Emotions and Context in Decision-making for a Socially Intelligent Agent

by

Aarti Malhotra

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Computer Science

Waterloo, Ontario, Canada, 2024

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Roland Goecke Professor, School of Information Technology & Systems, Faculty of Science & Technology, University of Canberra
Supervisor(s):	Jesse Hoey Professor, David R. Cheriton School of Computer Science, University of Waterloo
Internal Member:	Jeff Orchard Associate Professor, David R. Cheriton School of Computer Science, University of Waterloo Yuri Boykov Professor, David R. Cheriton School of Computer Science, University of Waterloo
Internal-External Member:	Mark Hancock Professor, Dept. of Management Science and Engineering, Faculty of Engineering, University of Waterloo

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Artificial intelligence (AI) research has largely focused on rational thinking, decision making, goal achievement, and reward maximization. Emotions have been considered non-essential, or even detrimental, when designing and implementing AI systems. With the advancement in affective computing research and the increasing adoption of AI agents as part of human society, there is a growing need to have a deeper connection between human and machine. Although significant efforts have been made in affective computing towards recognizing human emotions and generating human-like emotions, there has been less progress towards using emotions to guide decision and understanding human social context. This thesis focuses on emotions and context in decision-making, towards building socially intelligent agents, that are adaptive and emotionally aligned with humans.

I first conducted a systematic review of the literature on implemented systems for decision making that used emotions. I synthesized extracted data into four conceptual model types, viz. Matching, Appraisal and Coping, Decision-theoretic and Parametric and provided a process view of each type. Then, I implemented one such model as a brain-inspired neural model. The aim was to model the role of affect guiding decision-making, resulting in interactions that are similar to human interactions, while inhibiting some behaviors based on the social context. The model was implemented using Nengo, a python library for building and simulating large-scale neural models, using spiking neurons. I then investigated how to supply such a model with context, known to be a very important factor in emotional-based decision making. I proposed a computer vision spatio-temporal transformer model and its variations for joint learning and prediction, and evaluated on an existing Video Group Affect dataset. Improvements to social event prediction were shown by utilizing affective information. Finally, I considered a real-world care-giving scenario which demonstrates the potential of our model for establishing an emotional relationship and interaction between older adults, care partners, people living with dementia, and three exemplar robots.

The insights gained in this thesis may encourage AI and affective computing research to develop agents that can simulate human affective and decision-making mechanisms, and in the process understand humans better.

Acknowledgements

First and foremost, I would like to thank my supervisor, Prof. Jesse Hoey for his guidance, encouragement, and belief in me. He always brought the best in me. He inspired me to explore, search and research the outer world, and encouraged me to discover my 'self'.

I thank the committee members Prof. Jeff Orchard and Prof. Yuri Boykov, Prof. Mark Hancock and Prof. Roland Goecke for their feedback on my thesis work. I also thank my collaborators Prof. Abhinav Dhall, Dr. Terrance Stewart, Dr. Garima Sharma, Mr. Rishikesh Kumar, Dr. Jill Dosso, Rebecca Hutchinson for their amazing collaboration. My thanks to Prof. Shai Ben-David for giving me the opportunity to join the esteemed University of Waterloo and for the trust he had in me. I am also thankful to Prof. James Danckert and Prof. Myra Fernandes for their guidance in cognitive neuroscience. I thank Dr. Lisa Feldman Barrett for being an inspiration to me in continuing my quest for emotion research. She reminds me of my role model, Dr. Marie Curie. I am grateful to the team at PointClickCare, who supported me throughout my academic journey. I also thank my lab mates including Dr. Josh Jung and university colleagues for their support and camaraderie.

Finally, I would like to express my gratitude to the Almighty for His Grace. I feel blessed and indebted to my family for their unconditional support through thick and thin. The family comprising of grandparents Ms. Sushila and Mr. Satya Pal Malhotra, parents Ms. Meenakshi and Mr. Arun Malhotra, parents-in-law Ms. Shalini and Mr. Vijay Singh Thakur, brother Mr. Hemayur and his wife, Ms. Mehak and my loving and caring husband Mr. Satish Thakur.

Dedication

I dedicate this thesis to my SatGurus and my family.

Table of Contents

List of Figures xi

List of Tables xiii

1 Introduction 1

1.1 Socially Intelligent Agent . . . 3

1.1.1 Motivation . . . 4

1.1.2 Research Questions . . . 4

1.2 Contributions and Publications . . . 5

1.3 Chapter Summary . . . 8

2 Models of Emotions in Decision-making and Behavior 9

2.1 Background . . . 10

2.1.1 Psychology . . . 10

2.1.2 Neuroscience . . . 11

2.1.3 Sociology . . . 12

2.1.4 Related Surveys . . . 13

2.2 Conceptual View . . . 14

2.2.1 Learning and Adaptive . . . 16

2.2.2 Appraisal . . . 17

2.2.3 Role of emotion . . . 17

2.2.4 Choice Exploration . . . 18
2.2.5 Policy Computation . . . 18
2.3 Survey Methodology . . . 18
2.3.1 Identification of Research . . . 19
2.3.2 Selection of Studies . . . 20
2.4 Data Extraction . . . 22
2.4.1 Design and Contents of Data Extraction Chart . . . 23
2.4.2 Data Extraction Procedures . . . 23
2.4.3 Supplementary information . . . 24
2.5 Data Synthesis . . . 24
2.6 Discussion . . . 28
2.7 Conclusion . . . 29

3 Affect Control Theory **39**
3.1 Affective Semantic Space . . . 39
3.2 Social Interaction . . . 40
3.3 Deflection . . . 41
3.4 Transient impression formation . . . 42
3.5 Optimal Behavior . . . 43

4 Biologically-inspired model of ACT **45**
4.1 Human Brain and Neural Communication . . . 46
4.2 Cortex-Basal Ganglia-Thalamus loop . . . 47
4.3 Emotions in the brain . . . 47
4.4 Nengo . . . 48
4.5 Neural Model of ACT agent . . . 49
4.6 NeuroACT Simulation . . . 53
4.7 Results . . . 54
4.8 Conclusion and Future Work . . . 56

5 Social Context Understanding 58
5.1 Background 60
5.1.1 Social Event Prediction 61
5.1.2 Group Affect Prediction 61
5.1.3 Video Vision Transformers 62
5.2 Dataset 63
5.3 Network Architecture 64
5.3.1 Base Transformer 64
5.3.2 Multi-Task Learning Variations 65
5.4 Experiments 66
5.4.1 Train-Val-Test split 66
5.4.2 Experimental Setup 67
5.5 Results and Discussion 68
5.5.1 Quantitative Results 68
5.5.2 Qualitative Results 70
5.5.3 Limitations 72
5.6 Ethical Impact 72
5.7 Conclusion and Future Work 73

6 Application 74
6.1 ACT Simulation for Social Robot 74
6.2 ACT modelling 76
6.2.1 EPA profiles 76
6.2.2 EPA distance and robot judgement 77
6.2.3 Deflection and robot usefulness 78
6.3 Conclusion and Future Work 78

7 Conclusion **80**

7.1 Socially Intelligent Agent 80

7.2 Limitations of ACT 81

7.3 Ethical Impact of SIAs 81

7.4 Closing Thoughts 82

References **83**

APPENDICES **109**

A Systematic Review Search Queries **110**

A.1 Scopus 110

A.2 ACM Digital Library 111

A.3 PsycINFO 112

A.4 IEEE Xplore Digital Library 113

List of Figures

1.1 SIA in an environment 3
1.2 Outline of this thesis and corresponding chapters 5

2.1 Hypothetical decision-making scenario used to explain terms. 15
2.2 Five features of a decision-making affective agent. Each feature has a number of options, giving in total 144 possible models which we will reduce to 4 categories in section 2.5. 16
2.3 PRISMA flow diagram for our systematic review. 22
2.4 Distribution of papers by publication year 23
2.5 High level process view of agents of four model types (**Matching**, **Appraisal and Coping**, **Decision-theoretic** and **Parametric**) 25
2.6 Distribution of models by model type 28

3.1 3D affective semantic space of concepts called *EPA (Evaluation Potency Activity)* profile. The values of a dimension indicate its intensity on a scale ranging from -4.3 to 4.3 values. 40

4.1 A neuron structure, from [Barrett, 2017] 46
4.2 The cortex-basal ganglia-thalamus loop, from [Eliasmith, 2013] 48
4.3 Nengo GUI 49
4.4 Neural implementation of ACT 50
4.5 Example of behavior of NeuroACT 52
4.6 Human agent cooperates with AI agent, with change in emotion 55

4.7 Human agent defects with AI agent, with change in emotion 55
4.8 Human agent defects with AI agent, with change in emotion and identity . 56

5.1 Sample frames from VGAF dataset [Sharma *et al.*, 2021] video clips with social event search keyword and perceived affect label. 58
5.2 Illustration of network architectures for a. **ViTR**, b. **ViTR-AR** and c. **ViTR-AD** models. 64
5.3 Confusion Matrix for social event prediction for all three ViTR models. . . 69
5.4 Actual and predicted social event label - *Casual Friends Family Gathering*. Actual and predicted affect label - *Positive* 71
5.5 Actual and predicted social event label - *Group Activities*. Actual and predicted affect label - *Neutral* 71
5.6 Actual social event label - *Fighting*, predicted social event label - *Sports*. Actual and predicted affect label - *Negative* 71
5.7 Social event and affect prediction for some clips using ViTR-AR and ViTR-AD models. (a) and (b) are correct social event predictions, (c) is an incorrect social event prediction. All three clips have correct affect prediction. . 71

6.1 Robots used in the study. (a) AIBO; (b) Joy for All Cat; (c) MiRo-E. . . . 75
6.2 Sentiments associated with the three socially assistive robots. Means plus standard errors are shown. 76
6.3 Relationship between responses to three statements about the social robots (I would enjoy. . . , I would use. . . , I would find useful. . .) and respondents' self-robot EPA distance. Means and standard errors are shown. 77
6.4 Relationship between agreeing that a robot would be useful and ACT deflection for the concept of being assisted by the robot. Means and standard errors shown. 79

List of Tables

2.1 Part 1: Data extraction of five features of emotions in decision-making agent in papers analyzed. Papers related to same model are grouped by round brackets. The major row sections correspond model types, from top to bottom: **Matching**, **Appraisal and Coping**. Colors are model types, and exceptions are shown in cross-hatched grey color (otherwise the value is that at the top of the column in the same color). 30

2.1 (contd.) Part 2: Data extraction of five features of emotions in decision-making agent in papers analyzed. Papers related to same model are grouped by round brackets. The major row sections correspond model types, from top to bottom: **Decision-theoretic** and **Parametric**. Colors are model types, and exceptions are shown in cross-hatched grey color (otherwise the value is that at the top of the column in the same color). 31

2.2 Summary of model implementation characteristics for ***Matching*** model type 32

2.3 Summary of model implementation characteristics for ***Matching*** model type (contd.) 33

2.4 Summary of model implementation characteristics for ***Appraisal and Coping*** model type 34

2.5 Summary of model implementation characteristics for ***Appraisal and Coping*** model type (contd.) 35

2.6 Summary of model implementation characteristics for ***Decision-theoretic*** model type 36

2.7 Summary of model implementation characteristics for ***Decision-theoretic*** model type (contd.) 37

2.8 Summary of model implementation characteristics for ***Parametric*** model type 38

3.1 Sample ACT interactions 44

5.1 Social event categorization based on keywords used in search queries of the VGAF dataset [Sharma *et al.*, 2021]. 63

5.2 Dataset split for social event categories. 67

5.3 Dataset split for affect categories. 67

5.4 Experimental results for Social event prediction. 68

5.5 Classwise accuracies (in %) for social event prediction. Here, TS = TimeSformer, TSHR = TimeSformerHR, Q = Quarrelling, M = Meeting, SP = Sports, FL = Funeral, SH = Show, GA = Group Activities, P = Protest, FFG = Casual-FFG, C = Celebration Party, F = Fighting. 69

5.6 Experimental results for Affect prediction. 70

A.1 Scopus queries and results (October 20th, 2021, 11am) 110

A.2 ACM queries and results (October 27th, 2021, 5pm) 111

A.3 PsycINFO queries and results (October 20th, 2021, 1:30pm) 112

A.4 IEEE Xplore Digital Library queries and results (October 20th, 2021, 11:30am) 113

Chapter 1

Introduction

"All the world's a stage, And all the men and women merely players; They have their exits and their entrances, And one man in his time plays many parts"

- William Shakespeare

The above words by William Shakespeare beautifully describe the social reality and how we, as humans, play various roles and do our part. Emotion, affect, sentiment, feeling are a core part of our being and they guide our decisions and behavior. Knowledge of context in a situation gives people a holistic view and helps them in their day-to-day interactions. Understanding the nuances of human social interaction remains a challenging area of research, but there is a growing consensus that *social identity*, a person presents in a given context, is a critical aspect [Goffman and others, 1978; Schröder *et al.*, 2016; Konig *et al.*, 2016; Khan and Hoey, 2017; Konig *et al.*, 2017; Francis *et al.*, 2019]. Therefore, understanding the social context, emotions and behaviors displayed by humans, and being able to maintain an identity itself, are crucial skills for a Socially Intelligent Agent (SIA).

In human-human interaction, decades of anthropological and sociological research have shown the importance of heavily context-dependent social *identities* and roles [Goffman, 1959; Stets and Serpe, 2019; MacKinnon, 1994]. Identities are usually defined linguistically, and carry significant emotional weight. If a 'doctor' is advising a 'patient' in a 'hospital', both present embodied, with culturally defined and shared identities that are appraised by the other and by the self. The 'hospital' context pervades the entire interaction. The actors in this situation will adhere to social norms and prescriptions that are conditional on identities and context for selecting emotional signaling and concrete actions.[1] Doctors, for example, are socially expected (in western cultures) to be authoritative and directive. Patients, on the other hand, are expected to be deferential and accepting.[2] Should those same two individuals find themselves in another situation, say the patient is the judge in a case brought forward by the doctor (ignoring the apparent conflict of interest in this constructed situation), their expected behaviors may change. If behaviors do not correspond in this way to the identities at play, then tension is sure to result. Such tension can arise from uncertainties in the situation caused by environmental noise (e.g., someone mishears something), or by lack of information (e.g., the context is not well defined). For example, if the doctor *upbraids* the patient for not taking his/her medicine on time, it may impact the interaction negatively. If the patient chooses a behavior (e.g. *apologize*), then a reduction in this tension may be possible, and the interaction can be brought back to normal. If the same two people meet at a wedding reception or a sports event, the different context may give rise to uncertainty in what identities should be adopted.

Therefore, on top of maintaining an identity of their own based on the context and culture, agents also need the ability to detect the identity and emotions of a human with whom they are interacting. For example, if an agent is in a hospital setting and is given a task to take care of a patient, it should know that, in the context of the hospital, it should adopt the identity of a 'care-giver', and when it is interacting with a human, the knowledge that the other is a 'patient' and the setting is a 'hospital' will establish the social context in which the agent must act. If the human is a 'nurse' or a 'doctor', then the agent may have to interact differently. An agent given a task of teaching a 'student' in a 'classroom', on the other hand, should know that in the context of a classroom, it should take on the identity of a 'teacher' and that the human being will be acting according to the identity of a 'student'. Interactions that do not conform to the cultural prescription will result in restorative behaviors (as when the patient apologized to the doctor), or re-identifications (the patient starts thinking of the doctor as his 'tormenter' and himself as a 'victim'). These transient, context-dependent, identities can become salient and highly significant in a non-coherent interaction, from a cultural prescription perspective. Context in humans may involve any situational factors such as location, event, people, cultural, political, economical, etc., or their internal state of being. In this thesis, I focus on social event context and emotions to guide decision-making for a Socially Intelligent Agent (SIA).

[1] We distinguish between *norms* and *prescriptions* as the former are more rule-like (e.g. you must not 'butt' into a lineup, or, in ancient times, a man was expected to open a door for a woman), and may invoke sanctions (e.g. you must stop at a stop-sign even if there is no one around), while prescriptions are emotional (e.g. you must be deferential to your parents).

[2] While this may be true in a hospital context, the behavior expectations would be substantially different in a different social context and in other cultures [Henrich, 2020].

A comprehensive computational model of emotions in social interaction is *Affect Control Theory (ACT)* [Heise, 1987; MacKinnon, 1994; Heise, 2007; MacKinnon and Robinson, 2014], which includes components for identity, normative behaviors, prescriptive emotions, context, and transient sentiments. Although ACT is based on human empirical data from various cultural groups and is quite promising, it is yet to be adopted in AI applications. Some challenges for ACT-type models exist, which if tackled, may help towards its long-term adoption in various domains such as healthcare, home, business, academic etc., where human interactions and social contexts are important. In this thesis, I propose using ACT for SIA and capabilities to perceive social scene and infer social context and affect to make appropriate decision or behave with human(s) in an emotionally aligned manner.

1.1 Socially Intelligent Agent

Human beings live in a socio-cultural environment, which influences and shapes human perceptions, thoughts, actions and vice versa. This process of continuous and dynamic interaction keeps on evolving with time. In that sense, it was rightly said by Heraclitus that "*The Only Constant in Life Is Change.*" A human being who can adapt to the changing society and/or change the society such that it adapts to the human can survive better. With scientific and technological advancements, we are able to develop AI agents who can become part of this society, perform required tasks autonomously or in an assistive fashion, and/or interact with humans. But they still lack adaptability and understanding of the culture, context, and emotions. Figure 1.1 shows a simplistic view of the main theme of this thesis, which aims to explore the role of emotions and context in decision-making for an adaptive SIA in an environment.

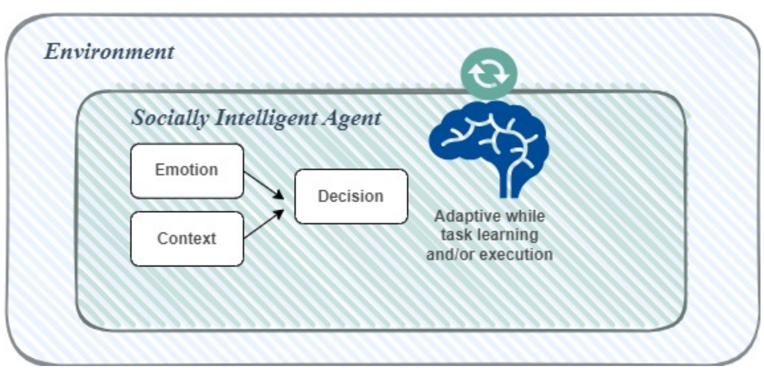


Figure 1.1: SIA in an environment

1.1.1 Motivation

Since my childhood days, I was inspired by a television series called *'Small Wonder'*[3] . It was about a robot named V.I.C.I. (Voice Input Child Identicant), built by a robotics engineer, who embodies it in a human-like form, like his daughter. He wanted the robot to learn and evolve from the social environment at his home with his family and neighbors. They portray her as their adopted daughter and not a robot. She interacts with people around her, helps the family with various kinds of tasks, learns behaviors and but lacks emotional connection. The emotions displayed by her are still robotic in nature. I was fascinated by the superhuman abilities that was given to a human-like robot and always dreamt of designing a better version of V.I.C.I., one that is more emotionally aligned with humans and socially intelligent. Building an SIA is the motivation behind this thesis.

1.1.2 Research Questions

This thesis outlines four aspects of SIA that can be considered while designing and developing such agents, as depicted in Figure 1.2. They are as follows: models of emotion in decision-making, brain-inspired implementation of such a model, context understanding, and modeling for a real-world scenario. There can be other aspects as well, but they are out of scope in this work. This thesis aims to explore below research questions (corresponding chapters provided in brackets):

RQ1: What are the existing computational models of emotion which have an AI implementation and consider emotion to guide decision-making and behavior? (Refer Chapter 2 for systematic review)

RQ2: Can ACT, which is an existing socio-psychological model of decision-making be implemented as a biologically-inspired neural model of decision-making for SIA? (Refer Chapter 3 for background on ACT and Chapter 4 for neural implementation of ACT)

RQ3: How well can a computer vision system infer social event context and perceived affect, as required by SIA, in a visual scene? Does affect improve social event classification? (Refer Chapter 5 for scene context understanding)

RQ4: How can a SIA be designed and simulated in a realistic scenario using ACT? (Refer Chapter 6 for modeling user study)

[3]https://www.imdb.com/title/tt0088610/

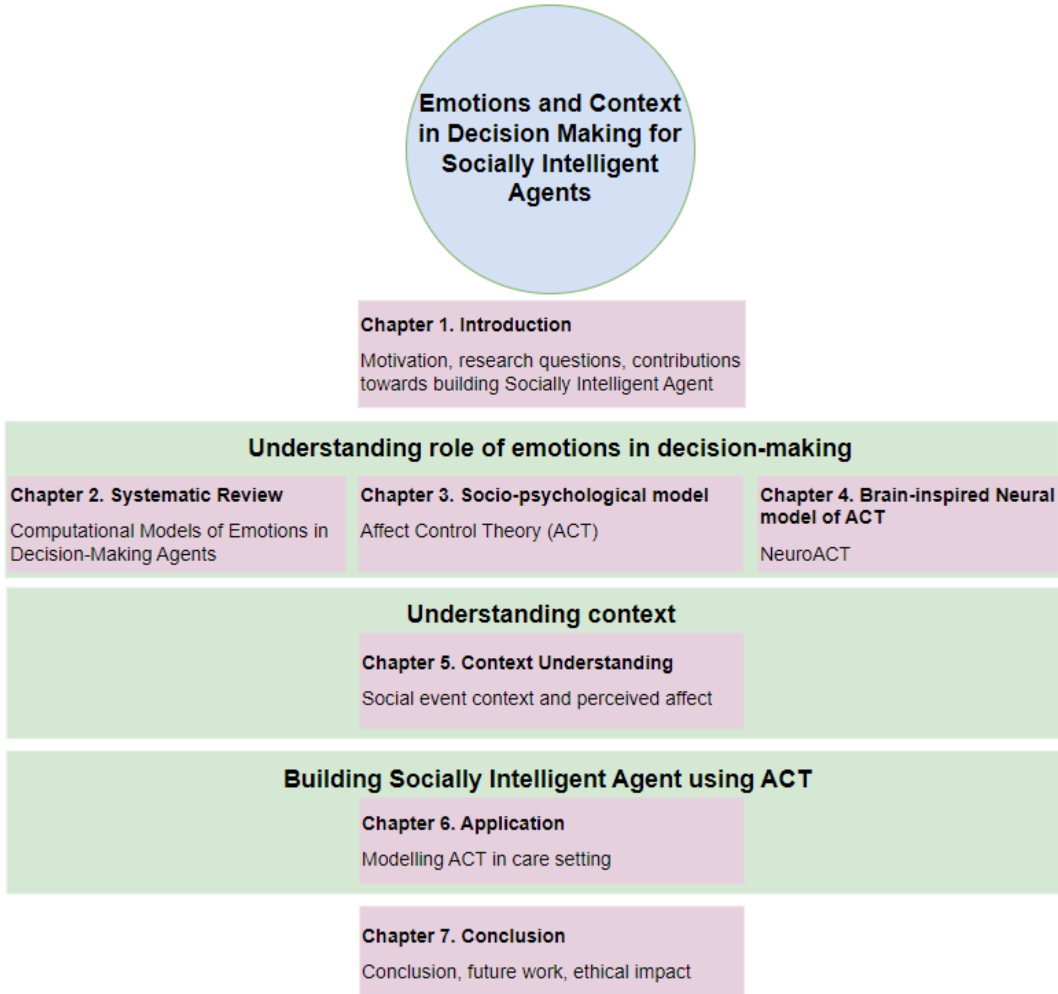


Figure 1.2: Outline of this thesis and corresponding chapters

1.2 Contributions and Publications

The main contributions based on the above research questions are:

1. Comprehensive systematic review of literature for computational approaches to emotions in decision-making in AI agents

2. Implementing a biologically-inspired neural model of emotions in decision-making in social interactions, i.e. ACT, to be used by SIA

3. Building a visual perception component for SIAs towards social scene understanding and infer social event context and perceived affect.

4. Modeling ACT in a real-world interaction scenario for care-giving robots (being the SIAs) interacting with patients with dementia, human care-givers and older adults

A list of particular chapters related to the above four contributions and the corresponding articles published and/or submitted with my collaborators are given below, along with details regarding my contribution.

- **Chapter 2: Models of Emotions in Decision-making and Behavior**

  Paper(s): **Aarti Malhotra**, Jesse Hoey, Rebecca Hutchinson and Gabby Chan. *Computational Models of Emotions in Decision-Making Agents: A Systematic Review*. Submitted to Journal of IEEE Transactions on Affective Computing. [Malhotra *et al.*, Under Review]

  Contributions: **AM**, JH devised the survey. RH and AM worked on the search query formation. RH queried all listed databases. **AM** and GC did title and abstract screening. **AM** did full text screening, data extraction, data synthesis, paper writing. JH reviewed queries, screening, data extraction, synthesis and paper writing.

- **Chapter 4: Biologically-inspired model of ACT**

  Paper(s): **Aarti Malhotra**, Terrence C. Stewart, and Jesse Hoey. *A Biologically-inspired Neural Implementation of Affect Control Theory*. International Conference on Cognitive Modelling, Toronto, 2020. [Malhotra *et al.*, 2020]

  Contributions: **AM** conceptualized the design and worked on the ACT implementation and simulation scenarios. TS assisted with gaining understanding of the Nengo library usage for the implementation, reviewed the code and provided feedback. **AM** did the paper writing and TS and JH reviewed the same.

