that while reasoning about our observations of others' (ie their expressions, actions), we make inferences about their mental states (ie goals, beliefs) based on how our own theory of how mind works. Making inferences about their intentions or possible future actions allows us to adjust our behavior accordingly [17]. Bayesian Theory of Mind suggests that Bayesian methods can be used to model our inference of agents' beliefs and desires based on observations [17].

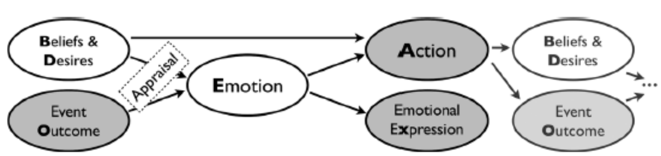

Fig. 3. A schema describing a sample reasoning, third-person-appraisal as a diagram [18]

Similarly, in the case of emotions, it is suggested that we interpret our observations of others' emotional expressions based on our intuitive theories of emotions, and accordingly predict how someone feels [18]. Ong et al term the emotional reasoning processes as affective cognition [19] and provide a taxonomy of affective cognition inferences, suggesting that a unified intuitive theory of emotion can explain these reasoning via Bayesian inference [18]. Their taxonomy includes processes of emotion recognition, third person appraisals (Figure 3), inferring cause of emotions, emotional cue integration, reverse appraisal, hypothetical reasoning and counterfactual reasoning, expressing them in terms of conditional probabilities [18].

E. Game Theory

When considering emotional expressions as signals that carry information, we also need to consider the nature of expression. Expression level might change based on many factors: personal habits, context, who the observers are and such. Although we might not be able to control all emotional signals (ie physiological ones), we might strategically alter how expressive we want to be in terms of altering our facial expressions, body language, verbal language and such. Thus, being able to model and predict the expressivity level might come in handy while decoding others' emotion signals. Such strategic choices based on expression might be explained by game theory.

Lewis's theory of signal systems can be taken as a starting point to connect emotion signaling to game theory [15]. Lewis signaling games [20] are those with two players, a sender and a receiver, who communicate with each other. At each turn, nature (can be considered as non-player agent) picks one state at random; sender observes this state and sends a signal to communicate it to the receiver; receiver, who doesn't have access to the state chooses an act based on the signal. There are equal number of states, signals and acts and one correct mapping of each state to an act. If the receiver chooses the correct act (ie makes a correct prediction of the nature's state based on sender's signal), both of the players get rewarded. Therefore, the game encourages players to come up with a language that allows communication of the state information. If they evolve a strategy that makes them more successful than chance, their language can be considered efficient; in the case they are able to form a one to one mapping of states to signals and signals to acts, they have formed a perfect signaling system which allows them to have a hundred percent success rate (assuming they always pick signals and acts according to the mapping) [21].

Barret studies an extension of these games called syntactic games, where there are more states (and actions) than signals and more than one signal available at each turn [21]. By extending the game in similar ways, it might be possible to model emotion communication as well. For example nature states can correspond to emotions felt by the sender, signals to expressions of emotion in different modalities (ie vocal tone, facial expression, body language) and acts as an appropriate way for the receiver to respond to the sender's emotional state. For adults, we can assume a prior knowledge of a language for each player and think of the game as refinement of languages of the sender-receiver pair to match each other. The existing reward system would mean that correct choices of receiver's acts benefits both of the players. As this might not always be the case in naturalistic scenarios the reward function can also be modified.

Integrating emotions to game theory might be useful for game theory's development as well; for example, O'Neill identifies the lack of theories of emotion and emotional action as a gap, suggesting that this might be causing some distortions in the current approaches [22]. Similarly, Ketelaar examines the role of emotions in decision making and suggests that emotions can be used to understand choices for different strategies [23].

IV. A Speculative Combination of Theories

In mathematics it is recognized that theories can be connected to each other, and can be used to explain the same phenomena in alternative ways. The theoretical approaches mentioned here can work in conjunction, explaining different components or levels of emotional experience; they can also be applied independently without conflicting with each other, offering different insights.