- **Chapter 5: Social Context Understanding**

  Paper(s): **Aarti Malhotra**, Garima Sharma, Rishikesh Kumar, Abhinav Dhall and Jesse Hoey. *Social Event Context and Affect Prediction in Group Videos*. In Workshop on Addressing Social Context in Affective Computing, ACII 2023. [Malhotra *et al.*, 2023]

  Contributions: **AM** conceptualized the study and built the hypothesis. **AM**, GS and RK annotated the dataset with event labels. **AM** worked on the open sourced code and created the main model and its variants. RK and GS reviewed the code in detail. **AM** did paper writing. All authors reviewed the paper.

- **Chapter 6: Application**

  Paper(s):

  Jill Dosso, Ela Bandari, **Aarti Malhotra**, Jesse Hoey, Francois Michaud, Tony Prescott, and Julie Robillard. *Towards emotionally aligned social robots for dementia: perspectives of care partners and persons with dementia*. The Journal of Alzheimer's & Dementia, 2022. [Dosso *et al.*, 2022b]

  Jill Dosso, Ela Bandari, **Aarti Malhotra**, Gabriella Guerra, Jesse Hoey, Francois Michaud, Tony Prescott, and Julie Robillard. *User perspectives on emotionally aligned social robots for older adults and persons with dementia*. The Journal of Rehabilitation and Assistive Technologies Engineering, 2022. [Dosso *et al.*, 2022a]

  Contributions: JD constructed the survey, led participant recruitment, performed data coding, analysis, and visualization, and wrote the first draft of the manuscript. EB and JR researched the literature, conceived the study, and designed and piloted early study materials. JR conceptualized the project and supervised the work. **AM** and JH performed data analyses related to Affect Control Theory. FM and TP provided study materials. GG performed data coding and checking. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

1.3 Chapter Summary

The rest of the chapters of this thesis are as follows. Chapter 2 provides a systematic survey of computational models of emotions in decision-making and behavior. It also gives some background on psychological, neuroscience and sociological perspectives. Chapter 3 provides background and details on ACT. Chapter 4 introduces biologically-inspired neural model of ACT. Chapter 5 gives details about a new predictive computer vision model of social event context and perceived affect in videos with group of people. Chapter 6 provides ACT modelling in a real-world scenario. Chapter 7 provides conclusion, limitations, future work and ethical impact.

Chapter 2

Models of Emotions in Decision-making and Behavior

> *"The intuitive mind is a sacred gift and the rational mind is a faithful servant. We have created a society that honors the servant and has forgotten the gift."*

- Albert Einstein

Emotions can influence decision-making, learning, and other aspects of human behavior [Bechara *et al.*, 2000; Feldman *et al.*, 2022]. Emotion (or *affect*) is a property of consciousness [Barrett and Satpute, 2019] and a part of every psychological phenomenon, even those that are not explicitly emotional [Hutchinson and Barrett, 2019]. Artificial Intelligence (AI) systems that involve affective social interactions with humans have largely focused on detecting and generating emotions. However, the use of emotions for decision-making in AI systems is a less well studied problem. In recent times, popular conversational AI chatbots like ChatGPT[1] have also stayed away from incorporating emotions. Modelling emotions in artificially intelligence systems could give insights about human affective process, and would be a step towards systems being more human-like and less robotic. Our view is that incorporating emotions into decision-making and behavior in SIAs may allow these agents to be more aligned with humans.

With the primary objective of understanding the relationship between emotions and decision-making in implemented systems, in this work we present a systematic review

[1] https://openai.com/blog/chatgpt

on the computational models that have leveraged emotions for decision-making in the form of an AI system or simulation. Based on extracted emotional and decision-making features, we perform data synthesis and categorize these into four high level model types, and provide a process view of how each type uses emotions in decision-making. We also provide supplementary information which may be useful for practical development.

This chapter is organized as follows: Section 2.1 provides background on psychological, neuroscience and sociological perspectives on how emotions may be influencing human decision-making. It also outlines related surveys. Section 2.2 introduces five features of emotions in decision-making agents used for data extraction and provides a conceptual view with a hypothetical scenario. Section 2.3 gives details on our survey methodology, which includes search strategy, inclusion/exclusion criteria, and selection process. Section 2.4 provides the data extraction procedure and content for the final selection of papers based on the five features and supplementary information. Section 2.5 focuses on data synthesis and model type categorization. It gives the process view of each type. Section 2.6 provides discussion and limitations of the survey. Section 2.7 concludes the chapter. Appendix A provides search queries.

2.1 Background

We give a brief background from various studies in psychology, neuroscience and sociology, that provide some evidence of influence of emotions in human decision-making.

2.1.1 Psychology

Psychological experiments have demonstrated that, in task performance, positive affect becomes experienced as efficacy and negative affect as difficulty, with predictable consequences for an individual's cognition and goals, e.g., a happy mood makes hills look less steep [Schwarz and Clore, 1983]. Bechara *et al.* [2000] have shown that emotions influence decision-making, learning, and other aspects of human behavior. Many views of emotion, such as appraisal theories, were based on a stimulus-response model that posits a response being a result of assigning meaning to a stimulus [Lazarus, 1991; Scherer, 2009; Roseman, 2011]. Some effects of fleeting incidental emotions over time on decision-making are explored in [Andrade and Ariely, 2009]. The study provided empirical evidence using a sequence of ultimatum and dictator games, for the enduring impact of transient emotions on economic decision-making, suggesting that people tend to behave consistently with past actions and cognitions, earlier choices, that are unconsciously based on a fleeting incidental emotion can become the basis for future decisions and hence outlive the original cause for the behavior (i.e., the emotion itself). Hoelzl and Loewenstein [2005] suggest that both anticipated regret and social takeover (i.e., in sequential decision-making, knowing that someone else might step in and assume an investment that one has decided to quit) can influence decision-making with regards to keeping an investment. Research such as [Hoelzl and Loewenstein, 2005] reviews the implications of moral standards and moral emotion for moral decisions and moral behavior. A study [Fong and Wyer Jr, 2003] found that the participants' decisions in a simulated investment situation and an academic situation were mediated by the emotional reactions they experienced in response to alternative decision outcomes. Haidt [2001] argued that people tend to form moral judgments rapidly and effortlessly and, in turn, generate rationales that support their affect-laden judgments. Emotion as information perspective can be found in [Frijda, 1988; Schwarz, 1990; Schwarz and Clore]. van Dijk and Van der Pligt [1997] reason that positive emotions, like elation, signal that everything is fine and no further action is required. On the other hand, negative emotions, like disappointment, signal a problematic state of affairs and a need for careful information processing.

2.1.2 Neuroscience

According to the classic James-Lange theory [Lange and James, 1967], the perception of an emotion provoking stimulus leads to an emotional behavior, which then leads to conscious experience of the emotion. Damasio supported the role of emotions in judgment and decision-making, and for human social action [Damasio, 1994]. His 'somatic marker hypothesis' is related to the learned neural bias based on social conventions and ethics, which helps in working with fewer alternatives and guides decision-making. The cases of Phineas Gage, Elliot and others suffering from frontal lobe lesions, have much in common when it comes to failure in the personal and social domain, due to a lack of affect associated with an event, although the knowledge of affect remains. These cases were clear illustrations of the significance of emotions in human decision-making and behaviors. Gray [1994] presents evidence supporting the existence of neuroanatomical components that mediate three fundamental behavior-coordinating systems: approach, avoidance and, fight or flight. These systems reflect corresponding categories of affective reactions: positive affect and approach behavior; negative affect and avoidance, and anger/aggression and fear. Davidson *et al.* [2000] presents evidence about the location of some of these systems in the brain: the behavioral approach system in the left anterior cortex, and the behavioral inhibition system in the right anterior cortex. Recent perspectives follow a constructivist view where the brain constructs concepts and categorizes incoming sensory events as similar based on past experiences. When past experiences of emotion are used to categorize the predicted sensory array and guide action, then one perceives that emotion [Barrett, 2016]. Emotions are varied in nature and cannot be localized to a single neural circuit. When it comes to the observed pattern for any single emotion category, variation is the norm. The mechanisms that implement constructions of emotions are not specific to the domain of emotion but operate across cognitions, perceptions, and action [Barrett and Satpute, 2019]. The impact of social factors on decisions has been demonstrated with many different decision tasks, and there have been numerous investigations of the neural circuitry and neurochemistry mediating these effects (see [Rilling and Sanfey, 2011; Phelps *et al.*, 2014] for reviews). A review of the overlap in the neural circuitry of race, emotion and decision-making is provided in [Kubota *et al.*, 2012]. Some studies have considered the impact of social roles or identities in social interaction. EEG-based hyperscanning studies showed asymmetric brain-coupling patterns of leader-follower participants in a dyad during coordinated movements [Dumas *et al.*, 2010]. The asymmetric phenomena were also emerged from some studies of decision-making in game contexts [Balconi and Vanutelli, 2016]. This asymmetric pattern of coupling may be explained by the differential roles of the partners during the interaction, and the participants may have different expectations for the assigned roles.

2.1.3 Sociology

Historical and cultural variability suggests that, to an important degree, subjective experiences and emotional beliefs are both socially acquired and socially structured [Thoits, 1989]. In this view, emotions are signals to the self, but also they are signals to others and are the objects of others' responses, primarily about social roles or 'identities'. This view of emotions starts by treating humans as primarily *social* animals and fundamentally group-oriented, rather than individualistic and rational ones. Strong and persistent ties in human networks are *relational* rather than *transactional* [Lawler *et al.*, 2009]. In this view, rationality exists at the level of groups of agents, not of individuals. Intelligence is defined by a social order that exists in a group and is internalized by each member through *affective dynamical structures* of *roles* or *identities*. Members of a group learn these structures as children, growing to assume a set of identities within the structures as adults. Members seek out other members of the group that play complementary roles, and enact a joint behavior for their chosen relationship. Small-scale breakdowns are handled through a restorative set of multi-modal communicative cues that are displayed in the voice, face, gestures, and body, and are commonly referred to as *'emotion'* signals. Larger-scale breakdowns are handled by cognitive skill in creating new structures that are reified and internalized by group members [Berger and Luckmann, 1967]. The dynamics of the role relationships, coupled with human ability to cognitively explore, in a time- and energy-bounded fashion, using reason and rationality, allow the entire group to build, maintain, enact, and transform a *social order* [Goffman, 1963] that is jointly optimal for survival.

Affect Control Theory (ACT) [Heise, 2007] operationalizes this viewpoint using a model of emotional coherence based in language that was founded on the control principle of [Powers, 1973], which states something very reminiscent of the free energy principle [Friston, 2010]: that people try to minimize incongruencies by controlling their perceptions. Heise tranposed this to the sentiment space of [Osgood *et al.*, 1957a], imposed a denotative structure from symbolic interactionism [Mead, 1934], and added affective dynamics [Gollob, 1974]. ACT is a computational model that has been used to predict classes of human behavior in a variety of settings [Heise and MacKinnon, 2010].

2.1.4 Related Surveys

We provide a brief overview of some existing works which have surveyed computational models and applications, that use emotions in autonomous agents. In [Rumbell *et al.*, 2012], twelve autonomous agents, incorporating an emotion mechanism into action selection are selected partially arbitrary and analyzed based on characteristics of the agent architecture, the action selection mechanism, the emotion mechanism and emotion state representation, along with the emotion model it is based on. A narrative summary of empathic virtual agents and robots as observers generating empathy in users or as targets evoking empathy in users, is given in [Paiva *et al.*, 2017]. A short review that focuses on how agents show emotion while communicating with other agents is provided in [Martınez-Miranda and Aldea, 2005]. A review of single-user and group recommender systems is provided in [Tran *et al.*, 2021], with a focus on influence of personality, emotions and group dynamics. A short review on academic emotions and student engagement models is provided in [Subramainan and Mahmoud, 2020]. A systematic review of works related to social robotics as a tool in autism therapy is given in [Pennisi *et al.*, 2016]. A brief review of a eight selected computational models of emotions compared by affective processes and theoretical foundations is given in [Rodriguez *et al.*, 2011]. A short review of recommender systems, which compares filtering techniques, affective mode and recommended items is provided in [Raheem and Ali, 2020]. A review on eleven architectures of models of emotions for intelligent agents in crisis simulations, comparing on the underlying theories, application, advantages and disadvantages is given in [Loizou *et al.*, 2012].

A survey of computational models of emotion in reinforcement learning (RL) agents, comparing emotion elicitation, emotion type and emotion function is provided in [Moerland *et al.*, 2018]. A review of intrinsic motivation and emotion in player modeling, focused on simulation-based game testing is outlined in [Roohi *et al.*, 2018]. A critical analysis of computational models, frameworks or architectures for artificial emotion generation was done in [Ojha *et al.*, 2021] to examine if they satisfy five properties of domain-independence, integration of mood, integration of personality, data-driven mapping of appraisals into emotion intensities, and ethical reasoning for emotion regulation. A review on affective recommender systems is done in [Katarya and Verma, 2016], classifying based on perspectives such as research gaps, nature, algorithm or method adopted, datasets, the platform on executed, types of information and evaluation techniques applied. A review on 27 studies based on e-learning environment used, the categories of the emotional states measured, the affective measuring methods, modalities and the major research purpose is done in [Mejbri *et al.*, 2022].

All the related surveys mentioned above have either hand-picked models or have done a non-systematic search for emotional models. In this chapter, we do a systematic search in four databases, perform detailed data extraction and data synthesis, giving both an in-depth view as well as a high level categorization of models. As far as we know, this is the first comprehensive systematic review of computational models of emotions, implemented in a system or simulation, and used for decision-making.

2.2 Conceptual View

In this section, we clarify the features used for classifying the surveyed papers. We first use a toy example to give our interpretation of terms, and then describe each such term as a feature that we subsequently array the surveyed papers along. Throughout we consider that we are dealing with a situated agent: one that has some mental model of: its own state and actions; its context (environment); its preferred states (rewards); and of the temporally deep relation between itself, its context and rewards. Features will be shown in **bold**, while other relevant terms are *italics*.

Consider a human (say Mr. Bean), who has to perform a task of driving his car from home to a supermarket. There may be many decisions that he will have to take in order to complete the task. This is his **'Policy'**, which consists of step-wise decisions for task completion based on context. Consider there are four routes A, B, C and D for Mr. Bean to choose from (Figure 2.1). These are Mr. Bean's **'Choices'**. If he has driven to that store before, he may have a mental map of the routes and any context that he remembers of it. If he has never been to that store, he may gain a mental model of his own self and the environment he is in, as he experiences for the first time. This is his internal model. If his self state and/or environment context changes, he may need to update his internal model to match. This will mean he is **'Learning and Adaptive'**. In Mr. Bean's case, if the shortest route A which he usually takes is now closed for construction, he will need to appraise the situation (termed as **'Appraisal'**), decide about which other route would be best for him, update the policy and also update his internal model for the changed environment. When deciding amongst three other options, say route B, C or D, he can use a mental simulation to explore some or all of the choices, or he can avoid the simulation and choose one (say at random) and go with it. Mr. Bean's emotions or affective feelings associated with any concept (be it person, place, behavior, choices, etc.) gives him a subjective signal, a *somatic marker* [Damasio, 1994], that may help in guiding him to make a choice. This is the **'Role of emotion'** in his situation. These markers, thought to be largely encoded in frontal lobes in the human brain, are considered to be associated with specific contexts or plans, and give fast evaluations of appraised situations. For Mr. Bean, if route C had a lot of potholes and was a bad experience for him before, he will have negative somatic markers associated with it. He can use these markers to simply eliminate route C from his available options, narrowing his focus to options B and D, and simplifying his subsequent decision process. Emotional process or affect in that sense is an ongoing process. It can be unconscious (which may not be appraised) or conscious (which may be appraised). If emotions are appraised after a decision is made, it may help in evaluating the decision and take corrective action.

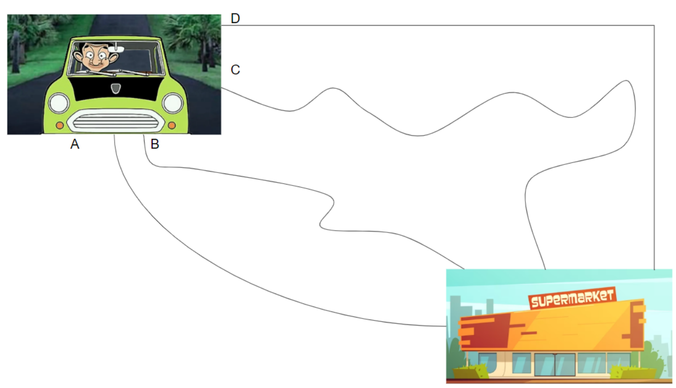


Figure 2.1: Hypothetical decision-making scenario used to explain terms.

Mr. Bean's example was a hypothetical scenario, deliberately simplified to help solidify the interpretation of terms we use in this chapter. If the situation is *social*, e.g., replace the supermarket with a social objective involving a group of agents, and the paths with social etiquette, then one can see that the 'narrowing' of paths using emotional markers may be critical to make the explored state space manageable in a short period of time, while being **socially intelligent**. In Figure 2.2, the five **features** are shown in task learning and execution, which will be a part of the main data extraction when we turn to analyzing the computational models of emotions in decision-making in section 2.5. Note that a learning and execution of a task can happen at the same time or in different phases, and that the word *Agent* in this study is referred to the affective decision-maker. A description of each of these features is given below.

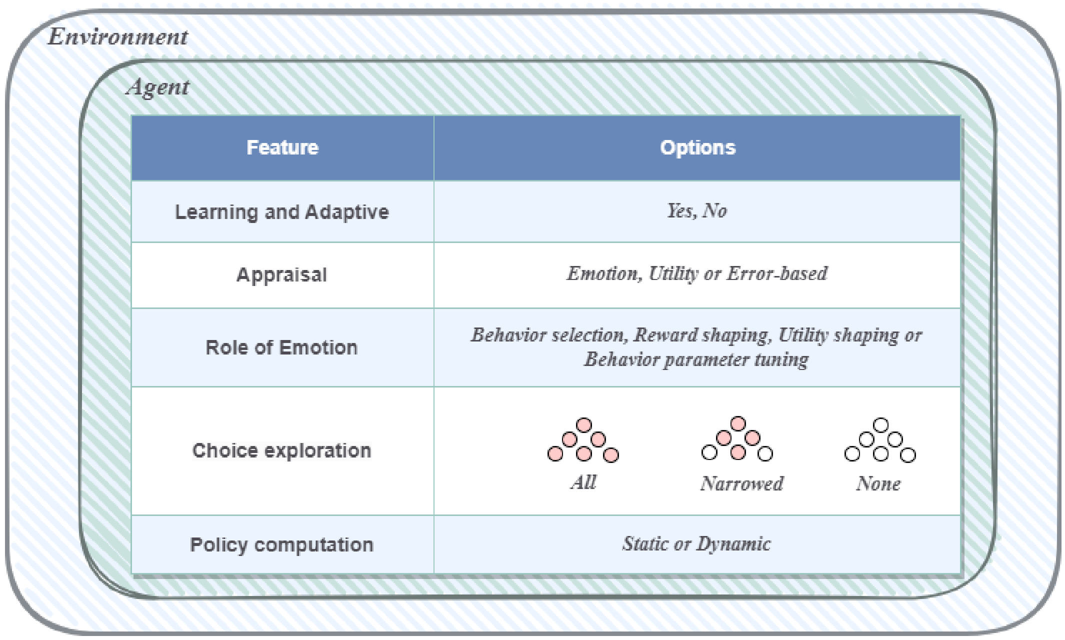


Figure 2.2: Five features of a decision-making affective agent. Each feature has a number of options, giving in total 144 possible models which we will reduce to 4 categories in section 2.5.

2.2.1 Learning and Adaptive

This indicates if the agent can learn and adapt to its environment. The options for this feature can be *'Yes'* or *'No'*. *'Yes'* would mean the agent can update or switch to a different internal model of the environment in case of changes. *'No'* would mean its internal model of the environment remains fixed.

2.2.2 Appraisal

This indicates assessment type of the situation or context by the agent. If the agent only assesses emotions (of itself or another agent/human), the appraisal is *'Emotion-based'*. If the agent assesses the value of choices or whether a goal is reached or not, then it is *'Utility-based'*. If the agent assesses the gap between its own prediction and the actual situation, then the appraisal is *'Error-based'*.

2.2.3 Role of emotion

This indicates whether the role of emotion is directly influencing *'Behavior selection'*, or indirectly influencing behavior selection via *'Reward shaping'*, *'Utility shaping'* or *'Behavior parameter tuning'*. Lisetti and Gmytrasiewicz [2002]'s work highlights seven different ways that emotions modify decision-making. Moerland *et al.* [2018] describes five ways in which emotions impact decision-making. Lisetti and Gmytrasiewicz [2002]'s list (prefixed by letter 'L') is given below:

(L1) emotions change perceptual processes;

(L2) emotions change action space (a) narrows action sets; (b) changes state probabilities to make states more/less likely predictions;

(L3) emotions change utility functions or desires;

(L4) emotions can define new goals, essentially an exploration bonus in RL

(L5) emotions can directly specify actions

(L6) social contagion

(L7) social rationality - modification of reward to include group goals

Moerland *et al.* [2018]'s list (prefixed by letter 'M') is given below. Here we also show the relationships between the taxonomies of Lisetti and Gmytrasiewicz [2002], Moerland *et al.* [2018] and our quadruple of emotion roles:

(M1) modification of the reward function and reward shaping; corresponds to Lisetti & Gmytrasiewicz's (L3) and to our *'Reward shaping'* or *'Utility shaping'*

(M2) modification of the state, e.g., augment the state with emotion and plan over it; (L1), (L2b)

(M3) meta-learning in which affect changes model hyperparameters like learning rate; (L4) and *'Behavior parameter tuning'*

(M4) action selection (L5) and *'Behavior selection'*, or the exploration/exploitation trade-off; (L4) and *'Behavior parameter tuning'*

(M5) epiphenomenal - sending emotional signals for the sake of sending emotional signals. (L6)

2.2.4 Choice Exploration

This provides the choices explored by the agent in a temporal depth in the future. It can be *'All'*, *'Narrowed'* or *'None'*. In a pure decision-theoretic framing of any situation, *'All'* is the only option. Exploring all choices to find the best option may be time-consuming. Some algorithms may, nevertheless, prune certain parts of the state space based on some heuristics and threshold settings, but these are non-emotional, such as estimates of expected utility. Exploring narrowed set of choices may make decision-making process faster. No choice exploration may mean that the agent has been directed to choose a particular option in a given situation.

2.2.5 Policy Computation

This focuses on policy computation being *'Static'* or *'Dynamic'*. *'Static'* means policies are pre-determined hand-coded strategies for what decision to take in a particular situation, whereas *'Dynamic'* means policies are computed based on the current situation. Policy computation may involve variable factors, resulting in a policy based on the situation. This can be more generalized as opposed to static policies that have to be defined for each situation an agent may encounter.

2.3 Survey Methodology

A systematic literature review methodology was chosen for this survey in order to find and examine computational models of emotion in decision-making (or behavior as purposeful action) in AI systems discussed in the published literature. This study followed Kitchenham's guidelines [Kitchenham, 2004] for conducting the review, except for in the study quality assessment stage as we were interested in all models regardless of their efficacy. Guidance on design was also found in the PRISMA 2020 [Page *et al.*, 2021] explanation and elaboration document.

2.3.1 Identification of Research

Information Sources

Four databases were searched on October 20th, 2021, to identify relevant papers. The same searches were rerun on June 12th, 2023 to identify relevant papers that were newly published or recently added to a database. The IEEE Xplore Digital Library and the Association for Computing Machinery Full-Text Collection were searched to find studies on affective computing from the computer science and software engineering disciplines. The PsycINFO database was chosen to identify relevant studies in psychological and neuroscience research, and it was searched via the American Psychological Association PsycNet interface. Scopus, an interdisciplinary database covering these areas and more, was also searched. Due to time constraints, we were not able to employ supplemental search techniques such as citation searching or handsearching. We also chose to focus on the scholarly published literature due to the broad nature of the topic, but the lack of a grey literature search could result in missing some models due to publication bias.

Search Strategy

The search process began by identifying the main concepts within the research topic: emotion, behavior, and artificial intelligent agent. Preliminary searches were conducted to identify related reviews as well as appropriate keywords. The search was constructed in the Scopus databases and possible keywords were tested using the AND NOT operator to determine what would be missed if they were not included. Nesting was used for the AI agent concept, for which there are some phrases that adequately describe both aspects of this concept but also some phrases and keywords which only describe the AI aspect or the agent aspect (ex. computational model) so they had to be ANDed together. See Appendix A for the complete database searches.

The search results were also checked against previously identified relevant papers to ensure those would be retrieved. The final Scopus search was adapted for the other databases' search interfaces and features. Phrasing and truncation symbols were used as needed and searches were limited to English when databases allowed.

2.3.2 Selection of Studies

We list here the inclusion and exclusion criteria, which were used for screening the papers.

Inclusion criteria:

- studies that describe computational models which make decisions based on an emotional input, and act accordingly;
- the model must be described to a degree where researchers can understand how it works;
- it must be implemented in a system that autonomously makes decisions;
- emotion, affect, sentiment or feeling must be involved in decision-making process the AI agent has;
- the AI agent can be a virtual agent or human, a bot, a computer simulation, or a robot or any form in which AI agent is implemented;
- written in English; and
- only peer reviewed research papers.