Here we report a sample speculative high-level combination of the aforementioned theories. We extend Broeken's set theoretical approach to make it more specific for emotion communication of a pair of people. we examine an agent observe another agent, infer their emotional state, goals, emotion strategy, potential changes in the agent's own emotional state due to their inference of the other agent's emotional state, agent express their emotional state based on their emotion strategy and goals. We think of expressions as encoded messages of the emotional state. The relevant sets are as follows:

- The set of emotional components E contains dimensions of the emotional model (for example valence and arousal for Russell's circumplex model of affect [3]).
- The set of potential emotional dimension values PI. Its elements are tuples (sets with two elements), each containing an emotion dimension and an associated value within the range defined for that dimension. Assuming in the broadest sense values are real numbers, we can state that $i \in PI, i \in E \times \mathbb{R}$
- The set of all possible goals G.
- The set of all possible emotional expressions X with two subsets, one corresponding to voluntary expressions VX and another corresponding to involuntary expressions IX. For the sake of simplicity for this speculation, we will only work with voluntary expressions letting $X = VX$
- The set of all observables, termed as the world W, a subset of it is PW which are those perceptible by the agent.
- The set of mental objects O, with subsets of objects perceived by the agent PO and the agent's beliefs $B \subseteq G \times PI \times F$ which contains an estimation of the other person's emotion strategy function, their goals, and their belief on our own emotion strategy function and their emotional state.
- Emotional state of the Agent I, which is the current emotional state of the agent expressed as a set of tuples representing values of emotion representation and is a subset of potential dimension values PI
- The set of perceptive processes P , which are modeled as functions $p \in P, p : PW \mapsto PO$, explained below
- The set all possible emotion strategy functions F, $f \in F, f : G \times PI \times F \mapsto VX$, explained below
- The set of emotion transformations T, which are modeled as functions $t : PIxPI \mapsto PI$, explained below

Let A and O be an arbitrary agents, A being the one whose processes we are examining and O being an agent A observes. Subscripts $_a$ and $_o$ define to which agent the element is related.

- Let the other agent's emotional expression be perceptible by the agent, such that $X_o \in PW_a$
- Agent's perception of external stimuli are mapped to mental objects by perceptive processes, $p : PW_a \mapsto PO_a, p \in P$. In this case we can define a function for processing emotional expressions, $p_a \in P, p_a : X \cap W \mapsto X \cap PO_a$; given another agent's emotion expression $X_o \in X$, it creates a mental object representing that expression $p_a(X_o) \in X_o$.
- Let $b_o \in B_a \subset PO_a, b_o \in F$ represent the agent's belief of the other agent's emotion strategy choice.
- Prediction of the emotional state of the other agent then can be performed by using the inverse of the predicted emotion strategy function, b_o^{-1}. As we don't know if b_o is bijective, it might not have an inverse. But for the sake of simplicity in this high level speculation, let's assume its existence. This function, $b_o^{-1} : (PO_a \cap X) \mapsto B \subset PO_a, b_o^{-1}(p(X_o)) \in G \times PI \times F$ takes in the agent's observation of the other agents expressions, $p(X_o)$, to infer that agent's predicted goals $b_o^{-1}(p(X_o)))_g$, emotional state $b_o^{-1}(p(X_o)))_i$ and their belief of our own emotion strategy function $b_o^{-1}(p(X_o)))_{b_a}$.
- Transformation of the agents emotional state, can have many potential inputs but in this case we consider it having the other agent's predicted emotional state and the current emotional state of the agent as the input such that $t_a : PI \times PI \mapsto PI, t_a((b_o^{-1}(p(X_o)))_i, I_a) \in PI$, mapping from a tuple of potential emotional states to potential emotional states.
- Expression of the emotional state is a result of the agent's emotion strategy function f_a, that takes in agent's own emotional state, predicted emotional state of the other agent, agent's belief of the other agent's emotion strategy and produces an emotion expression, $f_a \in F, f_a : G \times PI \times F \mapsto X, X_a \in X, X_a = f(G_a, I_a, b_o)$.

For a visual representation refer to Figure 4. In this setup, elements of F, act both as the encoding function and its inverse as the decoding functions from an information theoretical view, which is in line with the approach of ToM. Game theoretical and information theoretical principles can be applied to probe the communication aspect (ie elements of F), dynamical systems' principles can be applied to probe the transformations of the emotional state (ie elements of T) and ToM can be applied to probe our mental model (ie elements of O and F).