Exclusion criteria:

- studies that describe using an AI agent's emotions to influence human decision-making;
- models that use explicit rule based decision-making and behavior;
- studies that present a dataset (any modality) of emotion expressions and a classifier only;
- work that simply expresses emotion without any reason for it or any understanding of its effects;

- work that computes some form of emotional representation, but uses it only to verify a model of emotion;
- modelling only human's behavior guided by human's emotion;
- uses traits, moods and/or personality considerations in decision-making; and
- surveys, reviews, thesis or position papers (see section 2.1.4 for a review of related surveys).

Study Selection Process

In October 2021, 5360 results were exported into RefWorks and duplicates were removed. 4710 unique records were exported into Covidence to facilitate screening and data extraction. In June 2023, 6414 results were exported into Covidence and duplicates were removed. This added 1019 new unique results to review. Hence out of 11774 total exported search results, 5729 were screened at the title/abstract stage. See PRISMA flow diagram in Figure 2.3 for more details on our search and screening results. Two reviewers (AM and GC) independently screened the titles and abstracts of all the unique records added into Covidence. Any disagreements about relevance were resolved through consensus. Overall, 843 papers that met the selection criteria were moved into the full text screening stage, which was completed by one reviewer (AM). Based on the selection criteria, 714 papers were excluded at the full-text stage (e.g. 172 for being rule-based, 101 for lacking a description of a working model). Finally, 129 papers were included for final data extraction. Some studies outlined the same model, giving us a distinct set of 102 models. Figure 2.4 shows the distribution by publication year. There seemed to be an increase in papers on the topic of this survey intermittently till June 2023.

To define our scope of survey with good focus, we had stricter criteria regarding implementation in system or simulations. Papers like [Gratch and Marsella, 2004] and [Ahn and Picard, 2005] did not get included due to implementation being briefly discussed. Also, in search queries, we considered papers where the search terms 'behavior' and 'emotion' were within 5 words of each other, to limit the huge results. Papers like [Hudlicka, 2004] and [Malhotra *et al.*, 2020] did not appear in search results. The former was due to the word distance criteria and the latter was part of proceedings which was not indexed in the chosen databases.

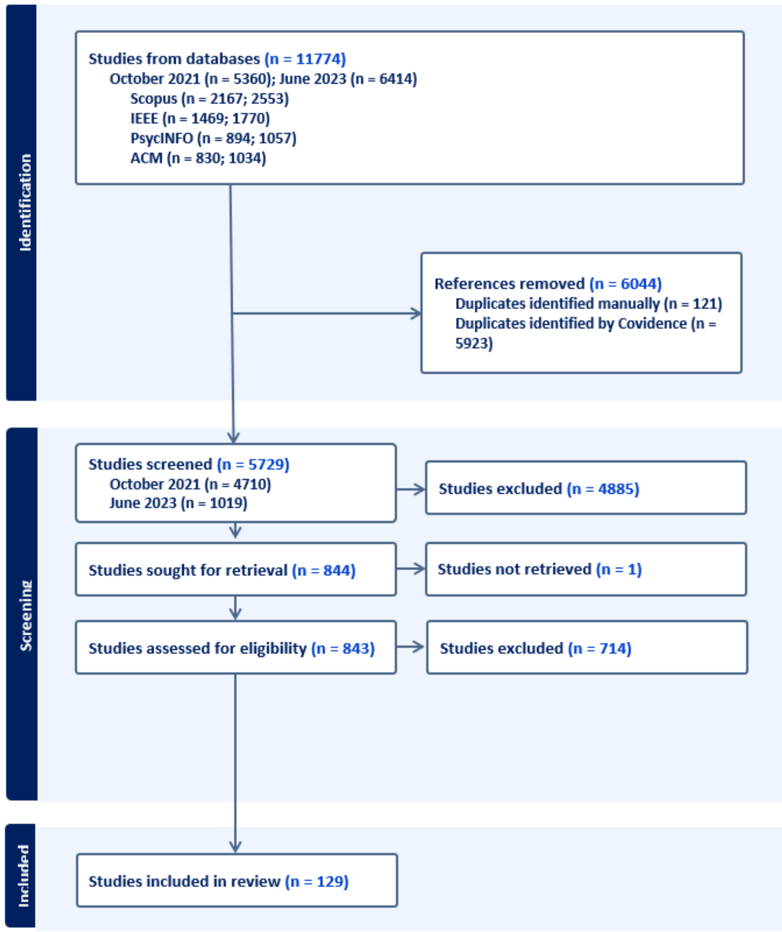


Figure 2.3: PRISMA flow diagram for our systematic review.

2.4 Data Extraction

We studied the included **129** papers that satisfied the selection criteria, and extracted the relevant emotional and decision-making features, viz. **Learning and Adaptive**, **Appraisal**, **Role of emotion**, **Choice exploration** and **Policy computation**, that were highlighted in section 2.2.

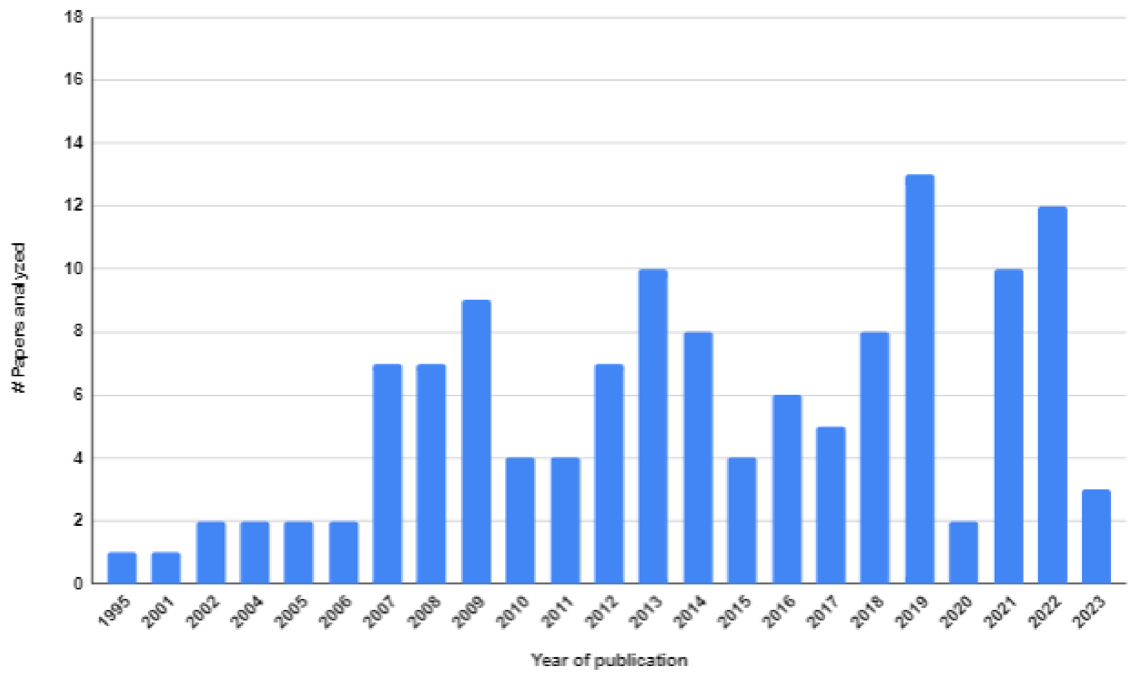


Figure 2.4: Distribution of papers by publication year

2.4.1 Design and Contents of Data Extraction Chart

Table 2.1 Part 1 and Table 2.1 Part 2 summarizes the emotional and decision-making features reviewed in section 2.2 for the **129** selected papers. It shows the values for the five features for all papers, grouped into four high-level categories marked with distinct background color. These groupings will be discussed in more detail in section 2.5. Some supplementary feature extraction is provided in section 2.4.3. Tables 2.2 - 2.8 show the supplemental information of each paper for the four model types.

2.4.2 Data Extraction Procedures

The data extraction was done manually by AM and reviewed by JH. In case of multiple publications on the same model, they were joined together in the supplementary extraction tables, resulting in 102 models in total, even though each of them were cited separately. An iterative approach was followed to get feedback from all authors as extraction progressed. Google shared drive was used to collaborate and place artifacts to review.

2.4.3 Supplementary information

Information (beyond the five main features) of seven additional features for each of the 102 models were extracted in tabular format for further reference. Note that these may be specific to the application used in the study. This can be found in tables 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8. Additional features considered are:

1. *Whose emotions?*
This indicates whose emotions the agent considers in the model. It can be *'Agent'* or *'Human'* or *'Both'*.

2. *Emotions*
This lists all the emotions that the agent considers for decision-making.

3. *Decision or behavior*
This provides the decision or an action that the agent takes in the model.

4. *Summary of emotion-decision link*
This provides application specific summary of emotional influence on decision.

5. *Computational approach*
This provides the computational approach taken for building the model.

6. *Code*
This gives details of the code specified for implementation. If not specified, *'NA'* is indicated.

7. *System/Simulation*
This provides if the implementation is in the form of a real world system or a simulation of a real world or some imagined scenario. In case there are multiple agents involved, it is specified accordingly.

2.5 Data Synthesis

Table 2.1 Part 1 and Table 2.1 Part 2 provide data extraction for emotions and decision-making features of 129 papers (102 models), based on the features listed in section 2.2. In

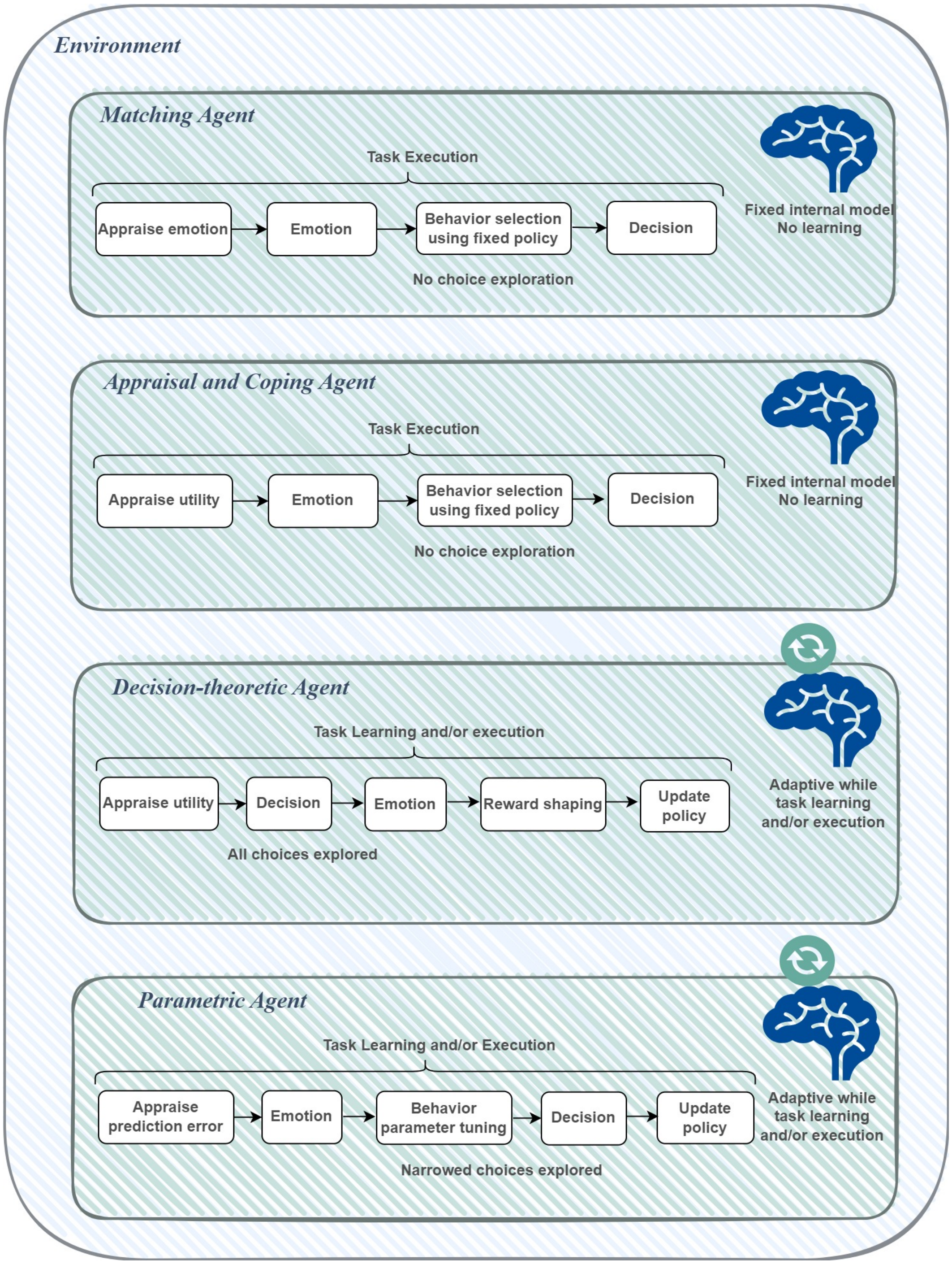


Figure 2.5: High level process view of agents of four model types (**Matching**, **Appraisal and Coping**, **Decision-theoretic** and **Parametric**)

the data synthesis, we categorized high level model types that surfaced based on distinguishing feature values. Accordingly, we labelled these high level model types as ***Matching***, ***Appraisal and Coping***, ***Decision-theoretic*** and ***Parametric***. A simplistic process view of agents of each model type in an environment is given in Figure 2.5. In the following, we provide an understanding of agents categorized under each model type and illustrate with a few sample papers. The words 'model' and 'agent' are used interchangeably.

Matching model appraises the emotion of itself or the human in a specific context, and it matches the behavior to use for that particular emotion, which is the decision. This model is relatively simplistic. Such agents have no learning feature and any internal model is fixed and not adaptive. Emotion-based appraisal is used as a guide to a static pre-defined policy. No choices are explored. In our earlier hypothetical scenario, if Mr. Bean is a matching agent, he may be pre-programmed such that if he is in a happy state, he should take route A, or if he is sad, he should take route B. He will not explore or evaluate any routes at any decision intersection. For example, a DJ agent in a bar setting [Rincon *et al.*, 2017], appraises the social emotional state of the audience, and based on that plays a song of a genre that will make them happy. If they are already happy, then the DJ will continue to play music of same genre. Each song is tagged with an emotion. Similarly in [Callejas *et al.*, 2014], a virtual recruiter system is designed to challenge or comfort a user based on anxiety level. Anxiety is appraised by an external source and dialog is generated based on a static policy. In [Sharma *et al.*, 2023], based on user emotion, a movie is recommended by the system. In [Denecke *et al.*, 2020], depending on the user's emotion, a mobile application with an integrated chatbot suggests activities or mindfulness exercises.

Appraisal and Coping models perform goal related utility appraisal, such that the situation and progress towards a goal is assessed, resulting in an emotion. Further corrective action is taken based on the appraised emotion. Agents of such type use a static pre-defined policy and no choices are explored. If Mr. Bean is an appraisal and coping agent, he will appraise the distance from the goal as he takes route A. If he is far away and getting late, he may feel frustrated, and take route B. He will not explore or evaluate any routes at any decision intersection. Such agents have no learning feature, and the internal model is fixed and non-adaptive. For example, in [Becker *et al.*, 2007], a virtual human 'Max' is simulated, wherein a situation-based coping behavior is used. A vacuum cleaner robot in [Gluz and Jaques, 2014b,a] appraises the situation and changes its affective reactions. In a robot navigation simulation [Zhang *et al.*, 2006], if the agent encounters a wall, or an obstacle or a goal, it leads to fear, frustration or happiness respectively. Based on the emotion, the agent avoids obstacles or goes to the goal. In a buyer and seller agent negotiation simulation [Qie *et al.*, 2022], the emotional value of the offer affects the counter-offer.

Decision-theoretic models perform utility-based appraisal, such that they appraise the rewards during task learning, and label emotion based on that. Policy computation is dynamic, such that actions are corrected based on the emotions, learning takes place and the internal model is adaptive. All the choices are explored to evaluate utilities. If Mr. Bean is a **Decision-theoretic** agent, he may learn by simulating all the routes to the supermarket. At each route choice, he may evaluate the distance to the goal and the rewards or cost that he accrues and what he feels based on that. Accordingly, he will choose the best route to follow. This may take a long time to compute the policy. But once learned, he may simply follow the policy. In [Homolová *et al.*, 2018], a game simulation is used, where an agent plays an ultimatum game with a human, and based on expected emotion as reward, accepts or reject offers. In [Mashayekhi *et al.*, 2022], a multi-agent normative society simulation is performed, wherein agents take prosocial actions based on guilt. In a social dilemma game simulation [Yu *et al.*, 2013, 2015], based on emotion as intrinsic reward, action is selected and utility values are updated. In an agent maze navigation simulation [Chao *et al.*, 2016], navigation action is changed based on emotional state as reward.

Parametric model types use error-based appraisal, such that they confirm predictions made about the environment, use emotions as a guide to narrow down the choice exploration, and take action according to a decision-theoretic policy only defined over those narrowed choices. Emotion influences behavior parameter tuning, which eventually leads to decision or behavior. Emotions are viewed primarily as models of noise/uncertainty. For example, negative emotional appraisals are generated by failed predictions (surprise), and lead generally to broader (more modeled uncertainty) posterior distributions, and thus to more exploration in reinforcement learning agents. Policy computation is dynamic and the internal model is adaptive. Learning happens during execution, even though the agent may start with an initial learned model. If Mr. Bean is a **Parametric** agent, he may start on route A based on his prior experience, but if it is now closed for construction, which causes a prediction error, and an associated emotional signal (of increased uncertainty) may guide him towards taking route D, which he feels more confident (good) about as next best option. He will update his internal model of his environment and his feelings as he executes the task. For example, in [Hesp *et al.*, 2021], based on reward prediction error, a policy is changed dynamically. In [Hoey *et al.*, 2016], an agent in social interaction changes self and/or other's identity and/or behavior based on the prediction error (called incoherence or deflection) of the interaction. In [Broekens *et al.*, 2007], reward-based exploitation or exploration behavior is used. Positive affect influences exploitation, whereas negative affect influences exploration. In a multi-agent simulation of choosing shopping locations [Han *et al.*, 2008], the location is selected based on the emotional value of utilities and error.

Figure 2.6 shows the distribution in percentages of models by model type. Out of 102 models, 28 were ***Matching*** type, 29 were ***Appraisal and Coping*** type, 32 were ***Decision-theoretic*** type and 13 were ***Parametric*** type.

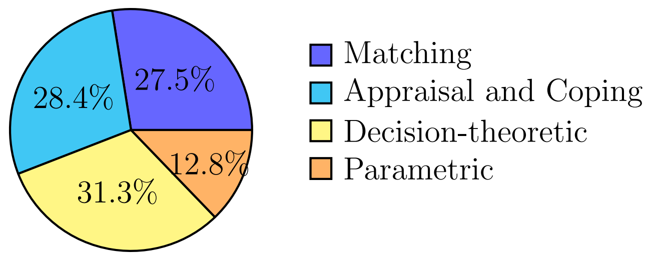


Figure 2.6: Distribution of models by model type

2.6 Discussion

To survey and analyze studies and systems involving emotions in decision-making, we selected five features for extraction based on our understanding of human decision-making, which is quite subjective and can be based on variety of factors (some of which the decision-maker may also not be aware of). This makes the search for unique characteristics challenging as well as interesting. Compared to many machine learning tasks, there is no standard benchmark for comparison of emotional models, making this analysis somewhat qualitative. In data synthesis, we categorized all surveyed papers into four high level model types, based on the features. Some models did have feature values which seem to be falling in multiple categories, but from a holistic perspective they were put into the majority model type. For example, the model in [Pereira *et al.*, 2017], uses a guilt payoff matrix with transition probabilities for decision to cooperate or defect in an iterated prisoners dilemma game simulation, which may mean it is 'Decision Theoretic', but because of the preset nature of its emotional appraisal, it seems more appropriate in the 'Matching' category.

Models of human emotions and decision-making can range from very simplistic to very complicated. AI systems are generally application and task focused. Depending on the goal of the system, the complexity of environment and the expectation from the decision-making agent, one can utilize one of the four model types, or have a hybrid model. The matching model type simply matches user's emotion (regardless of context) to behavior selection, but recognizing user's emotion itself is a challenging task and is context dependent. If one uses matching model type for a real-world application, there may be a need for a different emotion recognition module for each user context. The appraisal model type infers user's emotion based on limited set of situational utility contexts, and selects a behavior based on the emotion. The decision-theoretic model type bases emotions on rewards and generalizes across situations. The parametric model type bases emotions on prediction error, which generalizes to even more situations (including those where utility or reward is not well defined). In summary, the four model types, viz. Matching, Appraisal and Coping, Decision-theoretic and Parametric, range from simple to complex mechanisms, and specific to more generalized contexts.

This systematic review attempts to cover all the literature, but some sources may have been missed. These include grey literature (sources such as new articles, publications which do not have appropriate search registry). These were not searched due to time constraints.

2.7 Conclusion

In this systematic review, we analyzed 129 papers (resulting in 102 models) for computational models of emotions in decision-making and behavior, that are implemented as a system or a simulation. It is a challenging task to compare such models qualitatively, and to surface their characteristics. We used five key features (*learning and adaptive*, *appraisal*, *role of emotion*, *choice exploration*, and *policy computation*) for comparing 129 papers and clustering them into four high-level model types, viz., **Matching**, **Appraisal and Coping**, **Decision-theoretic** and **Parametric**. While much work has gone into matching and appraisal and coping mechanisms, these have deficits in terms of a lack of adaptivity, a fixed action selection mechanism and a disconnect from the situation (e.g. emotions are connected to behaviors through *ad hoc.* rules). Decision-theoretic models alleviate many of these issues, but are left using emotions as simple reward-shaping mechanisms, which does not alleviate the decision-theoretic requirement of exhaustive search through action space. Further, decision theoretic models for social situations are often based on shared reward functions, which would expand the search through action space even further. Finally, parametric models connect emotions with parameters and uncertainty, and thus with *model selection*, thereby alleviating this last problem and providing a clear method for narrowing computation over action spaces (by using models that only have such narrowed spaces). Such a mechanism for emotion's interaction with decision-making is well aligned with recent findings in neuroscience [Hesp *et al.*, 2021]. This survey may help researchers to frame literature on the topic, utilize models based on their application, and push the boundaries of emotion, cognition, decision-making and affective computing research in AI.

Table 2.1: Part 1: Data extraction of five features of emotions in decision-making agent in papers analyzed. Papers related to same model are grouped by round brackets. The major row sections correspond model types, from top to bottom: **Matching**, **Appraisal and Coping**. Colors are model types, and exceptions are shown in cross-hatched grey color (otherwise the value is that at the top of the column in the same color).

Studies	**Learning and adaptive**	**Appraisal**	**Role of emotion**	**Choice exploration**	**Policy computation**
Lesani *et al.* [2004], Callejas *et al.* [2014], Pereira *et al.* [2017], Alhijawi [2019], (Mochida *et al.* [1995]; Tsankova [2009]), (Mei [2016]; Ying and Zhentao [2016]), Novianto and Williams [2011], Taurah *et al.* [2019], Álvarez *et al.* [2019], Li *et al.* [2021], Siqueira *et al.* [2018], Kao *et al.* [2019], Papangelis *et al.* [2014], Rubilar *et al.* [2014], Lee and Kim [2018], Hoogendoorn *et al.* [2010], Kaklauskas *et al.* [2018], Mizgajski and Morzy [2019], Karabelnikova and Samsonovich [2021], Sarrafzadeh *et al.* [2008], Fischl *et al.* [2019], Polignano *et al.* [2021], Rincon *et al.* [2017], Sharma *et al.* [2023], Denecke *et al.* [2020]	No	Emotion	Behavior selection	None	Static
Pimentel and Cravo [2009]				Narrowed	
Johansson and Dell'Acqua [2012]		Utility			
Chen *et al.* [2022a]	Yes				
Liu and Ando [2008], (Esau *et al.* [2006]; N. Esau, L. Kleinjohann, and B. Kleinjohann [2007]), Antos and Pfeffer [2011], Lim *et al.* [2012], (Sharpanskykh and Treur [2013],Sharpanskykh and Zia [2014], A. Sharpanskykh and K. Zia [2012], Sharpanskykh and Treur [2010], A. Sharpanskykh and J. Treur [2013]), Becker *et al.* [2007], (Pontier and Siddiqui [2009]; Hoorn *et al.* [2008]; J. F. Hoorn, M. Pontier, and G. F. Siddiqui [2012]; Pontier *et al.* [2012]), (Gluz and Jaques [2014b,a]), Belkaid *et al.* [2017], (M. Belkaid, N. Cuperlier, and P. Gaussier [2015]; M. Belkaid, N. Cuperlier and P. Gaussier [2018]), (Faghihi *et al.* [2009, 2008, 2011]), Zhang *et al.* [2006], De Carolis *et al.* [2017], Belhaj *et al.* [2016], Donaldson *et al.* [2004], Kiryazov *et al.* [2013], Xie *et al.* [2019], Yongsatianchot and Marsella [2022], Prasad and Thomas [2022], Zhai *et al.* [2022], Maroto-Gómez *et al.* [2023]	No	Utility	Behavior selection	None	Static
Daglarli *et al.* [2009]	Yes				
Punithan and Zhang [2018]			Reward shaping		
Kim *et al.* [2004]	Yes				Dynamic
Jauffret *et al.* [2013]	Yes	Emotion, Error			
Cabrera-Paniagua *et al.* [2023]		Emotion	Utility shaping	All	
Qie *et al.* [2022]	Yes		Utility shaping		
Tzeng *et al.* [2021]	Yes			All	Dynamic
Lu *et al.* [2021]	Yes	Utility, Emotion			

Table 2.1: (contd.) Part 2: Data extraction of five features of emotions in decision-making agent in papers analyzed. Papers related to same model are grouped by round brackets. The major row sections correspond model types, from top to bottom: **Decision-theoretic** and **Parametric**. Colors are model types, and exceptions are shown in cross-hatched grey color (otherwise the value is that at the top of the column in the same color).

Studies	Learning and adaptive	Appraisal	Role of emotion	Choice exploration	Policy computation
Biancardi *et al.* [2019], (Liu *et al.* [2008]; C. Liu, K. Conn, N. Sarkar, and W. Stone [2007]; Conn *et al.* [2008]), Homolová *et al.* [2018], (Tripathi *et al.* [2019]; A. Tripathi, T. Ashwin, and R. M. R. Guddeti [2018]), Rach *et al.* [2021],Huang *et al.* [2020], Chen *et al.* [2022b], Moussa and Magnenat-Thalmann [2013], Chao *et al.* [2016], Schermerhorn and Scheutz [2009], Gadanho and Hallam [2001], Salichs and Malfaz [2012], Castro-González *et al.* [2013], Gomes *et al.* [2019], Feldmaier and Diepold [2014], Jiang and Wang [2019], Lu *et al.* [2016], Nasir [2018], Ficocelli *et al.* [2015], Zhou and Coggins [2002], (Yu *et al.* [2015, 2013]), Bagheri *et al.* [2021], Barthet *et al.* [2021, 2022], Wang *et al.* [2022], Mashayekhi *et al.* [2022], Sequeira *et al.* [2014]	Yes	Utility	Reward shaping	All	Dynamic
(Zhang and Liu [2009]; H. Zhang and S. Liu [2009])		Utility, Emotion			
(Hoefinghoff *et al.* [2013]; Hoefinghoff and Pauli [2013])		Emotion	Reward shaping, Behavior selection	Narrowed	
Chen and Wang [2019]			Reward shaping, Behavior selection		
(Matsuda *et al.* [2011]; Horio and Matsuda [2010]), Kuremoto *et al.* [2013]			Utility shaping		
Tsankova [2002]			Behavior selection		
Duell and Treur [2012], (Hoey *et al.* [2016]; Asghar and Hoey [2015]; Hoey *et al.* [2018]), Hesp *et al.* [2021]	Yes	Error	Behavior parameter tuning	Narrowed	Dynamic
Han *et al.* [2008], Broekens *et al.* [2007], Huang *et al.* [2018b], Wu *et al.* [2022], Jiang *et al.* [2022], Zhang and Zeng [2023]		Error, Utility			
(Morgado and Gaspar [2005]; L. Morgado and G. Gaspar [2005]; L. Morgado and, G. Gaspar [2007])		Error, Emotion			
Johansson and Dell'Acqua [2009]		Error, Emotion		All	
Hogewoning *et al.* [2007]		Utility			
(Lee-Johnson and Carnegie [2009]; C. P. Lee-Johnson and D. A. Carnegie [2007])		Utility	Behavior parameter tuning, Reward shaping	All	

Table 2.2: Summary of model implementation characteristics for ***Matching*** model type

#	Paper	Whose emotion?	Emotions	Decision or behavior	Summary of emotion-decision link	Computational Approach	Code	System and/or Simulation
1	Lesani *et al.* [2004]	Agent	nervousness	refuse or agree to task allocation	based on nervousness, agree or reject request from faulty agent	Formula based	JADE (Java)	Multi-agent simulation of fault situations
2	Callejas *et al.* [2014]	Human	anxiety	provide comfort or challenge to human	based on human's anxiety, adapt agent's dialog	Strategy rules	SAIBA	Virtual recruiter system
3	Pereira *et al.* [2017]	Agent	Guilt	cooperate or defect	based on the guilt, cooperate or defect	Evolutionary Game Theory (EGT)	NA	Iterated Prisoner's Dilemma simulation
4	Alhijawi [2019]	Human	like, love, haha, wow, wad and angry	recommend digital products	based on user satisfaction and emotion, recommend item	modified collaborative filtering, similarity of users	NA	Item recommendation simulation
5	Mochida *et al.* [1995], Tsankova [2009]	Agent	frustration state(indicating pleasantness and unpleasantness)	reach goal and avoid obstacles	frustration is used to avoid obstacles and reach goal	Braitenberg's architecture	NA	Autonomous mobile robot simulation
6	Mei [2016]; Ying and Zhentao [2016]	Human	happy, angry	grasp or avoid the object	when emotion is happy, then grasp, else avoid	Emotion-Driven Attention(EDA) model	Matlab	Robotic manipulator simulation
7	Novianto and Williams [2011]	Agent	valence, arousal	play red ball or drums or sleep dilemma	based on emotional preference, select action	probabilistic causal network, appraisal	NA	ASMO Robot bear simulation
8	Taurah *et al.* [2019]	Human	anger, contempt, disgust, fear, happiness, neutral, sadness, surprise	selecting time table for the student	based on the student's emotion, plan a timetable	NN, Microsoft Emotion API	Java Spring-Boot	MOOC system
9	Álvarez *et al.* [2019]	Human	tense, excited, cheerful, relaxed, calm, bored, sad, irritated, neutral	selection of song to play	based on an emotion to be induced to the user, select a song to play	Annoy algorithm (Approx Nearest Neighbor Search), Collaborative filtering	Java Spring webservice	DJ music recommender system
10	Li *et al.* [2021]	Human	58 emojis and context	dialogue responses	based on current context and expected emotion, generate dialogue	Transformer	NA	chatbot system
11	Siqueira *et al.* [2018]	Human	2D valence and arousal	decide when to extend conversations based on incoherence	emotional incoherence - driven dialogue	Appraisal, emotional coherence matching	NICO robot	DJ Robot interactive system with human
12	Kao *et al.* [2019]	Human	neutral, joy, sadness, fear, anger, surprise, disgust, non-neutral	dialogue generation	user emotion used for dialogue generation	generative model SeqGAN	NA	experimental simulation for chatbot
13	Papangelis *et al.* [2014]	Human	happiness,caring, depression, inadequateness, fear, confusion, hurt, anger, loneliness, and remorse	dialogue with the user	based on the emotional score, generate dialogue	Action Weights Learning (AWL), pattern matching	NA	Adaptive Dialogue System (ADS) PTSD assessment system
14	Rubilar *et al.* [2014]	Agent	threat	navigation	threat causes behavorist reaction	FSM	MODI (two-wheeled robot)	robot navigation system
15	Lee and Kim [2018]	Human	emotion intensity vector	recall emotional episode and interact with user (alerting, encouraging, showing empathy)	based on past memory of anticipated emotions, robot reacts	modified Deep ART network	Mybot robot	Robotic system interacting with human in kitchen

Table 2.3: Summary of model implementation characteristics for ***Matching*** model type (contd.)

#	Paper	Whose emotion?	Emotions	Decision or behavior	Summary of emotion-decision link	Computational Approach	Code	System and/or Simulation
16	Hoogendoorn *et al.* [2010]	Agent	positive, negative	option selection	emotion and intention influences that of the other person	Differential equations, somatic marker hypothesis, neurological mirroring	Matlab	simulations of leaders and followers in group decision
17	Kaklauskas *et al.* [2018]	Human	anger, surprise, happiness, disgust, sadness, fear, neutrality, valence, arousal	select video ad	based on viewers emotional expressions and physiological parameters, show ad video	Multiple Resource Theory	Microsoft Emotion API	video recommendation system for real estate buying
18	Mizgajski and Morzy [2019]	Human	PAD	news content recommendation	expected reaction based recommendation for similar items for user	affective user based collaborative filtering with targeting by expected reaction (AUBCF-WTBER)	Python	content recommendation system
19	Karabelnikova and Samsonovich [2021]	Agent	valence, arousal, dominance	selection of a partner, changing facial expression, dance patterns	emotion based moral schema and appraisal based feeling leads to bias in action selection	eBICA, appraisal, somatic marker	Python, Unreal engine	simulations of virtual pet and 3 virtual dance partner bots, and system experiment with 1 human and 2 bots
20	Sarrafzadeh *et al.* [2008]	Human	6 basic emotions by Ekman	tutoring actions	past action-emotion pattern matched to do next action	Fuzzy approach, pattern matching	NA	Affective Tutoring System
21	Fischl *et al.* [2019]	Agent	happy, neutral, distressed	drive towards or away from human or be stationary	based on human facial expression, agent emotion influences behavior to drive towards or away from human or be stationary	Rule based, NEF	Nengo, Python	human-robot interaction system
22	Polignano *et al.* [2021]	Human	6D vector(joy, anger, sadness, surprise, fear, disgust)	song suggestion	compute list of suggested songs for user based on his affective preferences and affective state	formula based affective coherence score	IBM Bluemix	affective music recommender system
23	Rincon *et al.* [2017]	Human	PAD	song selection	DJ agent calculates the social emotion of audience and decides the next song to achieve goal of happiness	ANN	JaCalIVE framework	DJ agent-human agents simulation
24	Pimentel and Cravo [2009]	Agent	positive, negative	task specific actions	based on somatic marker, chose action	Somatic Marker hypothesis	NA	Simulation of The Iowa Gambling Task (IGT), The Restaurant World (RW) task
25	Johansson and Dell'Acqua [2012]	Agent	sadness, happiness, fear	path finding	based on emotion, plan path finding	formula based	NA	path maneuvering simulation for NPC
26	Chen *et al.* [2022a]	Human	VAD	response dialogue	based on emotion, chat	Transformer-based	Pytorch	empathetic response in customer service
27	Sharma *et al.* [2023]	Human	anger, fear, sad, disgust, happy, surprise	movie selection	based on user emotion, recommend movie	Similarity, SVD, collaborative filtering	Python	movie recommender system
28	Denecke *et al.* [2020]	Human	fear, anger, grief	dialog generation	depending on the user emotion, activities or mindfulness exercises are suggested	Similarity	OSCOVA	SERMO, a mobile application with integrated chatbot

Table 2.4: Summary of model implementation characteristics for ***Appraisal and Coping*** model type

#	Paper	Whose emotion?	Emotions	Decision or behavior	Summary of emotion-decision link	Computational Approach	Code	System and/or Simulation
1	Liu and Ando [2008]	Agent	hopeful, elated, surprised, anxious, disappointed, painful	expressions, voice and posture	based on emotional appraisal, and anticipation, change action	Emotion Appraisal Module (EAM) and Growing Network (G-Net) module	brainCAD	Semi-embodied system (qViki)
2	Esau *et al.* [2006]; N. Esau, L. Kleinjohann, and B. Kleinjohann [2007]	Both	happiness, sadness, anger, fear, neutral	head movements, facial expressions and natural speech output	based on emotional state of human, agent's emotions and drives influence its behavior	Formula based	NA	Robot head MEXI system
3	Antos and Pfeffer [2011]	Agent	hope, fear, boredom	update priority of goals	based on the emotion, update the priority of goals	emotional operators	NA	Simulations of restless bandits and forgaging domain
4	Lim *et al.* [2012]	Agent	hope, fear, arousal	actions related to energy and well-being	based on emotions, take adaptive action	FAtiMA-PSI model	ORIENT (Java)	Simulation of NPC
5	Sharpanskykh and Treur [2013]; Sharpanskykh and Zia [2014]; A. Sharpanskykh and K. Zia [2012]; Sharpanskykh and Treur [2010]; A. Sharpanskykh and J. Treur [2013]	Agent	hope, fear	exit action of group of agents	based on neighborhood's preparation to emotional responses for each option, exit route is determined	Hebbian learning principle, formula based	LEADSTO, Netlogo	Simulation - multi-agent evacuation scenario
6	Becker *et al.* [2007]	Agent	PAD	reactive, deliberative situation-focused coping behavior	based on appraised emotion, coping behaviors are considered	Appraisal, Coping, BDI	NA	Simulation of virtual human Max
7	Pontier and Siddiqui [2009]; Hoorn *et al.* [2008]; J. F. Hoorn, M. Pontier, and G. F. Siddiqui [2012]; Pontier *et al.* [2012]	Agent	valence	hug, attack, change, avoid, Pontier *et al.* [2012] - moral decisions	based on highest expected emotional satisfaction, action is chosen	Silicon Coppélia (EMA, CoMERG, I-PEFiCADM)	Haptek player, Javascript	affective virtual agent system playing tic-tac-toe, Pontier *et al.* [2012] has Trolley and Footbridge dilemma simulation
8	Gluz and Jaques [2014b,a]	Agent	OCC joy, fear	affective reactions	based on appraisal, chose action	BDI, Appraisal, BDN	AgentSpeak - agent programming language	vacuum cleaner robot system
9	Belkaid *et al.* [2017]	Agent	boredom, frustration	control a robotic system in a visual search task	emotional metacontrol intervenes to bias the robot visual attention during active object recognition	ANN, Appraisal	robotic platform	robotic system in a visual search task
10	M. Belkaid, N. Cuperlier, and P. Gaussier [2015]; M. Belkaid, N. Cuperlier and P. Gaussier [2018]	Agent	pleasure, pain, motivation	perception of surrounding space	emotional modulation of the peripersonal space	dynamic neural field (DNF)	Webots, Promethe NN simulator	simulate robot approach motivation and perception of reachability
11	Faghihi *et al.* [2009], Faghihi *et al.* [2008], Faghihi *et al.* [2011]	Agent	valence	selecting behavior from memory	based on the emotion, reactive or deliberative action is taken	Emotional codelets in CTS architecture	NA	Robot-astronaut simulation

Table 2.5: Summary of model implementation characteristics for ***Appraisal and Coping*** model type (contd.)

#	Paper	Whose emotion?	Emotions	Decision or behavior	Summary of emotion-decision link	Computational Approach	Code	System and/or Simulation
12	Zhang *et al.* [2006]	Agent	frustration/sorrow, fear/dread, joy/happiness	navigation	based on emotion, avoid obstacle or go to goal	formula based	NA	Robot navigation simulations
13	Daglarli *et al.* [2009]	Agent	distress, relief, aggressive, enjoy	wander, head on, avoid obstacle, goal	emotions determine sequences of behaviors	HMM, Q-SOM neural network	Matlab	autonomous mobile robot simulation
14	De Carolis *et al.* [2017]	Both	human: VA based (anger, sadness, happiness, neutral), agent: OCC based	console, encourage, congratulate, joke, calm down	based on human emotion, robot emotions are appraised and chose dialogue/expression	Dynamic Belief Networks	NAO robot	Assistive social robot system
15	Punithan and Zhang [2018]	Agent	happiness, sadness	maze navigation actions	emotion and reward threshold based action and emotion loop	reward based	NA	maze traversal simulation
16	Belhaj *et al.* [2016]	Agent	joy, distress, hope, fear, sorry for, undefined	Walk, random walk, rest, ask for help self and other, no action	based on emotion, coping behavior is chosen	Appraisal and Coping	NA	RoboCupRescue (RCR) simulation agent in emergency situations
17	Kim *et al.* [2004]	Agent	happiness, sadness, anger, fear, neutral	73 behaviors in 10 behavior sets	agent's emotion influences the learning rate	reward based learning	OpenGL	synthetic character in the 3D virtual world simulation
18	Donaldson *et al.* [2004]	Agent	hunger, thirst, fear - 3 values (desirability, potentiality, intensity)	moving, thinking, panicking in path planning	based on intensity and threshold of emotional state, behavior changes	emotional A*, Appraisal	Python	map and maze traversal simulation
19	Jauffret *et al.* [2013]	Agent	frustration	navigation	based on frustration, change strategy to call for help	formula based prediction error	simulation software	robot navigation simulation
20	Kiryazov *et al.* [2013]	Agent	Arousal	involuntary (homeostasis) and deliberative (motivation based work or charge)	arousal impulse affects the speed movement and facial expression	formula-based, WASABI EE	iCub robot	service robot system and simulation
21	Xie *et al.* [2019]	Agent	fear, desire intensity	crossing a highway or waiting	fear and desire encoded into the agents decision-making formula for crossing highway or waiting	regression tree	NA	autonomous vehicle traffic simulation
22	Cabrera-Paniagua *et al.* [2023]	Agent	anger, joy, fear	selecting travel route	emotion used in utility of route	formula based	NA	simulation of flexible passenger transportation
23	Yongsatianchot and Marsella [2022]	Agent	emotional stress	evacuating or staying home, altering ones belief and altering ones goals	based on emotional stress, change belief or goals	Coping in POMDP	NA	evacuation simulation
24	Prasad and Thomas [2022]	Agent	fear	response to an attack	based on fear, select action	formula based	SPARQL	cyber security scenario simulation
25	Qie *et al.* [2022]	Agent	utility value as emotion	negotiation	emotional value of the offer affects counter-offer	Weber-Fechner law, cloud similarity, formula based	NA	Buyer and seller agent negotiation simulation
26	Zhai *et al.* [2022]	Agent	fear, happiness, delight, intimacy	path selection	number of velocity space samples depends on current emotional level of the robot	Particle swam optimization, formula based	Gazebo simulator, Rviz visualization interface	simulation of robot path travel
27	Tzeng *et al.* [2021]	Both	valence	comply or violate a norm	based on valence, comply or violate norm	BDI, Appraisal	Python	simulation of agent at grocery store
28	Maroto-Gómez *et al.* [2023]	Both	six basic emotions	emotional responses	based on emotion, interact	formula based	NA	social robot Mini system interacting with human
29	Lu *et al.* [2021]	Human	neutrality, pleasure, angry, sad	sleep, eat, work, entertainment	based on user emotional preference, select options of service	similarity based matching	Matlab	simulation of service robot

Table 2.6: Summary of model implementation characteristics for ***Decision-theoretic*** model type

#	Paper	Whose emotion?	Emotions	Decision or behavior	Summary of emotion-decision link	Computational Approach	Code	System and/or Simulation
1	Biancardi *et al.* [2019]	Human	warmth	dialog and non-verbal behaviors	based on warmth of human, adapt the dialog and non-verbal behaviors	RL (Q-Learning)	SAIBA compliant Greta/ Virtual Interactive Behavior platform	Virtual guide system
2	Liu *et al.* [2008]; C. Liu, K. Conn, N. Sarkar, and W. Stone [2007]; Conn *et al.* [2008]	Human	anxiety, engagement, and liking	robot motion,speed,shots, background music	based on the affective states of a child with ASD adapt its behaviors	RL (QV-learning)	Simulink software	Robot Basketball system for children with Autism
3	Homolová *et al.* [2018]	Human	positive (least to most)	accept or reject an offer	Based on expected emotion as reward, choose action	MDP	Matlab	Simulation of Ultimatum Game
4	Zhang and Liu [2009]; H. Zhang and S. Liu [2009]	Agent	positive, negative	navigation	based on expected reward, navigate	RL (ACLDM model)	NA	Robot navigation simulations
5	Tripathi *et al.* [2019]; A. Tripathi, T. Ashwin, and R. M. R. Guddeti [2018]	Human	3D vector (joy, sadness, fear)	video recommendation	based on the affective intensity of the present video and user actions, next video is recommended	RL (SARSA, Q-learning) and Deep bidirectional RNN	NA	Video recommender system
6	Rach *et al.* [2021]	Agent	valence, arousal	claim, argue, concede	select most appropriate emotional tone in dialogue, based on emotion	RL ($\epsilon - greedy$)	EVA 2.0	multi-agent dialogue simulation
7	Tsankova [2002]	Agent	fear, frustration	navigation	based on emotion, navigate avoiding obstacle	RL (Q-learning)	Matlab	Autonomous mobile robot simulation
8	Huang *et al.* [2020]	Agent	valence, novelty and motivational relevance	navigation	based on emotional reward, take action	RL model-based (MB) and model-free (MF)	V-REP robot simulator	static and random goal reaching task with mobile robot
9	Chen and Wang [2019]	Agent	empathy and counter-empathy	hunt independently or in teams in SG, cooperation, defect in PD, accept reject in UG	random action at times or based on empathetic utility, take specific action	Adaptive Empathic Learning, Q-learning	NA	multi-agent game simulation (Survival Game SG, Prisoners' dilemma PD, Ultimatum Game UG)
10	Chen *et al.* [2022b]	Agent	empathy	cooperation, defect in PD and 3-agent dilemma, accept reject in UG	based on empathetic and affective utility, take action	RL(Q-learning), Gradient Ascent, prospect theory	NA	multi-agent game simulation (Prisoners' dilemma PD, Ultimatum Game UG, 3-agent dilemma)
11	Moussa and Magnenat-Thalmann [2013]	Both	2 well-being emotions, 8 blame and praise emotions, and 4 empathic emotions	game play speech action selection	strongest emotion influences action selection	RL (Q-learning), Appraisal	Java, humanoid robot	humanoid robot system playing tic-tac-toe with human
12	Chao *et al.* [2016]	Agent	4D vector (hope, joy, fear, sad) and intensity	navigation	based on emotional state and reward, change action	RL, Q learning, Appraisal	NA	agent maze navigation simulation
13	Schermerhorn and Scheutz [2009]	Agent	positive, negative	whether to ask someone for directions	based on affective utility, choose action	POMDP, pattern matching	NA	preliminary robot simulation
14	Hoefinghoff *et al.* [2013]; Hoefinghoff and Pauli [2013]	Agent	frustration	action selection	frustration level is used as a threshold for the selection of actions	formula based	NA	Iowa gambling task simulation
15	Matsuda *et al.* [2011]; Horio and Matsuda [2010]	Agent	fear	cooperative behaviors	based on emotional judgement, take action	RL and Emotion learning (DRE), Monte Carlo	NA	multi-agent simulation of grid world

Table 2.7: Summary of model implementation characteristics for ***Decision-theoretic*** model type (contd.)

#	Paper	Whose emotion?	Emotions	Decision or behavior	Summary of emotion-decision link	Computational Approach	Code	System and/or Simulation
16	Gadanho and Hallam [2001]	Agent	happiness, sadness, fear, anger	avoid obstacles, seek light, wall following	emotional intensity changes, influence behavior changes	RL (Q-learning)	Khepera robot simulator	robot survival task simulation
17	Kuremoto *et al.* [2013]	Agent	Pleasure, Arousal	select movement direction	emotion into action selection	RL (Q learning)	NA	simulation of 2-hunter-1-prey problem
18	Salichs and Malfaz [2012]	Agent	happiness, sadness, fear	actions for food, water/medicine, world, other agent	emotion for learning action selection	RL (Q-learning), Appraisal	NA	simulation of role playing game CoffeeMud
19	Castro-González *et al.* [2013]	Agent	fear	escaping, play, idle, go to player, nearoff, dance, listening, stop, plugged, remain, charged	learned fear behaviors	Appraisal, Q-learning	Maggie robot	social robot system interaction
20	Gomes *et al.* [2019]	Agent	happy, neutral, unhappy (on ordered pair excitation-discouragement, satisfaction-disappointment)	walk, flipping, run, look to left, look to right, pick up	based on emotional impulse, change behavior	RL, Asynchronous Advantage Actor-Critic (A3C)	NA	maze simulation
21	Feldmaier and Diepold [2014]	Agent	2D valence, arousal	shortest path to goal	core affect used to bias the reward function to guide the decisions of an agent	RL, Dyna Q-learning, Appraisal	NA	maze traversal simulation
22	Jiang and Wang [2019]	Agent	regret	choose between agent's own detection results or requesting human service	based on cost of choices, regret influences decision-making	regret decision model	NA	human-multi-robot path planning simulation
23	Lu *et al.* [2016]	Agent	novelty, happiness index, control	learning speed, exploration	intrinsic emotional motivation based on novelty, happiness index, control, influences exploration and learning	RL (model-based VI)	NA	rat in maze simulation
24	Nasir [2018]	Both	happy, sad, neutral, angry	pursue drive, change emotion, be idle	emotion state of human and robot considered in optimal policy for convincing human for task completion	MDP	NA	socially assistive robot simulation
25	Ficocelli *et al.* [2015]	Both	Agent (happy, neutral, sad, angry), Human (3D body pose based)	verbal (speech) and nonverbal (intonation, facial expressions, and gestures) actions	Human affect affects drives, which affects robot emotions, using which the verbal and non verbal actions are output	RL (Q-learning)	C++	socially assistive robot-human interaction system to create the participant's activity schedule for the day
26	Zhou and Coggins [2002]	Agent	emotion variables hand coded?	feeding, drinking, playing	based on emotional values, select behavior	RL (Q-learning)	Khepera robot	robot feeding and drinking task system
27	Yu *et al.* [2015, 2013]	Agent	positive (joy),negative (fear, sadness, anger)	cooperate, defect	based on emotion as intrinsic reward, select action, update strategy and utility values	Multiagent RL	NA	Social dilemma game simulation
28	Bagheri *et al.* [2021]	Agent	anger, sadness, surprise, happiness, fear, and disgust	empathic behavior	emotion as reward marker for action	RL (Q-learning)	NA	Human robot interaction system for game playing
29	Barthet *et al.* [2021, 2022]	Agent	arousal	select game state	based on arousal, select game state	RL, K-NN	NA	rally-driving game simulation
30	Wang *et al.* [2022]	Human	smile, not-smile	empathetic behavior	based on emotional feedback, change policy	RL (Q-learning)	NA	human-robot interaction simulation
31	Mashayekhi *et al.* [2022]	Agent	guilt, conflict	prosocial actions	based on guilt, take prosocial actions	RL	NA	multi-agent normative society simulation
32	Sequeira *et al.* [2014]	Agent	novelty, valence, goal relevance, control	forgaging actions	emotional appraisals incorporated as reward feature	RL, POMDP, Appraisal	NA	six forgaging simulations