We have to note that all the aforementioned theories require

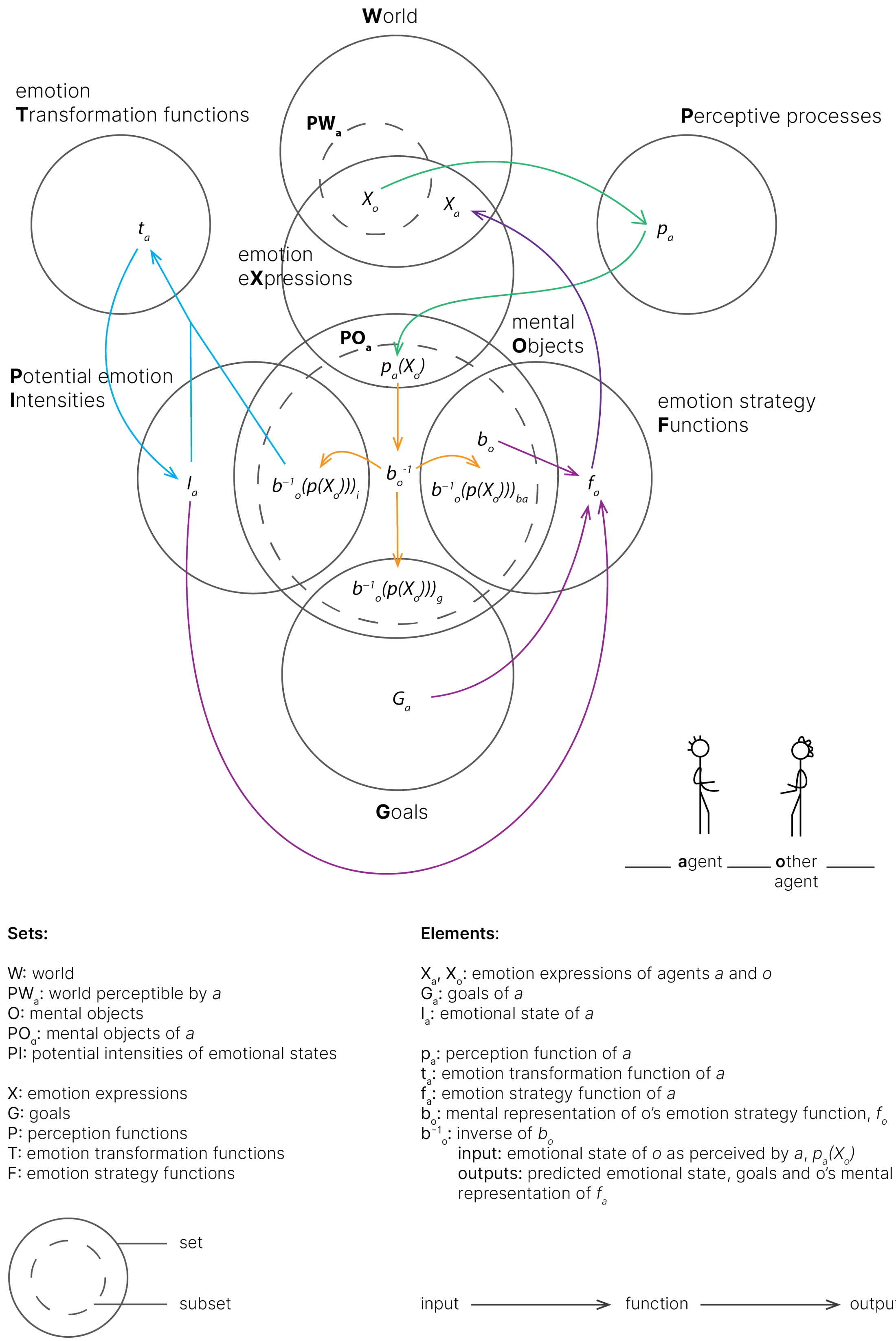

Fig. 4. A visual representation of the sets and the functions

definition of an emotional state. Recall that in Broekens' set theoretical approach [11] it is necessary to define emotion dimensions, in dynamical systems it is necessary to define the state of the system, in information theoretical approach it is necessary to define the information that is being transmitted and in Lewis' games it is necessary to define the set of states of the nature [20]. This brings us to the important question of modeling the affective space. However it is formed, experienced or communicated, how do we represent an emotional state? In Barrett's 2019 review on neuroscience of emotion, he suggests taking a constructivist approach, encouraging studies that "focus on whole-brain network dynamics along with time-varying, multivariate analytic approaches" [24]. His theoretical views align with the suggestion in theories we reviewed, as he calls for a dynamic systems based view that considers emotions as predictions [24]. This can lead to a data-driven approach based on neuroimaging.

V. Conclusion

In this paper we reviewed some of the challenges in defining and modeling emotions and some mathematical characterizations of emotional experiences. We then speculated on a way that combines the reviewed mathematical approaches, as an example. We think that it is promising to look at mathematical theories more deeply while modeling emotional experiences, as it sits between theory and computation. This approach can be thought of as developing a research-model [1] via a unification strategy [8]. We find the possibility of a cross-disciplinary approach bringing together philosophy, mathematics, affective and cognitive sciences very exciting. Such efforts would also be highly beneficial to HCI. For example once the emotional communication between agents is characterized, one of the agents can easily be replaced by an artificial one. This can contribute to programming affective responses of artificial agents, and inform the design of artifacts built for emotion regulation, reflection and mediation of human communication.

References

[1] E. Hudlicka, "From habits to standards: towards systematic design of emotion models and affective architectures," in *Emotion Modeling*. Springer, 2014, pp. 3–23.

[2] P. Ekman, "Basic emotions," *Handbook of cognition and emotion*, vol. 98, no. 45-60, p. 16, 1999.

[3] J. A. Russell, "A circumplex model of affect." *Journal of personality and social psychology*, vol. 39, no. 6, p. 1161, 1980.

[4] A. Scarantino and R. de Sousa, "Emotion," in *The Stanford Encyclopedia of Philosophy*, winter 2018 ed., E. N. Zalta, Ed. Metaphysics Research Lab, Stanford University, 2018.

[5] K. R. Scherer, "Appraisal theory." 1999.

[6] P. Goldie, "Emotions, feelings and intentionality," *Phenomenology and the Cognitive Sciences*, vol. 1, no. 3, pp. 235–254, 2002.

[7] P. H. Bucci, X. L. Cang, H. Mah, L. Rodgers, and K. E. MacLean, "Real emotions don't stand still: Toward ecologically viable representation of affective interaction," in *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 2019, pp. 1–7.

[8] R. Reisenzein, E. Hudlicka, M. Dastani, J. Gratch, K. Hindriks, E. Lorini, and J.-J. C. Meyer, "Computational modeling of emotion: Toward improving the inter-and intradisciplinary exchange," *IEEE Transactions on Affective Computing*, vol. 4, no. 3, pp. 246–266, 2013.

[9] R. Reisenzein, "Cognition and emotion: A plea for theory," *Cognition and Emotion*, vol. 33, no. 1, pp. 109–118, 2019.

[10] J. Bagaria, "Set Theory," in *The Stanford Encyclopedia of Philosophy*, spring 2020 ed., E. N. Zalta, Ed. Metaphysics Research Lab, Stanford University, 2020.

[11] J. Broekens, D. DeGroot, and W. A. Kosters, "Formal models of appraisal: Theory, specification, and computational model," *Cognitive Systems Research*, vol. 9, no. 3, pp. 173–197, 2008.

[12] G. Colombetti, *The feeling body: Affective science meets the enactive mind*. MIT press, 2014.

[13] N. Schwarz, "Feelings-as-information theory," *Handbook of theories of social psychology*, vol. 1, pp. 289–308, 2011.

[14] G. A. Van Kleef, C. K. De Dreu, and A. S. Manstead, "An interpersonal approach to emotion in social decision making: The emotions as social information model," in *Advances in experimental social psychology*. Elsevier, 2010, vol. 42, pp. 45–96.

[15] A. D. Kerr and K. Scharp, "The information in emotion communication," *arXiv preprint arXiv:2002.08470*, 2020.

[16] C. E. Shannon, "A mathematical theory of communication," *The Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.

[17] J. Pöppel and S. Kopp, "Satisficing models of bayesian theory of mind for explaining behavior of differently uncertain agents: Socially interactive agents track," in *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, ser. AAMAS '18. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems, 2018, p. 470–478.

[18] D. C. Ong, J. Zaki, and N. D. Goodman, "Computational models of emotion inference in theory of mind: A review and roadmap," *Topics in cognitive science*, vol. 11, no. 2, pp. 338–357, 2019.

[19] ——, "Affective cognition: Exploring lay theories of emotion," *Cognition*, vol. 143, pp. 141–162, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0010027715300196

[20] D. Lewis, *Convention: A philosophical study*. John Wiley & Sons, 2008.

[21] J. A. Barrett, "The evolution of coding in signaling games," *Theory and decision*, vol. 67, no. 2, pp. 223–237, 2009.

[22] B. O'Neill, "Approaches to modelling emotions in game theory," in *Proceedings of the Conference on Cognition, Emotion and Rational Choice, UCLA, April*, 2000.

[23] T. Ketelaar, "Ancestral emotions, current decisions: Using evolutionary game theory to explore the role of emotions in decision-making," *Evolutionary psychology, public policy and personal decisions*, pp. 145–168, 2004.

[24] L. F. Barrett and A. B. Satpute, "Historical pitfalls and new directions in the neuroscience of emotion," *Neuroscience Letters*, vol. 693, pp. 9–18, 2019, functional Neuroimaging of the Emotional Brain. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0304394017306171