Table 2.8: Summary of model implementation characteristics for ***Parametric*** model type

#	Paper	Whose emotion?	Emotions	Decision or behavior	Summary of emotion-decision link	Computational Approach	Code	System and/or Simulation
1	Duell and Treur [2012]	Agent	positive, negative	joint decision option selection	based on feelings and intentions, select joint decision option	Differential equations in temporal-causal network	LEADSTO	Multi-agent simulation
2	Hoey *et al.* [2016]; Asghar and Hoey [2015]; Hoey *et al.* [2018]	Both	3D EPA	Socially interactive behavior	based on the incoherence between fundamental sentiments and transient impressions, take adaptive action	POMDP	Python	Simulations
3	Hesp *et al.* [2021]	Agent	valence	action selection	affect influences agent's confidence in action selection	Deep active inference, MDP	Matlab	T-maze agent simulation
4	Johansson and Dell'Acqua [2009]	Agent	sadness, fatigue, anger, fear happiness	walk towards food, stand still, escape, explore, eat food, go to enemy, attack enemy, go to health pack	based on emotion, change behavior parameters that affect action	formula based	NA	Simulation of food searching agent while avoiding enemy agent scenarios
5	Han *et al.* [2008]	Agent	positive, negative	location choice for a shopping activity	based on the emotional value of utilities, select the location choice	Probabilistic utility based formulas	NA	multi-agent simulation of choosing shopping locations
6	Broekens *et al.* [2007]	Agent	valence	exploration vs exploitation	positive affect influences exploitation and negative affect influences exploration	RL, MDP, Boltzmann distribution	NA	Alternating-Goal and Candy task simulation
7	Huang *et al.* [2018b]	Agent	valence	robotic arm reaching behavior	based on emotional reward, take action	RNN with Hebbian, Oja update rules	OpenSim software, Jaco robotic arm	Simulation and system for musculoskeleton robotic arm target reaching task
8	Morgado and Gaspar [2005]; L. Morgado and G. Gaspar [2005]; L. Morgado and, G. Gaspar [2007]	Agent	joy, anger, fear, sadness	path to the target	based on emotional dispositions, cognitive processes are affected	Agent Flow Model, A* planner	NA	simulation of rescue agent
9	Lee-Johnson and Carnegie [2009]; C. P. Lee-Johnson and D. A. Carnegie [2007]	Agent	fear, anger, surprise, happiness, sadness	path planning and reactive, deliberative process	emotion-modulated path planning	formula based, A*	Matlab, C	MARVIN robot path planning simulation
10	Hogewoning *et al.* [2007]	Agent	valence	exploration vs exploitation	Artificial affect is used to directly control the amount of exploration by coupling valence to the β in the Boltzmann distribution used in action-selection	SOAR-RL (Q-learning), Hybrid-$\chi 2$ Method (Schweighofer-Doya, and Broekens-Verbeek method)	NA	Cue-inversion and Candy task simulation
11	Wu *et al.* [2022]	Agent	positive, negative	negotiation behavior, change of offer	based on valence, choose to explore or exploit	Weber-Fechner's law, Q-learning	NA	supply chain buyer, seller agent simulation
12	Jiang *et al.* [2022]	Agent	regret	driver lane change decision under collision risk	risk bias influences lane change	PRED model	NA	lane change simulation
13	Zhang and Zeng [2023]	Agent	positive, negative	turn left or right	emotional signals modulate or inhibit action selection	operant conditioning, memristor and rolls emotion model	Analog circuit	robot navigation out of maze simulation

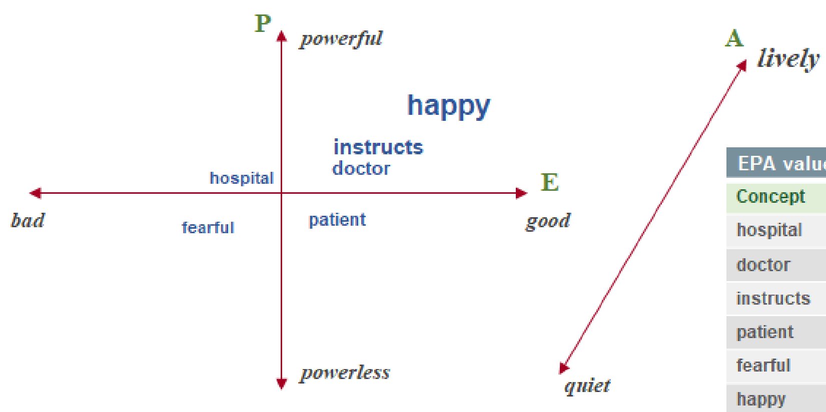
Chapter 3

Affect Control Theory

Affect Control Theory (ACT) is a socio-psychological model, that proposes social perceptions, actions, and emotional experiences of humans are governed by a psychological need to minimize incoherence between culturally shared, context-free sentiments about social situations, and context-dependent sentiments resulting from the dynamic behaviors of interactants in those situations [Heise, 2007]. This can be a powerful framework in designing and developing SIA, that can take a social identity itself, consider the culture and context it is in, while interacting with humans, and emotionally align its behavior. In this chapter, we first give background on affective semantic space used by ACT in section 3.1 and then give details on social interaction grammar of ACT, deflection, transient impression formation and optimal behavior computation in rest of the sections.

3.1 Affective Semantic Space

Context-free and contextual sentiments (termed “fundamental sentiments” and “transient impressions”, respectively) can be reduced to a real three-dimensional vector in an affective space. As found through cross-cultural studies of human sentiments, these three dimensions represent evaluation (how good vs. bad something is), potency (strong vs. weak) and activity (active vs. inactive) [Osgood *et al.*, 1957b]. The three-dimensional vector in the affective semantic space is referred to as an *EPA (Evaluation Potency Activity) profile*. Reducing sentiment to such a simple measure is found to account for over half of the variance in sentiment across cultures, and is therefore hypothesized to be an organizing principle of human socio-emotional experience [Osgood *et al.*, 1975]. Affect control theorists have compiled datasets of thousands of words along with average EPA ratings and variances obtained from survey participants [Heise, 2010].[1] For example, most English speakers agree that professors are about as nice (E) as students, however more powerful (P) and less active (A). The corresponding EPA profiles are (1.61 1.58 0.35) for professor and (1.49 0.31 0.75) for student.[2] Note that the profile attached to an identity indicates the sentiments as opposed to the function of the identity. In general, within-cultural agreement about EPA meanings of social concepts is high across subgroups of society, and cultural-average ratings from as little as a few dozen survey participants have shown to be stable over extended periods of time [Heise, 2010].



EPA values for few concepts from Indiana dataset

Concept	E	P	A
hospital	-0.92	0.12	-0.19
doctor	1.90	0.69	0.05
instructs	1.85	1.65	0.30
patient	0.90	-0.69	-1.05
fearful	-1.64	-0.94	-1.15
happy	2.92	2.43	1.96

Figure 3.1: 3D affective semantic space of concepts called *EPA (Evaluation Potency Activity)* profile. The values of a dimension indicate its intensity on a scale ranging from -4.3 to 4.3 values.

3.2 Social Interaction

ACT models the formation of transient impressions from interaction events with a minimalist grammar of the form 'Setting-Actor-Behavior-Object' ($S - A - B - O$). Each of these have an EPA profile: 'Setting' refers to the location context (e.g. 'hospital') or social event context (e.g. 'wedding'), while 'Actor' and 'Object' (as the linguistic 'object' of the behavior) refer to the identities (e.g. 'doctor' and 'patient'). The incoherence between out

[1] The EPA profiles in this thesis are from the Indiana 2002-4 dataset [Francis and Heise, 2006]. Code and datasets can be found at `affectcontroltheory.org` and at `bayesact.ca`.

[2] The values range by convention from -4.3 to +4.3 [Heise, 2010].

of context sentiments for each element in the grammar, and in context impressions when the four elements are perceived together can be used as an optimization loss to compute the best (most emotionally coherent) behaviors and identities. That is, ACT can answer questions of the form 'what should a doctor do to a patient in a hospital?' ('confer with') or 'who would harass a policeman in a protest' (a 'delinquent').[3] The answers to these questions are the identity and behavior that reduce emotional incoherence the most.

3.3 Deflection

According to ACT grammar [Robinson *et al.*, 2006], the fundamental sentiment vector f (represented by over-bar) is represented as follows:

$$f = (\bar{A}_e,\ \bar{A}_p,\ \bar{A}_a,\ \bar{B}_e,\ \bar{B}_p,\ \bar{B}_a,\ \bar{O}_e,\ \bar{O}_p,\ \bar{O}_a) \tag{3.1}$$

and the transient impression vector τ (represented by caret) evoked by an event is given by:

$$\tau = (\hat{A}_e,\ \hat{A}_p,\ \hat{A}_a,\ \hat{B}_e,\ \hat{B}_p,\ \hat{B}_a,\ \hat{O}_e,\ \hat{O}_p,\ \hat{O}_a) \tag{3.2}$$

Here A, B, and O represent the actor, behavior, and object respectively, and the subscripts e, p, and a represent the evaluation, potency, and activity respectively.

In ACT, sum of squared Euclidean distances between fundamental sentiments and transient impressions is called deflection D:

$$D_i = (f_i - \tau_i)^2 \tag{3.3}$$

Here, D_i are deflections for actor (A), behavior (B), and object (O) on the response dimensions of evaluation (e), potency (p), and activity (a). That is, i indexes over A_e, A_p, A_a, B_e, B_p, B_a, O_e, O_p, O_a (and also over S_e, S_p, S_a when settings are being considered). Deflection arises when impressions produced by an event differ from fundamental sentiments. Deflection that cannot be resolved produces psychological stress, which is a serious condition that can undermine one's health. Deflection is related to unlikelihood: the more deflection an event produces, the more that event seems stranger, more surprising, more unique and even inconceivable. Consider for example, a professor who yells at a student. Most observers would agree that this professor appears considerably less nice (e), a bit less potent (p), and certainly more active (a) than the cultural average fundamentals of

[3] These optimal behaviors and identities use ACT and the Indiana 2002-2004 dataset.

a professor. ACT treats the dynamics of emotional states and behaviors as continuous trajectories in affective space. Deflection minimization is the prescribed mechanism, but if deflection gets too high, it can result in re-identification of actor and/or object.

3.4 Transient impression formation

The transients existing after an event can be predicted from the transients that precede the event by the equation given by $\tau = tM$, where M is the matrix of prediction coefficients estimated in impression-formation research, with one column for each post-event transient being predicted. For example, Matrix M is 20 x 9, consisting of coefficients estimated from U.S male data on ABO. Vector t contains pre-event transients along with interaction terms that have been found to have predictive value in empirical analyses. Vector t given below is 1 x 20, hence τ is 1 x 9. (Refer to [Heise, 2007])

$$\begin{aligned} t = (1,\ \ \bar{A}_e,\ \ \bar{A}_p,\ \ \bar{A}_a,\ \ \bar{B}_e,\ \ \bar{B}_p,\ \ \bar{B}_a,\ \ \bar{O}_e,\ \ \bar{O}_p,\ \ \bar{O}_a,& \\ \bar{A}_e\bar{B}_e,\ \ \bar{A}_e\bar{O}_p,\ \ \bar{A}_p\bar{B}_p,\ \ \bar{A}_a\bar{B}_a,\ \ \bar{B}_e\bar{O}_e,& \\ \bar{B}_e\bar{O}_p,\ \ \bar{B}_p\bar{O}_e,\ \ \bar{B}_p\bar{O}_p,\ \ \bar{A}_e\bar{B}_e\bar{O}_e,\ \ \bar{A}_e\bar{B}_e\bar{O}_p) \end{aligned} \tag{3.4}$$

Note that ‘Setting’ is not considered here for simplicity. To show an example of how M and t affects the calculation of τ, the following equation shows the post-event Actor’s evaluation dimension estimated using the impression equations (considering non-zero values of first column of M which related to $\hat{A}_e$):

$$\begin{aligned} \hat{A}_e = -0.26 + 0.41\bar{A}_e + 0.42\bar{B}_e - 0.02\bar{B}_p - 0.10\bar{B}_a + 0.03\bar{O}_e& \\ +0.06\bar{O}_p + 0.05\bar{A}_e\bar{B}_e + 0.03\bar{A}_e\bar{O}_p + 0.12\bar{B}_e\bar{O}_e& \\ -0.05\bar{B}_e\bar{O}_p - 0.05\bar{B}_p\bar{O}_e + 0.03\bar{A}_e\bar{B}_e\bar{O}_e - 0.02\bar{A}_e\bar{B}_e\bar{O}_p& \end{aligned} \tag{3.5}$$

The coefficients in the above equation indicate the factors and the degree to which they contribute towards the post-event evaluation of the actor. For example, the positive coefficient on pre-event evaluation of actor $\bar{A}_e$, means that the good actors are evaluated more positively (in E) and bad actors are evaluated more negatively (in E), with a factor of 0.41. The positive coefficient on combination terms like pre-event behavior and object evaluation $\bar{B}_e\bar{O}_e$ means that the actors are evaluated more positively (in E) if they are observed doing good things to good people, or bad things to bad people, but more negatively (in E) if they are observed doing bad things to good people or good things to bad people, with a factor of 0.12. Similarly, the other dimensions can be calculated for $\hat{A}$, $\hat{B}$ and $\hat{O}$ giving τ value as mentioned in eq. (3.2).

3.5 Optimal Behavior

Action selection in an interaction would be based on any institutionally-appropriate, feasible, and sentiment-affirming behavior. For example, in a medical setting, there would be a doctor-patient interaction, where doctor's identity is generally considered as quite good and potent and somewhat active with an EPA profile as(1.90 0.69 0.05), whereas a patient identity is considered a bit good, less powerful and quite weak with an EPA profile as (0.90 -0.69 -1.05). The sentiment-affirming behavior for a doctor would be to treat or instruct the patient, so that his impression is maintained as good. If he does acts of yelling, cruelty etc., his impression will be bad and will cause deflection and conflict. An event seems more unlikely, uncanny, or unique as deflections (D) are larger. In ACT, the EPA profile for the optimal behavior is regarded as the one that minimizes the unlikeliness of an event, defined below. Note that i indexes over A_e, A_p, A_a, B_e, B_p, B_a, O_e, O_p, O_a (and also over S_e, S_p, S_a when settings are being considered).

$$k + \sum_{i=A_e}^{O_a} w_i D_i \tag{3.6}$$

where k is a constant and w stands for summation weights. From eq. (3.6) and (3.3), we have

$$k + \sum_{i=A_e}^{O_a} w_i (f_i - \tau_i)^2 \tag{3.7}$$

Minimizing unlikeliness or maximizing normality is obtained by setting partial derivatives of the above equation to zero and solving for behavior terms, giving us the suggested optimal behavior (for details refer [Heise, 2007]).

A software tool that implements ACT in Java is Interact[4]. It has a dictionary of various datasets across six nations, ranging from 1977 to 2007, and consists of EPA profile ratings for words (actor, behavior, object and settings), rated by male and female raters. This is useful in cross-cultural and historical analysis. In addition, it also displays interactant's facial expressions. Some sample interactions using Interact tool are shown in Table 3.1.

A recent extension of ACT, called *Bayesian Affect Control Theory (BayesACT)* [Hoey *et al.*, 2016; Schröder *et al.*, 2016; Hoey *et al.*, 2021], puts uncertainty at center stage by explicitly modeling variance in emotional sentiments about identities and contexts.

[4]Available for download at https://affectcontroltheory.org/resources-for-researchers/tools-and-software/interact/

Table 3.1: Sample ACT interactions

Emotion	**Actor**	**Behavior**	**Emotion**	**Object**	**Deflection**	**Socially aligned**
calm	doctor	instructs	fearful	patient	1.3	Y
angry	doctor	instructs	fearful	patient	3.1	N
calm	doctor	yells at	fearful	patient	6.8	N
angry	doctor	yells at	fearful	robber	0.7	Y

BayesACT generalises ACT by maintaining multiple hypotheses of behaviors and identities simultaneously as a probability distribution. It uses a Partially Observable Markov Decision Process (POMDP) [Åström, 1965] to combine uncertainty at multiple levels with decision-making. BayesACT accounts for the dynamic fluctuation of identity meanings for self and the other during interactions, elucidates how people infer and adjust meanings through social experience, and shows how stable patterns of interaction can emerge from individuals' uncertain perceptions of identities. For a more complete presentation of BayesACT, and a more detailed analysis of the 'doctor' example, see [Hoey *et al.*, 2021].

For long-term adoption of ACT in human-agent social interaction, there is a need to design, simulate and validate social agents in real-world settings. In this thesis, we first implement an ACT agent as a brain-inspired neural model and simulate a doctor-patient interaction and a few game-play scenarios. Since ACT is originally text based, we attempt a visual predictive system that can infer social event context and perceived affect in a given scene. We model ACT in a care giving application where pet-robots assist older adults and persons with dementia. Predicting identities and behavior is out of scope of this thesis.

Chapter 4

Biologically-inspired model of ACT

"What I cannot create, I do not understand"

- Richard Feynman

Social neuroscience has primarily focused on a sense of self identity and how a person's mind creates a perception of another person, whereas affective neuroscience has focused mainly on mapping emotions in the brain [Barrett and Satpute, 2013]. Social interaction in humans involve emotions and decision-making in a given context. We can develop socially intelligent agents in a non-biological fashion to replicate behavior of certain individuals, or a group. But it may not replicate the human brain processes involved in simulating such behavior. This is analogous to producing a flying behavior. We can get inspired by birds and build a plane that can fly, based on physics and engineering, but without having exact same wings and flying mechanism of birds. On the other hand, if we were to build a bird-like flying agent and implement bird-like processes of flying behavior, we could do it in a biologically-inspired fashion. This chapter is an attempt to provide an SIA with a brain-inspired model of emotion and decision-making, and to combine social and affective domain perspectives, as a step towards dissolving their artificial boundaries. Human processing may have lot of variation, and we are still discovering human brain and its mechanisms. Building a biologically-plausible model may help understand some human processes better. This could potentially be implemented in an artificial brain someday. In this chapter, we implement Affect Control Theory (ACT) as a brain-inspired model, towards gaining deeper understanding of the neural mechanisms of emotions in decision-making.

In section 4.1, we first give some background on human brain and neural communication. In section 4.2, we outline brain components that may be involved in affective decision-making, and then give a brief perspective of emotions in the brain in section 4.3. Section 4.4 introduces brain simulation library called Nengo and its modules. In section 4.5, we present a brain-inspired neural model of ACT using Nengo. We demonstrate a few social interaction simulation scenarios in section 4.6 and present the results in section 4.7. Conclusion and future work is provided in section 4.8.

4.1 Human Brain and Neural Communication

The human brain is an amazing part of the body that orchestrates our overall functioning in an energy efficient manner. At a very high level, the brain under our skull is made up of cortical sheet, sub-cortical areas (including basal ganglia, thalamus, brain stem and other nuclei) and cerebellum [Barrett, 2017; Eliasmith, 2013]. The human brain consists of approximately 100 billion neuron cells [Eliasmith, 2013]. To understand how the brain works, it is important to understand the communication between neurons, which are electro-chemical in nature.

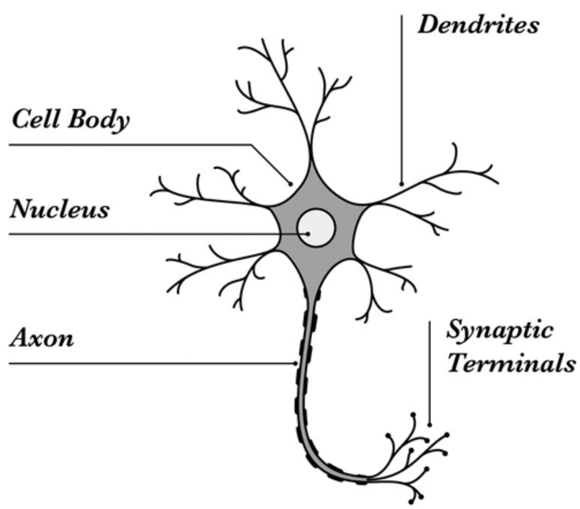


Figure 4.1: A neuron structure, from [Barrett, 2017]

A structure of a single neuron as depicted in [Barrett, 2017] is shown here in Figure 4.1. A neuron cell body (called soma) consists of branch-like structures called dendrites on one end and a root called axon terminal on the other end. The gap between one neuron and another is called synapse. If the current in a neuron cell body exceeds some threshold, a voltage ‘spike’ is generated, such that an electrical signal is passed from one neuron to the adjacent neuron via axon of the source neuron, releasing chemicals (called neurotransmitters) into synapse. These are picked up by dendrite receptors (special proteins) of the following neuron. Neurotransmitters can excite or inhibit the receiving neuron. The current in the receiving dendrite is called as postsynaptic current (PSC), which changes the firing rate of neurons. In this way, multiple neurons can fire and influence many other neurons, causing information signal flow in the brain. The neurons are heterogeneous in nature, such that responses for the same current injected to the same kind of neurons can be different [Eliasmith, 2013].

4.2 Cortex-Basal Ganglia-Thalamus loop

The cortex has neurons, that are believed to map representation of brain states to basal ganglia. The output of basal ganglia is the selection of an action with highest utility, which disinhibits appropriate thalamus areas, resulting in the relay of information back to the cortex. This is called the cortex-basal ganglia-thalamus loop. A model of an action selection in basal ganglia was presented by [Gurney *et al.*, 2001], and converted to a computational model in [Eliasmith, 2013] as shown in Figure 4.2. The matrix M_b specifies a known context, which when computed as a product with current cortical state, provides a similarity measure to the known contexts. The output of basal ganglia goes through thalamus, back to cortex with matrix M_c, which specifies appropriate cortical state as the consequence of the selected action. Some evidence suggests that single actions can be selected without basal ganglia, but chains of actions seem to involve basal ganglia [Aldridge *et al.*, 1993]. In this chapter, we consider the role and influence of basal ganglia in decision making and action selection.

4.3 Emotions in the brain

There have been various studies to find fingerprints of emotions in the brain. A classic belief is that the sub-cortical region called amygdala is related to fear processing. But evidences have shown that its activity can be related to emotional as well as non-emotional events such as feeling of pain, learning something novel, or meeting new people. Modern neuroscience suggests that no single brain area is dedicated to emotion. When current world

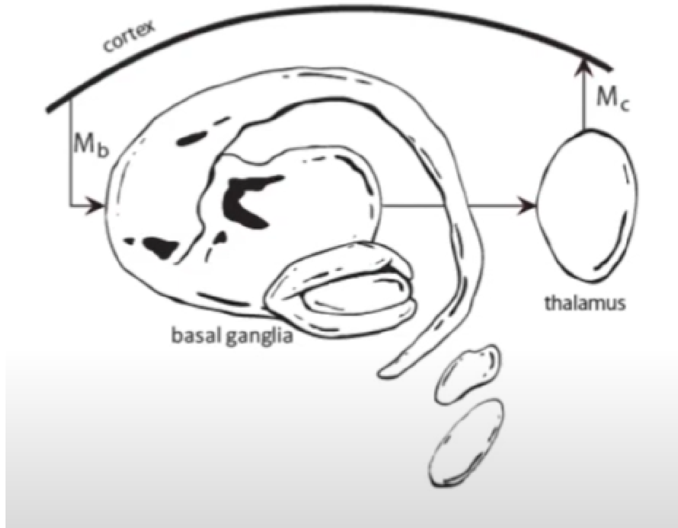


Figure 4.2: The cortex-basal ganglia-thalamus loop, from [Eliasmith, 2013]

context and body state causes brain to predict the next state and action, based on past experiences, the changes that take place in the body may be labelled as a particular emotion by the perceiver, based on the culture and conditioning. This is brain's construction of an instance of emotion [Barrett, 2017]. In this chapter, we consider emotional meanings of concepts as part of cortical representation.

4.4 Nengo

Nengo [Bekolay *et al.*, 2014] is a python based neural simulator, that follows Neural Engineering Framework (NEF) [Eliasmith and Anderson, 2004]. NEF outlines three principles of representation, transformation and dynamics, that enable construction of large-scale neural models. Representation means that population of neurons represent information with time-varying vectors of real numbers. The corresponding Nengo object is called an '**Ensemble**'. Transformation between two neuron populations is set as synaptic connection weights of a function between them. This is computed as product of decoding weights for that function in the first population, encoding weights for that function in the second population and any linear transformation. The corresponding Nengo object is called a '**Connection**'. Dynamics is when recurrent connections are introduced for persistent activity in the neural systems. The corresponding Nengo object could be some of the '**Networks**', or if Ensemble is connected to itself. Other Nengo objects which are useful are: '**Probe**', which gathers simulation data for analysis, '**Model**' that encapsulates a Nengo model, '**Node**' represents sensory and/or motor outputs. All the Nengo objects defined above were used in implementing our neural model. The Nengo interface used for demonstrating examples in this chapter is shown in Figure 4.3. Here, the vision components are a vector representation of perceived input, mapped in the cortex. We used five vision inputs from a stimulus (called as 'stim') related to ACT, for the demonstration of social interaction scenarios, which will be discussed in the following section.

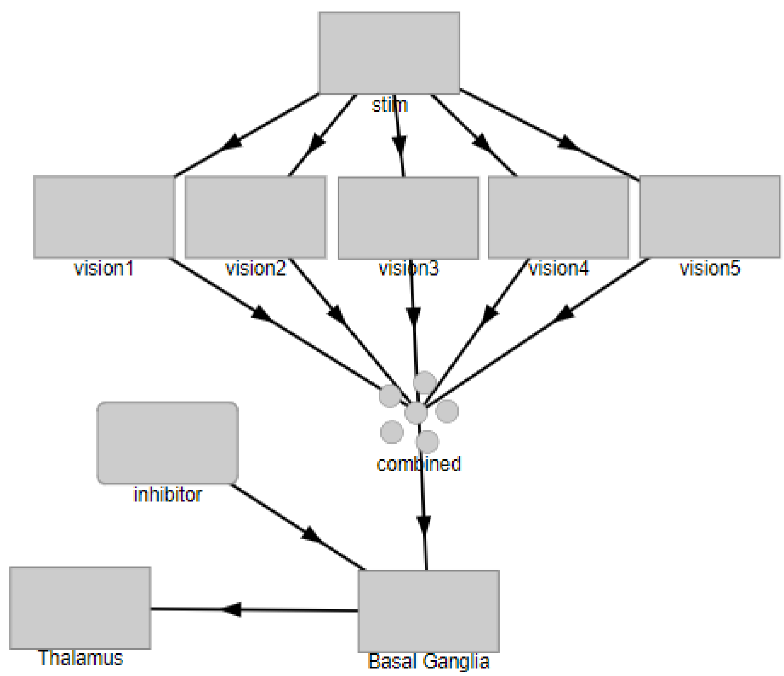


Figure 4.3: Nengo GUI

4.5 Neural Model of ACT agent

The novel contribution of this chapter is to take the underlying mathematics of ACT as discussed in Chapter 3 and implement it using the spiking neurons. The goal here is to show that the calculations required by ACT can be implemented by spiking neurons, using anatomical structure that fits the cortex, basal ganglia, and thalamus. In particular, it is striking that the overall form of the theory maps very well onto a neural model of the cortex-basal ganglia-thalamus loop that has been previously used to model a variety of tasks [Eliasmith *et al.*, 2012].

The core part of the algorithm that is modelled here and its relation to the neural model of the brain is shown in Figure 4.4. In this work, the mechanisms for maintaining and tracking the EPA values of the current situation is not modelled; rather, focus is on the calculation of deflection and hence unlikeliness, given the event perception from an object's (AI agent) perspective and time t. That is, given the EPA values of the current situation, the question is: what action should be performed by the object of the event?

This maps well onto the traditional roles of the cortex, basal ganglia, and thalamus. Neurons in the cortex (1 in Figure 4.4) will represent the EPA values of the interaction, viz. Modifier for Actor (Mod_{actor}), Identity of the Actor (I_{actor}), Behavior of the Actor (B_{actor}), Modifier for Object (Mod_{object}), Identity of the Object (I_{object}). Setting is not considered here for simplicity. The connections between cortical neurons and basal ganglia neurons (2) will compute eq. (3.7), the basal ganglia (3) will find the action with the largest deflection minimizing utility value, and the thalamus (4) will activate that particular action.

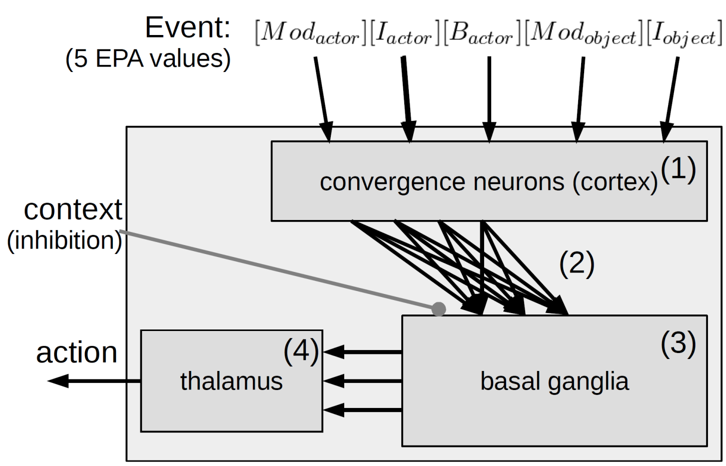


Figure 4.4: Neural implementation of ACT

While the overall mathematical function of this system is easy to describe and implement, it will be shown how spiking neurons can perform these operations. In particular, here Neural Engineering Framework (NEF; [Eliasmith and Anderson, 2004]) is used, which is a general method for finding how to connect simulated neurons so as to get the best approximation of any given algorithm. In general, the idea here is that the activity of groups of neurons can be thought of as representing vectors, and the connections between groups of neurons can be thought of as computing functions on those vectors. If we know the set of functions that we want to compute then we can perform a sequence of local optimizations (one for each set of connection weights) that will find the best approximation of the algorithm, given whatever type of neurons we want to use (including spiking and non-spiking neuron models).

For the basal ganglia and thalamus, we can make use of already-existing models of how to use the NEF to implement exactly the function that is desired here: a system that takes in a set of values from eq. (3.7) and determines which one is the largest utility, say U, outputting that information to the thalamus. This has been previously shown to both map on well to the anatomy of the basal ganglia and to exhibit realistic reaction times [Stewart *et al.*, 2010]. This system has been used in many previous models, including models of the bandit task [Stewart *et al.*, 2012] and the large functional brain model Spaun [Eliasmith *et al.*, 2012]. The same is used here without adjusting any parameters. Also, an inhibitory "context" input is used, that provides a large negative value for any actions that should not be considered.

While the basal ganglia and thalamus model take care of computing which of the action values has the largest deflection minimizing utility U (i.e. which action should be taken), this still leaves the question of how to have neurons calculate the eq. (3.7) values for each action, given the basic EPA values.

Since this is simply a function, it is possible to train a neural network to approximate that function. However, the general challenge of neural networks is that if the function being approximated is too complicated, we will need a very large neural network to do this (either very deep or very broad, or both). Importantly, the networks generated using the Neural Engineering Framework have been analyzed in terms of the class of functions that they are good at approximating when using a Leaky Integrate-and-Fire neuron model with the default distribution of tuning curves [Eliasmith and Anderson, 2004]. This analysis indicates that these neurons are best at approximating functions that consist of linear combinations of low-degree polynomials. Crucially, this is *exactly* the form of the calculation being done here (see eq. (3.5)). This means that we can use small numbers of neurons (here we use 1500) with the same parameter settings as has been used in the other biological models to approximate this function.

An example of the overall behavior of the resulting model is shown in Figure 4.5. The

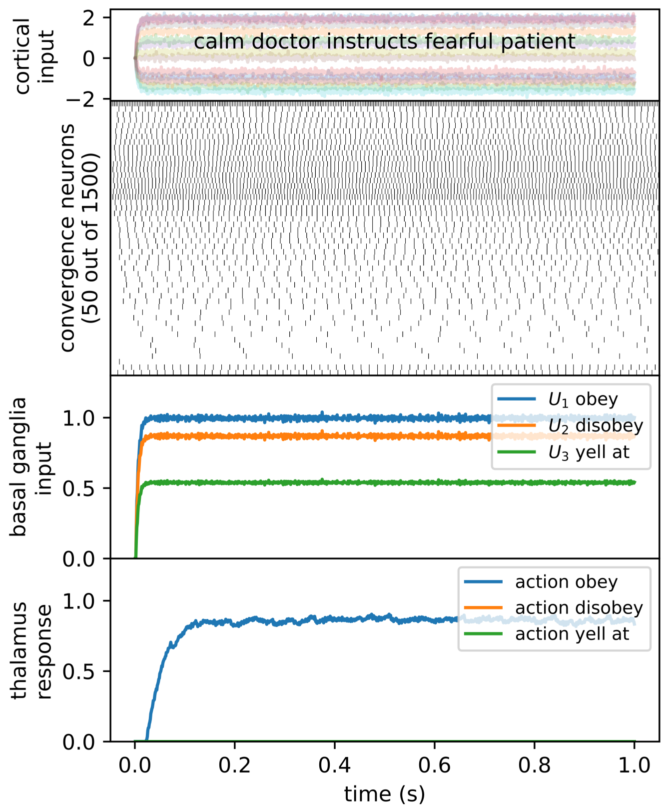


Figure 4.5: Example of behavior of NeuroACT

input is the EPA values for each of the 5 relevant terms. In this case, the situation is

$$[calm][doctor][instructs][fearful][patient]$$

and the corresponding input EPA values are [1.97 1.32 -1.4][1.9 0.69 0.05][1.85 1.65 0.3][-1.64 -0.94 -1.15][0.9 -0.69 -1.05]. These values are fed into the convergence neurons. These connections are completely random, meaning that any particular input will produce some random pattern of neural activity that is unique to that input. From that activity, the connection weights from the convergence neurons to the basal ganglia compute the eq. (3.7) function for all of the different actions in parallel. For simplicity, here we only plot three of those actions: 'obey', 'disobey' and 'yell at'. Finally, the basal ganglia model finds the largest of these activity values (i.e., 'obey') and directs that result to the thalamus, so the object of the event, which is the patient in this case, can perform for better interaction. This is also the optimal behavior according to the mathematical model.

4.6 NeuroACT Simulation

To simulate NeuroACT model for social interaction involving affect, decision-making and behavior, a single play of prisoner's dilemma game scenario was used. Of the two players involved; one represents a simulated human player agent (Actor) and the other represents NeuroACT AI agent (Object). In the play round, each player can decide to either give two coins to the other player (cooperation strategy) or take one coin (defection strategy) from a common pile. Players can maximize their individual returns by defecting, or they can jointly maximize their scores through mutual cooperation. In simulation scenarios, the AI agent perceives the emotional state, identity and behavior of the human agent, and outputs the optimal behavior it would choose ('give' or 'take') based on the ACT prescription of deflection-minimization. Given that the human player has acted, the simulation outputs what action the AI agent takes in that situation. The decision-making dynamics over the time scale are demonstrated, such that if the perceived emotion of the human agent changes during the play round, the AI agent changes its strategy as well. Change in behavior can also result if the perceived identity of the actor and/or object changes. For first two simulations, identity of both the players was kept as 'stranger', but for the third simulation, the identity was changed from 'stranger' to a 'friend'. EPA profiles used for identity, modifiers and behaviors are as below:

$$[happy] : [2.92, 2.43, 1.96]$$

$$[angry] : [-1.45, -0.30, 1.13]$$

$$[stranger] : [0.02, -0.09, -0.23]$$

$$[friend] : [2.75, 1.88, 1.38]$$

$$[gives\ to]\ : [1.60, 1.47, 1.55]$$

$$[takes\ from]\ : [-1.40, 1.62, 1.50]$$

Inhibition: The dictionary of EPA profiles was used from Indiana 2002-4 dataset, which consists of 500 behaviors related to various contexts, out of which we chose 2 behaviors for the game context viz., 'gives to' based on phrase 'sell something to' and 'takes from' based on phrase 'capture' for the simulation. If there was no mechanism of inhibitory neurons, AI agent would have selected a deflection-minimizing behavior out of 500 options, but in our case, it selects between 'give' or 'take' behavior only and others get inhibited.

To demonstrate the behavior of the model and show its ability to use neurons to perform similar calculations as found in the standard Affect Control Theory, we provide cortical input of 5 sets of EPA values representing a particular situation. Since neurons require time to respond, we hold this input constant for 0.5 seconds and then present a new situation. In particular, we manually adjust the recognized emotion from 'happy' to 'angry', as this causes ACT to produce a different action. It should be noted that, in this example, the 'object' is meant to correspond to the NeuroACT AI agent itself.

Scenario 1: Human agent cooperates with AI agent, with change in emotion

Perception at time $t \leq 0.5$:

[*happy*][*stranger*][**gives to**][*happy*][*stranger*]

Perception at time $t > 0.5$:

[***angry***][*stranger*][**gives to**][*happy*][*stranger*]

Scenario 2: Human agent defects with AI agent, with change in emotion

Perception at time $t \leq 0.5$:

[*happy*][*stranger*][**takes from**][*happy*][*stranger*]

Perception at time $t > 0.5$:

[***angry***][*stranger*][**takes from**][*happy*][*stranger*]

Scenario 3: Human agent defects with AI agent, with change in emotion and identity

Perception at time $t \leq 0.5$:

[*angry*][*stranger*][**takes from**][*angry*][*stranger*]

Perception at time $t > 0.5$:

[***happy***][***friend***][**takes from**][***happy***][***friend***]

4.7 Results

Results for the simulation runs for scenarios 1 and 2 are shown in Figure 4.6 and 4.7 respectively. In both scenarios, the resultant behavior changes from 'give' to 'take' on

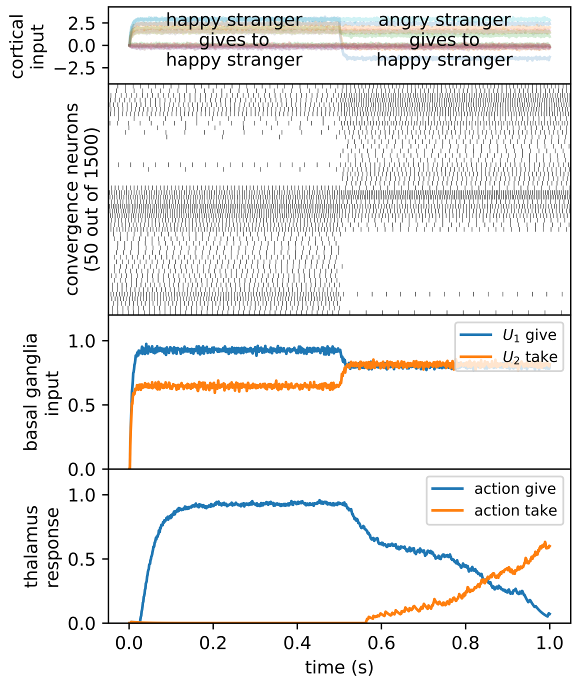


Figure 4.6: Human agent cooperates with AI agent, with change in emotion

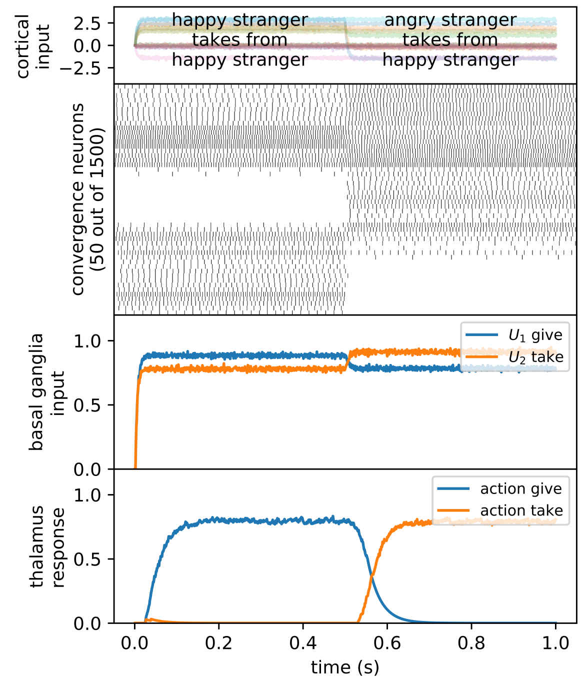


Figure 4.7: Human agent defects with AI agent, with change in emotion

perceiving the emotion of the human agent that changes from 'happy' to 'angry', given the affective dynamics. In scenario 1 (Figure 4.6), the change in behavior seems slower and more deliberate than in scenario 2 (Figure 4.7), where the change is faster and somewhat automatic. This may be due to the actor's behavior being more positive in scenario 1 as compared to scenario 2. In scenario 3 (shown in Figure 4.8), the identity of both human and AI agent are perceived as 'stranger' in an angry state, and the human defects with the agent. In this case agent also decides to defect. But when the perception changes from an angry 'stranger' to a happy 'friend', the agent decides to cooperate. Code for this work is available at *https://github.com/aarti9/neuroact*

NeuroACT shows how affect influences decision-making and behavior. The behavior chosen by the model matches with its non-neural counterpart in choosing the optimal behavior as prescribed by ACT. The ability of the neural model to handle time dimension is important for the temporal order of information processing similar to human brain circuit [Gupta and Merchant, 2017].

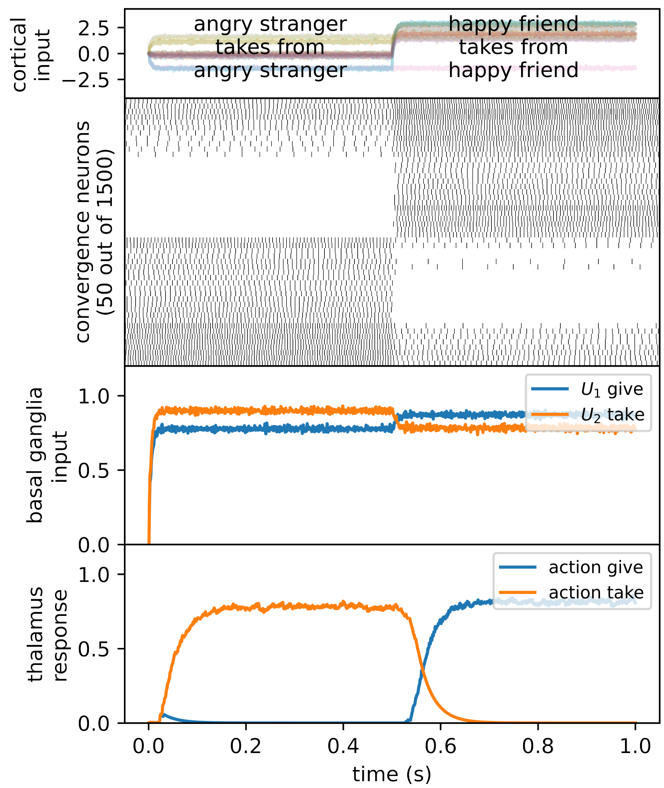


Figure 4.8: Human agent defects with AI agent, with change in emotion and identity

4.8 Conclusion and Future Work

Social interaction is a challenging area to replicate in brain simulations. NeuroACT is a novel contribution implementing affective social interaction in spiking neurons. It is a generalized and extensible neural model of ACT, capable of providing an AI agent with the ability to interact with the other AI agents or humans. Input is an interaction perception and output is an optimal behavior selection. This is a step towards making socially intelligent AI agents.

A specific doctor-patient interaction is demonstrated as an example for the model. Simulation of a single play in prisoner's dilemma game is provided. This can be iterated as well, taking into account that in the next round of play, the actor and the object change. NeuroACT can be used to model any other interaction. Future enhancements can include settings for additional context, such as location or social event. The model can be expanded using similar methods to predict the emotion and generate re-identification of the actor and the object post-interaction. This system can be enhanced by incorporating some sensorimotor signals to integrate with physical robots. Some other improvements could be considered involving a working memory component for the agent to utilize experience from the previous interactions. The input to the model is a generic input, which can incorporate visual, textual, or auditory forms, as all would eventually translate into verbal concepts. Advances in neuroimaging techniques like hyperscanning to study the inter-brain synchronization [Liu *et al.*, 2018] in social interaction may give more insight into the neural mechanisms at play.

Chapter 5

Social Context Understanding

"It's not the events of our lives that shape us but our beliefs as to what those events mean."

- Tony Robbins

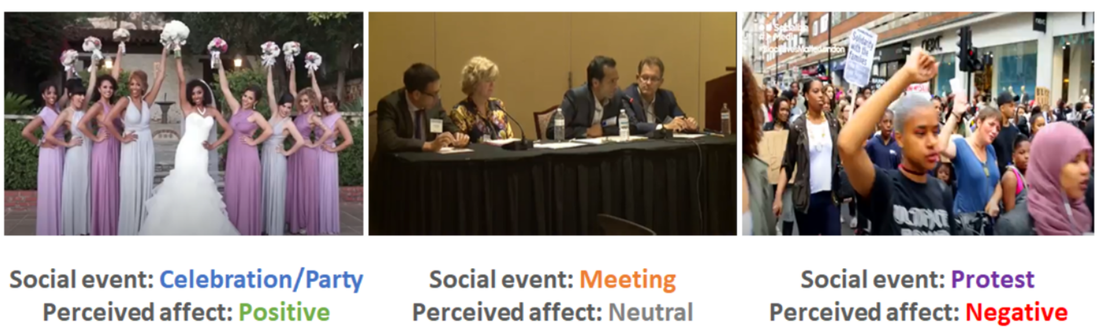


Figure 5.1: Sample frames from VGAF dataset [Sharma *et al.*, 2021] video clips with social event search keyword and perceived affect label.

Social situation and relationship amongst people impacts the perception of inferred affect [Dhall *et al.*, 2015b]. Social event context can be helpful for affective AI systems in making context-sensitive decisions [Marpaung and Gonzalez, 2017]. If AI agents have to co-exist with humans, it is imperative for the agents to understand the human social context. In this work, we aim to develop a visual system for the SIA to be able to predict social event context and perceived affect in a social scene. This can be leveraged by NeuroACT and other ACT-type models and be used for other applications. Computer vision research in affective computing has seen a lot of advances, from early research on facial expression recognition to recent context-aware emotion perception. Group affect recognition [Baveye *et al.*, 2015; Dhall *et al.*, 2012, 2015a; Guo *et al.*, 2018; Huang *et al.*, 2018a; Quach *et al.*, 2018; Tan *et al.*, 2017] and group social event detection [Ahmad *et al.*, 2016; Bossard *et al.*, 2013; Wang *et al.*, 2016; Won *et al.*, 2017; Xiong *et al.*, 2015] have also gained attention and are active areas of research.

Social event and affect prediction in a group setting is a challenging task, as not only the specific detail of individual's action matters, but also the overall context and affect play an important role for overall scene understanding. Many early studies in affective computing were focused on emotion detection using only facial expressions, but the importance of understanding the context has been underlined in many psychological studies [Barrett *et al.*, 2011, 2019; Martinez, 2019]. Context may involve various facets such as location, social events, culture, politics, social identities of people, etc. Here, we focus on social event context as a step towards deeper, more human-like, understanding of a scene. Social event contexts such as a group setting can have various correlated facial expressions and actions. For instance, people in a wedding or a sports victory scene may be crying, but they may be tears of joy. Considering only facial expressions or only actions may make it non-trivial to predict the correct event and the perceived affect. Many group videos can be perceived as multiple categories of context as well. Family group members may meet casually, but end up fighting or quarrelling over some issue. A wrestler in a sport may shake hands with the opponent at the beginning of a match and in the end hug him/her as a courtesy gesture.

Contextual and affective information, if tapped into, can give insights and guide artificial intelligence towards being more human-like in nature. A dataset that captures group affect label information in videos of different social events is the Video Group AFfect (VGAF) dataset [Sharma *et al.*, 2021]. The dataset has three classes (*positive, neutral* and *negative*) for perceived group affect, which are referred as *'affect label(s)'*. Fig. 5.1 shows few sample video frames from the VGAF dataset with social events such as *Celebration/Party*, *Meeting* and *Protest*. Social event context here includes overall scene as well as human activity in a group setting. Videos are also labelled as *positive*, *neutral* and *negative* respectively. Applications of prediction tasks on this dataset include inferring social event and perceived affect from the scene for SIA, video captioning, automatic video annotation, video search and retrieval, visual question answer (VQA), surveillance, event based affective forecasting, and the ability to give agents the capability to understand the scene, make better decisions and behave appropriately according to the situation.

For social scene videos, we needed a method to capture the spatial as well as the temporal aspect of the scene, along with short-term and long-term dependencies. One such method is the transformer network [Vaswani *et al.*, 2017]. Transformer based models have outperformed in many tasks in NLP and computer vision, showing improvements over traditional convolution-only models. They have been successful in image related tasks and are extendable to videos with capabilities for handling multi-modal information. In this chapter, the focus is on multi-task learning for social event classification in group videos using a transformer network, involving spatial and temporal transformer encoders only. Perceived group affect information is also utilized and shown improvements to the base model. The contributions are as follows:

- We introduce ten social event category labels for the VGAF dataset [Sharma *et al.*, 2021] clips. The dataset originally had perceived group affect labels only.
- We use a spatio-temporal transformer network for social event and perceived affect classification task on the dataset and show that it out-performs a convolution-only network and another spatio-temporal transformer-only network.
- We propose multi-task joint learning of social event and perceived group affect in videos, based on two strategies:
  1. Affect information as a regularizer with spatio-temporal transformer only, and
  2. Affect information with spatio-temporal transformer and a convolution classifier
- From experimental evaluations, we show that the use of affect improves social event classification. We also show that a spatio-temporal transformer based model outperforms the state-of-the-art for group affect prediction on the VGAF dataset using the visual modality only.

This chapter is organized as follows: Section 5.1 gives some background information. Section 5.2 gives details about the dataset. Section 5.3 introduces proposed overall architecture and variations in models. Section 5.4 provides details about experiments. Section 5.5 presents results and discussion. Section 5.6 gives ethical impact of this work. Section 5.7 provides a conclusion and future work.

5.1 Background

We provide background on social event and group affect prediction. We also discuss video vision transformers which is a motivation for our proposed method.

5.1.1 Social Event Prediction

Automatic social event detection is explored in computer vision for many different applications. One of the early works on automatic social event detection was done by Bossard et al. [Bossard *et al.*, 2013], where they proposed a dataset (PEC) comprising of 61,364 photo collections from 14 different categories and used hidden Markov based model to recognize events in the collection. In contrast to this work, Wang et al. [Wang *et al.*, 2016] identified an event from an image to indicate the importance of each image for the application of creating a photo album. Further, a method proposed by Xiong et al. [Xiong *et al.*, 2015] also accounted for the interactions of humans and objects to recognize events. The proposed approach is validated on the WIDER dataset which contains 60,000 images from 40 different event categories. In another study, Huang et al. [Huang *et al.*, 2010] proposed a method for emotion recognition only on the textual data. The study focused on the relation between emotion and event which give rise to a particular emotion. Events can be recognized from the data of different modalities, however, there is often a media gap between various modalities of the data [Zhou *et al.*, 2020]. Having temporal data seems beneficial for social event recognition.

Social event related research and datasets which are compared to VGAF are mostly image based such as EiMM [Mattivi *et al.*, 2011], SED [Reuter *et al.*, 2013], CUFED [Wang *et al.*, 2016], PEC [Bossard *et al.*, 2013], USED [Ahmad *et al.*, 2016], WIDER [Xiong *et al.*, 2015]. Video based datasets related to social events have been mainly used for surveillance and violence detection such as VSD [Constantin *et al.*, 2020], ViF [Hassner *et al.*, 2012], complex events [Jiang *et al.*, 2011] or only specific event [Conigliaro *et al.*, 2015]. To the best of our knowledge, VGAF is the only unconstrained video dataset that has a variety of different social events along with labelled affect information.

5.1.2 Group Affect Prediction

Group affect recognition in a visual scene poses some interesting challenges as a lot of interactions happen together. In social psychology, group affect was analyzed from a top-down and bottom-up perspective in [Barsade and Gibson, 1998; Kelly and Barsade, 2001], both playing an important role in contributing to the detection of collective affective state. One of the initial works on automatic group-level affect combining the top-down and bottom-up approaches can be found in [Dhall *et al.*, 2012]. The focus was on the happiness intensity estimation in images of group of people based on social context using ‘in the wild’ database called HAPpy PEople Images (HAPPEI). The two approaches in the context of attention mechanism applied to image captioning and visual question answering tasks is studied in [Anderson *et al.*, 2018]. To analyze the overall affect of a group of people, the common practice is to combine the facial and background features [Abbas and Chalup, 2017], [Rassadin *et al.*, 2017]. Some other works on video based group affect can be found in [Baveye *et al.*, 2015], [Mittal *et al.*, 2021]. In this chapter, the focus is to use the group affect information and check if it improves the social event recognition. VGAF [Sharma *et al.*, 2021] is a relevant group dataset, having videos of different social events, labelled with positive, negative and neutral group affective categories.

5.1.3 Video Vision Transformers

Transformer network based on self-attention proposed in [Vaswani *et al.*, 2017] has proven success in machine-translation and has been adopted in other NLP and computer vision tasks as well. Videos can be represented as a sequence of image frames, similar to a sentence being a sequence of words. To understand a video scene better, it is important to capture long-range temporal and contextual information along with short-range local information. Convolution based networks tend to focus more on localization, while transformer based networks can pay attention to overall context as well. Many video related tasks that use transformers include group activity recognition [Gavrilyuk *et al.*, 2020], video classification [Chen *et al.*, 2018], [Girdhar *et al.*, 2019]. [Wang *et al.*, 2020]. An image model called "Vision Transformer" (ViT) [Dosovitskiy *et al.*, 2020] had introduced a process to extract patches from an image and embed them with positional information into spatial transformer. Later works such as ViViT [Arnab *et al.*, 2021], TimeSformer [Bertasius *et al.*, 2021], STAM [Sharir *et al.*, 2021] adopted similar process to video by extending the self-attention mechanism from 2D image to the spatio-temporal dimension, considering video as a sequence of patches extracted from the individual frames. As in ViT, each patch is linearly mapped into an embedding and augmented with positional information. This makes it possible to interpret the resulting sequence of vectors as token embeddings which can be fed to a Transformer encoder, analogously to the token features computed from words in NLP.

Most of the research in social event and group affect prediction focuses on the individual task using images only, like[Bossard *et al.*, 2013] for event and [Dhall *et al.*, 2015b; Kosti *et al.*, 2017] for affect. Hence, there is a scope to explore the relation between social event and affect prediction. The recent success of different video based transformer methods inspired us to explore the effectiveness of the same on VGAF dataset. To the best of our knowledge, this is the only video classification work based on social event context and affect in group videos, and that shows some correlation of affect and social event concepts.

5.2 Dataset

We used the Video Group AFect (VGAF) dataset [Sharma *et al.*, 2021], which is one of the first video datasets in the wild, containing videos of different social events, and labelled with *'Positive'*, *'Neutral'* and *'Negative'* labels for perceived group-level affect (*'affect labels'*). The videos were collected from YouTube using specific event keywords, having a 'Creative Commons CC BY' license. The dataset consists of 4,183 video clips of 5 second duration each. The clips are derived from 326 original YouTube videos after segmentation [Sharma *et al.*, 2021]. In this work, the event keywords that were used for video search, were combined into 10 classes of *'Quarrelling', 'Meeting', 'Sports', 'Funeral', 'Show', 'Group Activities', 'Protest', 'Casual Family Friends Gathering'* ('Casual FFG' used as short form), *'Celebration Party' ('Celebration' used as short form), 'Fighting'*. Details of the search keywords and the social event categories are given in Table 5.1. Event categorization was done by a team of three researchers (AM, GS, RK), who also verified each video and the assigned event category label. Most of them had consensus. In case of few disagreements, consensus was arrived at by discussing the rationale behind each rater's categorization.

Table 5.1: Social event categorization based on keywords used in search queries of the VGAF dataset [Sharma *et al.*, 2021].

Social event	**Keywords**
Casual Family Friends Gathering	Family get-together, Friends conversation, Festivals
Celebration Party	Birthday party, Celebration, Crowd cheering, Match winning, Wedding
Fighting	Friends fighting, Street fight, Boxing, Crowd fighting, Fighting
Funeral	Condolence meeting, Funeral
Group Activities	Community service, Group performance, People on street, Religious gathering, Classes, Group dance, Group task, March-past
Meeting	Event announcement, Group discussion, Interview, Meeting, Conversation, Discussion, Press conference, Seminar, Speech
Protest	Stone pelting, Violent Protest, Protest
Quarrelling	Argument over discussion, Argument
Show	Concert, Live shows, TV shows, Talk shows
Sports	People watching match, Wrestling

5.3 Network Architecture

We propose a spatio-temporal transformer architecture for video understanding, based on the factorised encoder model as discussed in ViViT [Arnab *et al.*, 2021] and STAM model [Sharir *et al.*, 2021]. We named it as **Vi**deo**T**ransforme**R** or **'ViTR'**. We further modified it to do Multi-Task Learning (MTL), utilizing affective information. We adopted two different strategies to achieve MTL, which are discussed in section 5.3.2. Our intuition was to utilize affect information as an influence on the social event to disambiguate social event scenes. The architecture using affect as an indirect influence is given in strategy 1, which is **A**ffect as a **R**egularizer (**'ViTR-AR'**) and that using affect as a direct influence is given in strategy 2, which is **A**ffect **D**ependent (**'ViTR-AD'**). For schematic of these network architectures refer Fig. 5.2.

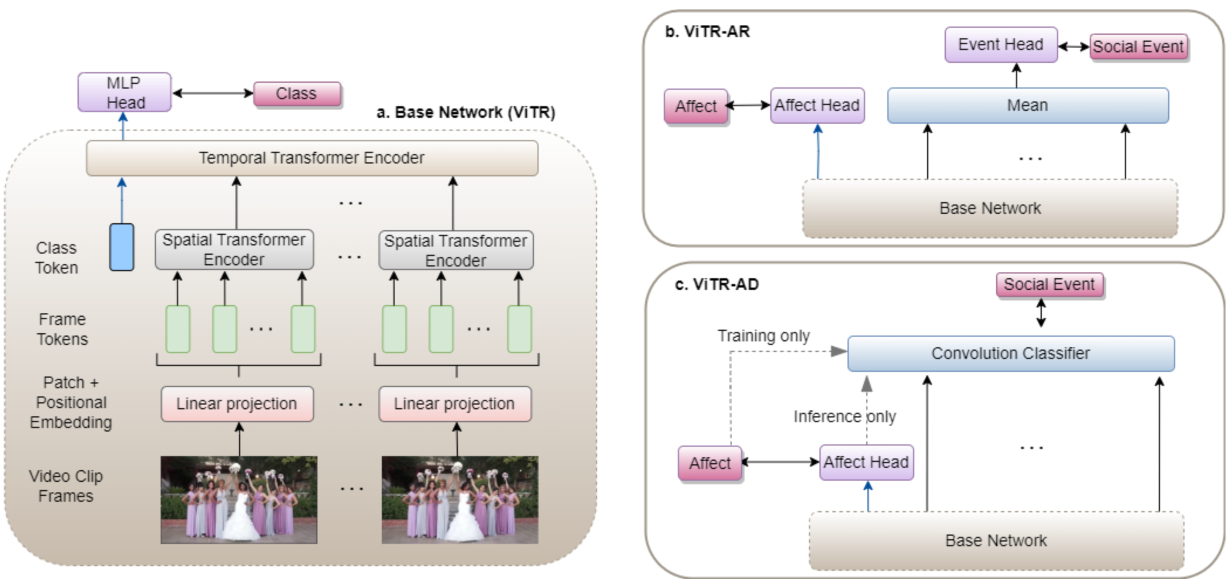


Figure 5.2: Illustration of network architectures for a. **ViTR**, b. **ViTR-AR** and c. **ViTR-AD** models.

5.3.1 Base Transformer

The base transformer 'ViTR', depicted in Fig. 5.2 a. consists of spatial and temporal transformer encoder, similar to the idea in Model 2 of ViViT [Arnab *et al.*, 2021] and STAM [Sharir *et al.*, 2021]. The video clip is divided into frames, preserving the temporal sequence. The frames are linearly projected into frame tokens and become an input to the spatial encoder. The spatial encoder incorporates interactions between tokens extracted from the same temporal frame. The frame-level representations from all spatial encoders are forwarded through a temporal encoder, which consists of temporal transformer layers. This takes care of interactions between tokens from different temporal frames. Finally, the output token from this temporal encoder helps with the overall video classification. This base network is used for either social event or affect classification.

5.3.2 Multi-Task Learning Variations

We were interested in utilizing affect information and understanding the impact on social event prediction. To this end, we jointly trained the base network on affect and event labels using multi-task learning. We modify our base architecture by applying two strategies, as discussed in the following sub-sections and depicted in Fig. 5.2 b. and c.

Strategy 1: In this architecture 'ViTR-AR', illustrated in Fig. 5.2 b., two modifications were made to the base network. First, the output class token was used to predict affect using single layer perceptron (affect head). Second, the mean of output frame tokens was taken, which remained unused in our base network, and passed that to the event head, which predicts the social event. Here, affect and event classification share the base network and were trained in parallel, there is no direct dependency of affect on event or vice versa. Affect acts as a regularizer via loss function.

Strategy 2: In strategy 1, event and affect classification are done by separate classification heads and one does not directly depend on other. In strategy 2 ('ViTR-AD'), affect information is used in a hierarchical fashion for social event classification. Here we use a convolution network and classifier as an extra output network. Convolution classifier is a simple network, which takes two inputs (affect class and output frame tokens), and predicts event class as seen in Fig. 5.2 c. While training, we use the ground truth affect class as a feature to train the social event class. At the time of inference, we first predict the affect class and concatenate the predicted affect class along with output frame tokens to the convolution classifier network. Here, the predicted event directly depends on the predicted affect.

The cross-entropy loss function used in the experiments[1] for a given input x (unnormalized raw value i.e. logits), ground truth target class index y, number of classes C, is

[1]https://pytorch.org/docs/stable/generated/torch.nn.CrossEntropyLoss.html

given as:

$$L_(x, y) = -\log\left(\frac{\exp(x_{\mathrm{y}})}{\sum_{i=1}^{C}\exp(x_{\mathrm{i}})}\right) = -x_{\mathrm{y}} + \log\left(\sum_{i=1}^{C}\exp(x_{\mathrm{i}})\right) \tag{5.1}$$

Loss function used for multi-task learning in both the strategies is chosen as a combined loss of affect and event loss calculations as below:

$$L_{total} = L_{affect} + L_{event}$$

Here L_{total} is the total loss, which is calculated as a sum of L_{affect} (cross-entropy loss function between predicted and target affect) and L_{event} (cross-entropy loss function between predicted and target social event).

5.4 Experiments

All experiments were performed on a compute cluster with P100 and V100 machines. Pre-trained models were used to initialise weights and hence no data augmentation was used. Along with our three transformer based models (ViTR, ViTR-AR and ViTR-AD)[2], we also trained and tested a convolution-only action-based model 'R(2+1)D', where we used Resnet50 as the backbone, pre-trained on Kinetics 700 [Kataoka *et al.*, 2020][3] and froze first three layers and retrained on our dataset. We also trained and tested on transformer based model 'TimeSformer' [Bertasius *et al.*, 2021], with their base version and High Resolution (HR) version of pre-trained weights. Code for this work is available at *https://github.com/aarti9/VideoSocialContext*

5.4.1 Train-Val-Test split

The dataset split for social event categories and affect categories is given in Table 5.2 and Table 5.3 respectively. A total of 2661 training (Train), 766 validation (Val) and 756 testing (Test) clips were used for both event and affect classification tasks. The split used in this work, is the same as used in the [Sharma *et al.*, 2021]. It is to be noted that there is no overlap of video clips among these splits after segmentation.

[2] Backbone code and pre-trained weights were used from https://github.com/Alibaba-MIIL/STAM

[3] The R(2+1)D backbone code and pre-trained weights was used from https://github.com/kenshohara/3D-ResNets-PyTorch

Table 5.2: Dataset split for social event categories.

Category	Train	Val	Test	Total
Quarrelling	608	119	112	839
Meeting	310	206	42	558
Sports	55	49	11	115
Funeral	142	3	55	201
Show	356	124	194	674
Group Activities	305	81	47	433
Protest	330	54	46	430
Casual FFG	224	62	108	394
Celebration Party	176	48	30	253
Fighting	155	20	111	286
Total	2661	766	756	4183

Table 5.3: Dataset split for affect categories.

Category	Train	Val	Test	Total
Positive	802	302	217	1321
Neutral	923	280	309	1512
Negative	936	184	230	1350
Total	2661	766	756	4183

For all the experiments, training was done taking only Train data, and then Train and Val data combined to see the effect of data increase. For the experiments done on Train + Val data, same hyperparameters were used as in the case of Train only. Test accuracy is reported on Test data only.

5.4.2 Experimental Setup

For the experiments, we used features from 16 frames per clip, each frame being of dimensions 224 x 224 in color after normalization. This was based on [Sharma *et al.*, 2021], where it was found that out of 8, 16, 32 frames, 16 was the optimal one for affect prediction. Also, our backbone pre-trained model is STAM [Sharir *et al.*, 2021], which has 16 and 64 frames as best performing ones. The loss function was cross-entropy, batch size was 2 (due to memory constraints), optimizer was Stochastic Gradient Descent, with a learning rate of 0.0002, momentum was 0.9 and weight decay was 1e-9. Each model was trained for 20 epochs. The model parameters were 119M.

5.5 Results and Discussion

5.5.1 Quantitative Results

The experimental evaluations for the social event models are presented in Table 5.4. The top-1 test accuracy achieved for the R(2+1)D model was the lowest amongst all models in comparison, reaching around 30.42% when trained on Train + Val dataset. The two TimeSformer models performed better than R(2+1)D, achieving up to 47.48% accuracy. The ViTR models performed better than both R(2+1)D and TimeSformer models, when trained on Train dataset. The base ViTR model was comparable to the base TimeSformer model when trained on Train + Val set, but ViTR-AR and ViTR-AD outperformed both R(2+1)D and TimeSformer, achieving 48.15% and 50.13% accuracy respectively. Within our proposed models, ViTR-AR and ViTR-AD performed better than the base ViTR overall and ViTR-AD had the highest accuracy amongst all models tested. The improvements in accuracy hint a possibility of affective knowledge being helpful for social event learning. The affect information in ViTR-AR model helps regularize training, making it more stable. Overall, transformer-based models performed better than the convolution-only model, as transformer networks are able to capture longer-range temporal information along with local spatial information.

Table 5.4: Experimental results for Social event prediction.

Model	**Train**		**Train + Val**
	Val Acc.	**Test Acc.**	**Test Acc.**
R(2+1)D [Kataoka *et al.*, 2020]	54.18	25.40	30.42
TimeSformer [Bertasius *et al.*, 2021]	56.92	42.46	45.63
TimeSformerHR [Bertasius *et al.*, 2021]	56.79	40.74	47.48
ViTR	61.23	44.97	45.50
ViTR-AR	60.97	46.56	48.15
ViTR-AD	56.92	47.09	**50.13**

The classwise accuracies for models trained on Train + Val set, can be found in Table 5.5 and the confusion matrices for the three models on social event prediction can be found in Fig. 5.3. Some observations are noted below:

Table 5.5: Classwise accuracies (in %) for social event prediction.
Here, TS = TimeSformer, TSHR = TimeSformerHR, Q = Quarrelling, M = Meeting, SP = Sports, FL = Funeral, SH = Show, GA = Group Activities, P = Protest, FFG = Casual-FFG, C = Celebration Party, F = Fighting.

Model	Q	M	SP	FL	SH	GA	P	FFG	C	F	Overall
R(2+1)D	29.46	66.67	-	16.36	54.12	61.70	58.69	7.41	-	-	30.42
TS	36.61	73.81	90.91	12.73	54.64	51.06	43.48	62.96	50	20.72	45.63
TSHR	45.53	59.52	72.72	16.36	57.21	48.93	45.65	54.62	36.67	36.93	47.48
ViTR	23.21	80.95	27.27	30.91	60.31	46.81	45.65	50.93	66.67	26.13	45.50
ViTR-AR	26.79	78.57	27.27	23.64	69.07	48.94	45.65	49.07	53.33	34.23	48.15
ViTR-AD	25.00	76.19	18.18	27.27	65.98	48.94	34.78	61.11	63.33	45.05	50.13

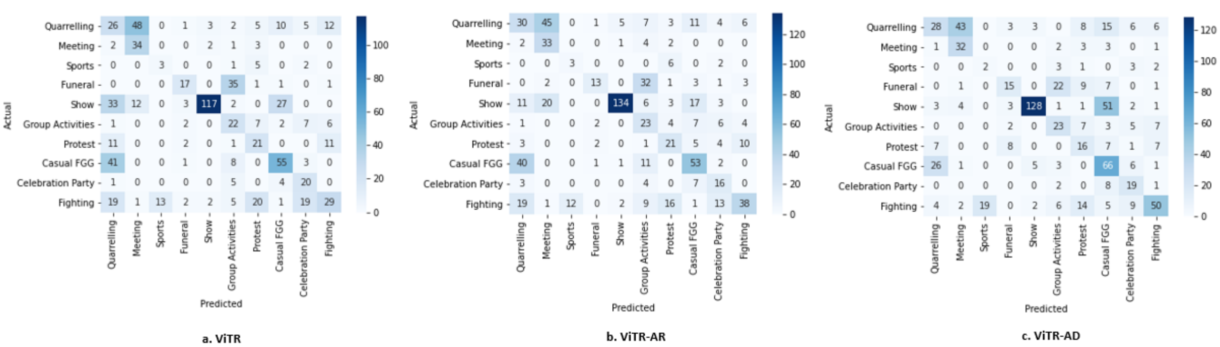


Figure 5.3: Confusion Matrix for social event prediction for all three ViTR models.

- ViTR-AR and ViTR-AD models have better overall accuracy for social event prediction than the base ViTR model. Using affect to predict event helped the model perform better. Table 5.5 shows that the ViTR-AR model has high classwise accuracy compared to ViTR-AD, possibly due to the latter using inferred affect (which may not be exact as the ground truth) to predict event, while the former uses affect in regularizing only.

- There seems confusion in the models for *'Quarrelling'* and *'Meeting'* as many labeled as quarrelling are predicted as meeting. This may be due to single modality of visual input being considered. Incorporating language and audio modalities may help the models to align better. Then again, meetings are often not that different from quarrelling!

- For the *'Sports'* category, due to minimal data, the variability in the prediction accuracy becomes very high across all the models.

- The challenges in distinguishing some categories like *'Quarrelling'* with *'Protest'* and *'Fighting'* is seen in the confusion matrix. This might be due to the similarity in these classes.

Table 5.6: Experimental results for Affect prediction.[3]

Model	**Train**		**Train + Val**
	Val. Acc.	**Test Acc.**	**Test Acc.**
LSTM-based [Sharma *et al.*, 2019]	51.43	44.84	-
VGAFNet [Sharma *et al.*, 2021]	59.00	53.83	57.01
ViTR	61.88	63.10	61.91
ViTR-AR	-	62.57	**63.23**
ViTR-AD	-	60.85	58.73

[3] ViTR-AR and ViTR-AD results use inferences from the same jointly trained model from Table 5.4.

We also tested the ViTR model on affect prediction (Table 5.6) and achieved top-1 test accuracy of 63.23% for ViTR-AR model when trained on Train + Val set. The state of art on this dataset, using full frame features (same as ours), is 53.83% and 57.01% on Train and Train + Val dataset respectively. ViTR models show improvements and provide a new state-of-the-art for group affect prediction on VGAF.

5.5.2 Qualitative Results

In Fig. 5.7, we show a few clips of correct and incorrect predictions made by ViTR-AR and ViTR-AD models. As seen, the first clip shows a casual friends gathering and the second clip is a gym class where people are doing a group activity. Both were predicted correctly. The third video clip has actual label as *'Fighting'*, but predicted label was *'Sports'*. The

Figure 5.4: Actual and predicted social event label - *Casual Friends Family Gathering*. Actual and predicted affect label - *Positive*

Figure 5.5: Actual and predicted social event label - *Group Activities*. Actual and predicted affect label - *Neutral*

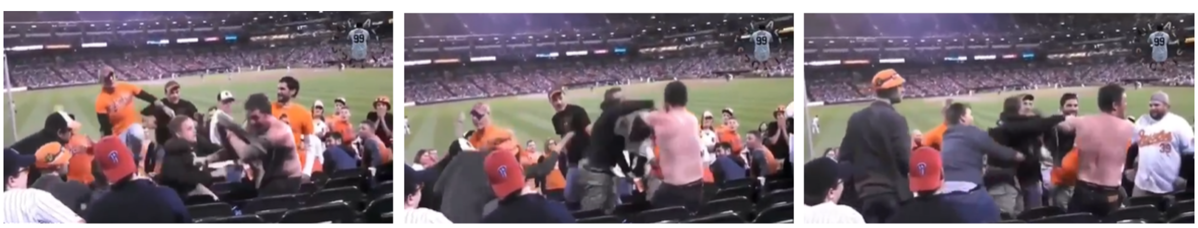


Figure 5.6: Actual social event label - *Fighting*, predicted social event label - *Sports*. Actual and predicted affect label - *Negative*

Figure 5.7: Social event and affect prediction for some clips using ViTR-AR and ViTR-AD models. (a) and (b) are correct social event predictions, (c) is an incorrect social event prediction. All three clips have correct affect prediction.

scene shows a sports ground in the background, but in the forefront, a few people are fighting with each other. It is challenging if there are mixed contexts in the scene, one could label it either ways, based on what one attends to whilst one predicts. All three clips were correctly predicted for affect label, i.e. *'Positive'*, *'Neutral'*, *'Negative'* respectively.

5.5.3 Limitations

The size of the dataset is small and could benefit from increasing the class-wise video clips and also making it more balanced. The unconstrained nature of the videos do help in making models less biased, but also poses challenges with noise in the background. Also, collecting data for crowd analysis research is a challenge due to ethical concerns. When it comes to social context and affect, there is so much variation in perception and experience, that it is challenging for AI to take into account all the factors. More in-depth research into understanding human factors contributing to the scene analysis will help in advancing the research field. This work could be useful for robots in group settings with humans, devices for people with vision issues, automatic video captioning or annotation, etc. Depending on the devices used, computational processing constraints may become a limitation for some on-device training. Although our main goal is to jointly learn and predict event and affect, we have provided single task results as well for our methods. From methods perspective, we provided comparison to convolution-only and another transformer method, but there could be other networks which can be trained on this dataset.

5.6 Ethical Impact

All the videos of the dataset have creative commons license and do not contain any personally identifiable information. The findings of this work have no explicit potential negative applications. However, use of the proposed method may lead to biased predictions as the dataset is small and some categories may be insufficiently balanced. The dataset includes videos of different social events from varying contexts in the wild. Detection of a social event and perceived affect may be misused to increase negative impacts to society. For example, if a system detects protests, some political parties may manipulate the scene for their benefit. Some scenes maybe difficult to classify in one category, which may result in misclassification.

5.7 Conclusion and Future Work

In this work, we introduced ten social event category labels for the VGAF dataset and proposed spatio-temporal transformer based networks for video classification to predict social event context and perceived group affect. The experimental results show that transformers perform better than pure convolution-based models for social event and affect prediction on the VGAF dataset. We performed multi-task learning to jointly predict social event and group affect and showed that affect prediction improves the social event prediction accuracy. Although the models show promise on the current VGAF dataset, increasing data for training by augmentation may help. We only considered holistic visual features in this work. In future work, some more cues like speech, language, sound, location information, objects in the scene may be helpful to improve the models. We believe that detecting social event context in videos is a step towards overall scene understanding for artificial intelligent agents, making them socially intelligent.

Chapter 6

Application

"Robotics and other combinations will make the world pretty fantastic compared with today."

- Bill Gates

A Socially Intelligent Agent can take form of a computer system, vitual agent or a robot. It can be used for various applications and domains such as healthcare, finance, education, retail etc. A unique quality of a human being is to be able to adapt and switch between different contexts and yet perform tasks efficiently. There is still lot of research to be done for an SIA to have such a quality. For now, in each of the applications that an SIA is developed for, there is a need to carefully design the agent such that it understands the context, identities at play and appropriate decisions and behaviors for that situation. ACT and its variations of BayesACT and NeuroACT provide such a model useful in prescribing appropriate behavior, while inhibiting inappropriate behaviors. In this chapter, we present a study of an SIA application in the form of a social robot that can be used as a socially assistive agent in a healthcare setting.

6.1 ACT Simulation for Social Robot

Socially assistive robots are devices designed to aid users through social interaction and companionship. Social robotics promise to support cognitive health and aging in place for older adults with and without dementia, as well as their care partners. However, while new and more advanced social robots are entering the commercial market, there are still major barriers to their adoption, including a lack of emotional alignment between users and their robots. ACT allows for the computational modeling of emotional alignment between the user and the socially assistive robot to be interpreted using a single, shared framework.

As part of an ACT simulation study for social robots, I collaborated with researchers at University of British Columbia, who conducted a Canadian online survey [1] , capturing attitudes, emotions, and perspectives surrounding pet-like robots among older adults (n=171), care partners (n=28), and persons living with dementia (n=7) [Dosso *et al.*, 2022a,b]. Still images of three commercially available pet-like social robots (Fig. 6.1) and a short description were presented in the online survey: Sony's AIBO [Pransky, 2001], Hasbro's Joy for All Cat, also called JustoCat [Brecher, 2020], and MiRo-E [Prescott *et al.*, 2017]. AIBO is a dog-like social robot. JustoCat is cat-like social robot with realistic looking fur and facial expressions. MiRo-E does not resemble any one animal but has features from many different animals. Participants were asked to rate these robots and themselves on a scale from -10 to +10 on the three ACT dimensions of EPA. These ratings were scaled to be consistent with EPA standards (-4.3 to 4.3). Overall study design, ethics approvals, recruitment for survey participants, data analysis for aspects other than ACT was done by other collaborators. My contribution was primarily around ACT modelling for a hypothetical scenario of *'robot assists person'*, deflection calculations for each participant, using the participant's sentiment ratings for self (the person) and each of the robot's identity. I was also involved in data analysis related to sentiments/ACT, writing about ACT and review of the manuscript.

Figure 6.1: Robots used in the study. (a) AIBO; (b) Joy for All Cat; (c) MiRo-E.

[1]This work was conducted in accordance with the Declaration of Helsinki and was approved by the University of British Columbia Behavioral Research Ethics Board (approval number H19-03308). Participants provided written informed consent before accessing the online survey.

6.2 ACT modelling

6.2.1 EPA profiles

To model the emotional relationship between person and robot, using ACT, we created sentiment profiles for each of the three socially assistive pet-like robots (Fig. 6.2). A repeated-measures ANOVA was conducted with Robot (AIBO, JustoCat, MiRo) and Sentiment (Evaluation, Potency, Activity) as within-subjects factors, Group (Care Partner, Healthy Older Adult, Person Living with Dementia) as a between-subjects factor, and sentiment scores as the dependent variable. We found a significant main effect of Sentiment (F(2, 234)=3.76, p=.025, η^2 = .031), with Evaluation (E) scores (i.e. "goodness") being the highest-rated among EPA. We also found a Robot x Sentiment interaction ($F(4, 468) = 13.65, p < .001, \eta^2 = .102$); the robots had significantly different sentiment profiles from one another, suggesting that participants viewed them as unique identities. **JustoCat was considered the most positive amongst the three robots.** There were no other main effects nor interactions; notably, participants did not rate the robots in a way that was significantly different across groups.

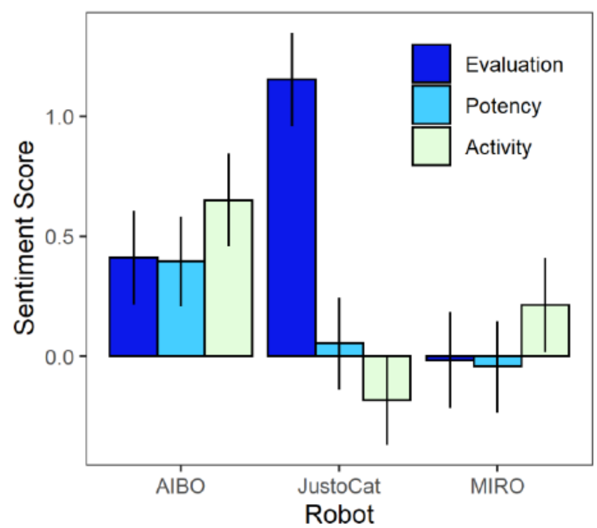


Figure 6.2: Sentiments associated with the three socially assistive robots. Means plus standard errors are shown.

6.2.2 EPA distance and robot judgement

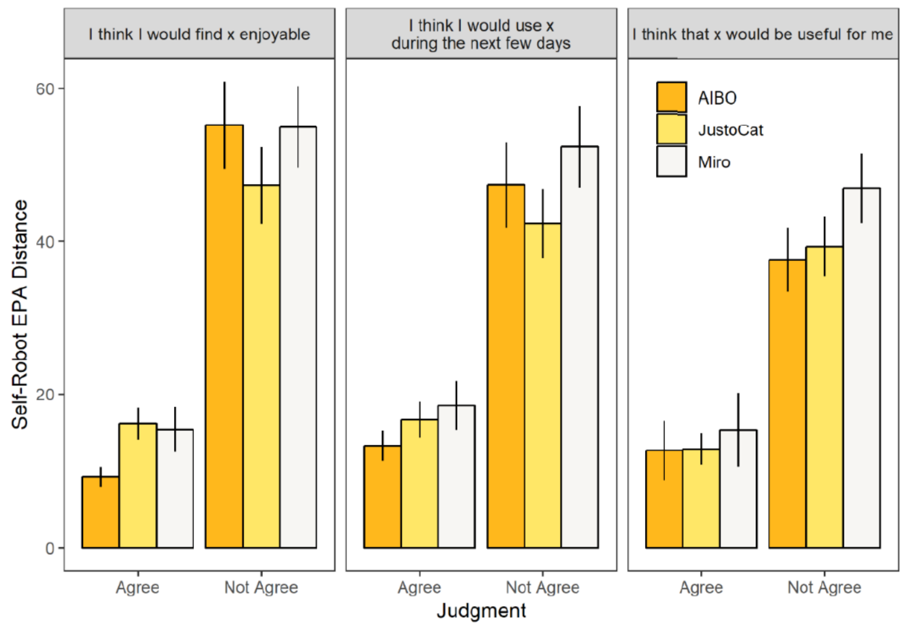


Figure 6.3: Relationship between responses to three statements about the social robots (I would enjoy. . . , I would use. . . , I would find useful. . .) and respondents' self-robot EPA distance. Means and standard errors are shown.

We calculated the EPA distance for each rater from the identity associated with each robot (R_e, R_p, R_a) to their rating of their self identity (SI_e, SI_p, SI_a), by using the formula

$$EPA\,distance = (R_e - SI_e)^2 + (R_p - SI_p)^2 + (R_a - SI_a)^2 \tag{6.1}$$

In this formula, S is the rater's ratings of themself, R is the rater's ratings of the robot, and e, p, and a are the three ACT dimensions. A larger EPA distance indicates a greater discrepancy between the rater's identity and their perception of the social robot's identity.

We looked at the EPA distances between each rater's identity (their ratings of themselves on the three EPA dimensions) and their ratings of each robot (Fig.6.3). We asked: does EPA distance (i.e., congruency between one's own identity and one's perception of the robot) correspond to differences in one's expectations and intentions around robot use? We performed a series of linear regressions to evaluate whether self-robot EPA distance for a particular robot was a significant predictor of agreement with positive statements about that robot. For example, does a greater congruency between the rater's own identity and their perception of JustoCat's identity predict higher agreement with the statement "I think I would find JustoCat enjoyable"? We found that EPA distance significantly predicted agreement with all three statements for all three robots (all $p < .001$). EPA distance between self and robot explained between 4.9% and 27.4% of the variance in agreement with robot statements (all $R^2 > .048, R^2 < .275$). In other words, participants whose identities were similar to those of the robots were more likely to indicate that they would find the robots enjoyable, use them during the next few days, and find them useful than those whose identities were dissimilar. The highest proportions of variance explained was for the three statements "I would think I would find AIBO/JustoCat/Miro enjoyable"

6.2.3 Deflection and robot usefulness

Finally, we used ACT to model an interaction *'robot assists person'*. We used EPA values of behavior "assists" as (2.20 1.64 0.75) from 'Indiana 2002-4' dataset [Francis and Heise, 2006] and ACT's impression formation equations to calculate the deflection D as mentioned in eq. (3.3) associated with each rater and each robot, that would be predicted if the robot (actor) were to engage in the behavior of 'assisting' a person (object). Using a series of linear regressions, we found that a smaller deflection associated with the action "assists" predicted agreement with the statement "I think that [robot] would be useful for me" (refer Fig. 6.4). The sample size of the participants was not balanced for each of the participant groups, which is one of the limitations of this study.

6.3 Conclusion and Future Work

Turning to the emotion modeling data, ratings of the three robots produced three sentiment profiles that were quite different from one another, suggesting that participants viewed them as having unique identities. The calculated EPA distance between a rater's judgments of themself and their judgments of a robot was very strongly linked to an anticipation that the robot would be enjoyable and useful. This is promising evidence that the ACT measures used in this work were able to capture dimensions of participants' identities that predict their real-world behavior and experiences with robots. Similarly, when our model indicated that the concept "robot assists person" was highly congruent for a particular respondent and robot, that respondent was more likely to agree that the robot would be useful – another piece of evidence validating ACT as a promising model of human-robot emotional

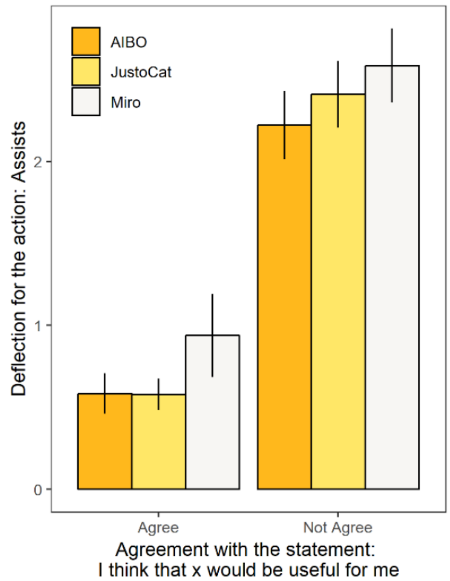


Figure 6.4: Relationship between agreeing that a robot would be useful and ACT deflection for the concept of being assisted by the robot. Means and standard errors shown.

alignment. Taken together, this work examined perceptions of social robots among a sample of older adults across a range of lived experiences. This can help in choosing an agent amongst different options available to the user. Furthermore, we demonstrated that computational modeling of emotional alignment between humans and robots was possible for this type of sample.

Future work can include same application of a socially assistive agent, but with an actual robot. Any human-agent interaction can be modelled with ACT and validated with a user study. The EPA ratings for identities, emotions, behaviors, settings can be collected from target user groups of an application, and can be used by the SIA to interact well with that group. Ethical concerns like target user privacy, user's mental state while interacting, user's background, user's expectations from the agent, user's preference for salutation/identification by the agent, etc., need to be considered when designing or deploying such applications.

Chapter 7

Conclusion

"The more socially intelligent you are, the happier and more robust and more enjoyable your relationships will be."

- Daniel Goleman

7.1 Socially Intelligent Agent

Social interaction is an important part of day-to-day life in human beings. Affect and emotions play a key role in decision-making and behavior. The nuances of an interaction between individuals can create harmony between individuals or create discord.

This thesis aims towards building Socially Intelligent Agent (SIA), that can align with humans emotionally and can be useful in domains such as healthcare, home, business, academics and many more, where AI agent needs to interact with humans, while making a meaningful connection, and providing humans with a good interaction experience. It focuses on emotions and context in decision-making for SIA. It does a systematic review of computational models of emotions in decision-making and synthesizes data into four types. It then uses one such model called Affect Control Theory (ACT), implements it in a brain-inspired neural model called NeuroACT and simulates social interaction scenarios. It builds visual perception components for SIA for social scene understanding, to infer social context and perceived affect, with an aim to behave in an emotionally aligned and coherent manner. Lastly, it models ACT for SIA in a real-world scenario.

A next step towards building SIA (who has an identity defined), would be to use contextual information in a visual scene, detect the interacting identities and their behavior and select an appropriate behavior using ACT. There can be various extensions to each of the research contributions in this thesis. The NeuroACT can include visual processing modules once available in the library and the output of the model can be connected to some motor system of an AI agent for action to be taken. Advances in neuroscience can help in evolving the model of decision-making further. The advantage of brain-inspired Nengo models are that they can be used on a neuromorphic hardware, which are considered to be efficient in processing [DeWolf *et al.*, 2020]. The vision classification model ViTR and its variants can be improved, speech and language modalities may be experimented with if improves accuracy. It can incorporate identities and behavior prediction classifiers. Finally, ACT can be modelled into different real-world scenarios for different target groups and validation study can be conducted.

7.2 Limitations of ACT

Although ACT is a strong framework for social interaction, the datasets for EPA ratings are average population ratings based on survey data collected from people from few different countries in the past. The affective ratings of social identities are highly dependent on the dataset used (population surveyed) and they are averages in ACT, making any bi-modal distributions give highly skewed answers. BayesACT [Hoey *et al.*, 2016; Schröder *et al.*, 2016; Hoey *et al.*, 2021] would potentially solve both these problems by (1) including a learning mechanism for social identities as a free energy minimization process; and (2) allowing affective meanings to be encoded as probability distributions, not as points. As times change, the society structure, social norms, meaning of concepts also evolve. An application using ACT model needs to be able to evolve with time. The ratings for each concept in the EPA dictionaries are text based only. An attempt is made in this thesis to have visual classifier to infer social context and perceived affect in a scene. We still need to convert it into text form to be able to use ACT grammar. Inference for identities and behavior need to be done for ACT interaction to work for SIA.

7.3 Ethical Impact of SIAs

Developing Socially Intelligent Agents can benefit society by incorporating more AI agents for social good, in assistive and adaptive way. But it also brings forth potential of misuse by humans for humans. Social engineering [Postnikoff, 2020] is one form of attack where human data can be collected and used through agents that pretend to be some believable identity. SIAs can be used to manipulate humans in social interaction. They can be programmed to take decisions that may not adhere to the social norms and may not be moral and ethical for society. As AI adoption increases, we need to be aware of the risks and ensure appropriate steps are taken to safeguard privacy and safety of everyone. Some design considerations for SIA are discussed in [Malhotra and Hoey, 2021].

Despite of the risk of misuse of SIA, building agents that are socially intelligent would be less dangerous than building agents that are only rational. A rational agent here means a non socially intelligent agent, even if it incorporates altruism as utility function, that penalizes the agent for being selfish. It may be challenging to quantify the amount of altruism that an agent should have, and also tricky to know who to be altruistic and who not to be with. As an illustration, in Prisoner's dilemma game, the optimal action for a rational agent is to defect, but a socially intelligent agent will consider the social identity of self and the other, say a 'friend' or a 'stranger', and take an appropriate action. This may lead to better cooperation. In a real-world social interaction, an SIA can be developed to not only understand the task at hand, but also understand the context, emotions of the humans, social norms and prescriptions, and behave appropriately in any situation, while learning and adapting to the environment.

7.4 Closing Thoughts

Designing and developing SIA that is emotionally aligned and adaptive with humans, while understanding the context in any situation, and adhering to social norms, can benefit society at large. The research presented in this thesis is a step towards that direction. Role of emotions and context are considered in decision-making by using ACT model for SIA. The insights gained in this thesis may encourage AI and affective computing research to develop agents that can simulate human affective and decision-making mechanisms, and in the process understand humans better.

References

A. Sharpanskykh and J. Treur. An adaptive affective social decision making model. In *Biologically Inspired Cognitive Architectures 2012*, pages 299–308. Springer, 2013.

A. Sharpanskykh and K. Zia. Emotional decision making in large crowds. In *Advances on Practical Applications of Agents and Multi-Agent Systems*, pages 191–200. Springer, 2012.

A. Tripathi, T. Ashwin, and R. M. R. Guddeti. A reinforcement learning and recurrent neural network based dynamic user modeling system. In *2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT)*, pages 411–415. IEEE, 2018.

K. J. Åström. Optimal Control of Markov Decision Processes with Incomplete State Estimation. *J. Math. Anal. App.*, 10:174—-205, 1965.

Asad Abbas and Stephan K Chalup. Group emotion recognition in the wild by combining deep neural networks for facial expression classification and scene-context analysis. In *ACM International Conference on Multimodal Interaction*, pages 561–568, 2017.

Kashif Ahmad, Nicola Conci, Giulia Boato, and Francesco GB De Natale. Used: a large-scale social event detection dataset. In *Proceedings of the 7th International Conference on Multimedia Systems*, pages 1–6, 2016.

Hyungil Ahn and Rosalind W Picard. Affective-cognitive learning and decision making: A motivational reward framework for affective agents. In *International conference on affective computing and intelligent interaction*, pages 866–873. Springer, 2005.

J Wayne Aldridge, Kent C Berridge, Mark Herman, and Lee Zimmer. Neuronal coding of serial order: syntax of grooming in the neostriatum. *Psychological Science*, 4(6):391–395, 1993.

Bushra Alhijawi. Improving collaborative filtering recommender system results and performance using satisfaction degree and emotions of users. In *Web Intelligence*, volume 17, pages 229–241. IOS Press, 2019.

Pedro Álvarez, A Guiu, José Ramón Beltrán, J García de Quirós, and Sandra Baldassarri. DJ-Running: An Emotion-based System for Recommending Spotify Songs to Runners. In *icSPORTS*, pages 55–63, 2019.

Peter Anderson, Xiaodong He, Chris Buehler, Damien Teney, Mark Johnson, Stephen Gould, and Lei Zhang. Bottom-Up and Top-Down Attention for Image Captioning and Visual Question Answering. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6077–6086, 2018.

Eduardo B Andrade and Dan Ariely. The enduring impact of transient emotions on decision making. *Organizational Behavior and Human Decision Processes*, 109(1):1–8, 2009.

Dimitrios Antos and Avi Pfeffer. Using emotions to enhance decision-making. In *Twenty-Second International Joint Conference on Artificial Intelligence*, 2011.

Anurag Arnab, Mostafa Dehghani, Georg Heigold, Chen Sun, Mario Lučić, and Cordelia Schmid. Vivit: A video vision transformer. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 6836–6846, 2021.

Nabiha Asghar and Jesse Hoey. Intelligent Affect: Rational Decision Making for Socially Aligned Agents. In *UAI*, pages 72–81, 2015.

Elahe Bagheri, Oliver Roesler, Hoang-Long Cao, and Bram Vanderborght. A reinforcement learning based cognitive empathy framework for social robots. *International Journal of Social Robotics*, 13:1079–1093, 2021.

Michela Balconi and Maria E Vanutelli. Competition in the brain. The contribution of EEG and fNIRS modulation and personality effects in social ranking. *Frontiers in psychology*, 7:1587, 2016.

Lisa Feldman Barrett and Ajay B. Satpute. Large-scale brain networks in affective and social neuroscience: towards an integrative functional architecture of the brain. *Current Opinion in Neurobiology*, 23:361–372, 2013.

Lisa Feldman Barrett and Ajay B. Satpute. Historical pitfalls and new directions in the neuroscience of emotion. *Neuroscience Letters*, 693:9–18, 2019.

Lisa Barrett, Batja Mesquita, and Maria Gendron. Context in Emotion Perception. *Current Directions in Psychological Science*, 20:286–290, 10 2011.

Lisa Feldman Barrett, Ralph Adolphs, Stacy Marsella, Aleix M Martinez, and Seth D Pollak. Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological science in the public interest*, 20(1):1–68, 2019.

Lisa Feldman Barrett. The theory of constructed emotion: an active inference account of interoception and categorization. In *Social cognitive and affective neuroscience*, 2016.

Lisa Feldman Barrett. *How emotions are made: The secret life of the brain*. Houghton Mifflin Harcourt, 2017.

Sigal G Barsade and Donald E Gibson. Group emotion: A view from top and bottom. In *D. H Gruenfeld (Ed.), Research on managing groups and teams, Vol. 1. Composition*, pages 81–102, US, 1998. Elsevier Science/JAI Press.

Matthew Barthet, Antonios Liapis, and Georgios N Yannakakis. Go-Blend behavior and affect. In *2021 9th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*, pages 1–8. IEEE, 2021.

Matthew Barthet, Ahmed Khalifa, Antonios Liapis, and Georgios N Yannakakis. Play with Emotion: Affect-Driven Reinforcement Learning. In *2022 10th International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 1–8. IEEE, 2022.

Yoann Baveye, Emmanuel Dellandrea, Christel Chamaret, and Liming Chen. LIRIS-ACCEDE: A video database for affective content analysis. *IEEE Transactions on Affective Computing*, 6(1):43–55, 2015.

Antoine Bechara, Hanna Damasio, and Antonio R. Damasio. Emotion, decision making and the orbitofrontal cortex. *Cerebral cortex*, 10 3:295–307, 2000.

Christian Becker, Stefan Kopp, and Ipke Wachsmuth. Why emotions should be integrated into conversational agents. *Conversational informatics: an engineering approach*, pages 49–68, 2007.

Trevor Bekolay, James Bergstra, Eric Hunsberger, Travis DeWolf, Terrence C Stewart, Daniel Rasmussen, Xuan Choo, Aaron Russell Voelker, and Chris Eliasmith. Nengo: a Python tool for building large-scale functional brain models. *Frontiers in neuroinformatics*, 7:48, 2014.

Mouna Belhaj, Fahem Kebair, and Lamjed Ben Said. Modeling and simulation of coping mechanisms and emotional behavior during emergency situations. In *Agent and Multi-Agent Systems: Technology and Applications*, pages 163–176. Springer, 2016.

Marwen Belkaid, Nicolas Cuperlier, and Philippe Gaussier. Emotional metacontrol of attention: Top-down modulation of sensorimotor processes in a robotic visual search task. *PloS one*, 12(9):e0184960, 2017.

Peter L. Berger and Thomas Luckmann. *The Social Construction of Reality*. Doubleday: Anchor Books, 1967.

Gedas Bertasius, Heng Wang, and Lorenzo Torresani. Is Space-Time Attention All You Need for Video Understanding? *arXiv preprint arXiv:2102.05095*, 2021.

Beatrice Biancardi, Chen Wang, Maurizio Mancini, Angelo Cafaro, Guillaume Chanel, and Catherine Pelachaud. A Computational Model for Managing Impressions of an Embodied Conversational Agent in Real-Time. In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 1–7. IEEE, 2019.

Lukas Bossard, Matthieu Guillaumin, and Luc Van Gool. Event recognition in photo collections with a stopwatch HMM. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1193–1200, 2013.

David B Brecher. Use of a robotic cat to treat terminal restlessness: a case study. *Journal of palliative medicine*, 23(3):432–434, 2020.

Joost Broekens, Walter A Kosters, and Fons J Verbeek. On affect and self-adaptation: Potential benefits of valence-controlled action-selection. In *International Work-Conference on the Interplay Between Natural and Artificial Computation*, pages 357–366. Springer, 2007.

C. Liu, K. Conn, N. Sarkar, and W. Stone. Online affect detection and adaptation in robot assisted rehabilitation for children with autism. In *RO-MAN 2007-The 16th IEEE International Symposium on Robot and Human Interactive Communication*, pages 588–593. IEEE, 2007.

C. P. Lee-Johnson and D. A. Carnegie. Emotion-based parameter modulation for a hierarchical mobile robot planning and control architecture. In *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2839–2844. IEEE, 2007.

Daniel Cabrera-Paniagua, César Azola, and Rolando Rubilar-Torrealba. Using affective criteria in the decision-making of cognitive agents on flexible passenger transportation domain. *Journal of Ambient Intelligence and Humanized Computing*, 14(8):10715–10729, 2023.

Zoraida Callejas, Brian Ravenet, Magalie Ochs, and Catherine Pelachaud. A computational model of social attitudes for a virtual recruiter. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, pages 93–100, 2014.

Álvaro Castro-González, María Malfaz, and Miguel Angel Salichs. An autonomous social robot in fear. *IEEE Transactions on Autonomous Mental Development*, 5(2):135–151, 2013.

Feng Chao, Chen Lin, Jiang Kui, Wei Zhonglin, and Zhai Bing. Research on decision-making in emotional agent based on reinforcement learning. In *2016 2nd IEEE International Conference on Computer and Communications (ICCC)*, pages 1191–1194. IEEE, 2016.

Jize Chen and Changhong Wang. Reaching cooperation using emerging empathy and counter-empathy. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, pages 746–753, 2019.

Yunpeng Chen, Yannis Kalantidis, Jianshu Li, Shuicheng Yan, and Jiashi Feng. A2-Nets: Double Attention Networks. *arXiv preprint arXiv:1810.11579*, 2018.

Ensi Chen, Huan Zhao, Bo Li, Xupeng Zha, Haoqian Wang, and Song Wang. Affective feature knowledge interaction for empathetic conversation generation. *Connection Science*, 34(1):2559–2576, 2022.

Jize Chen, Dali Zhang, Zhenshen Qu, and Changhong Wang. Modeling adaptive empathy based on neutral assessment: a way to enhance the prosocial behaviors of socialized agents under the premise of self-security. *Applied Intelligence*, 52(6):6692–6722, 2022.

Davide Conigliaro, Paolo Rota, Francesco Setti, Chiara Bassetti, Nicola Conci, Nicu Sebe, and Marco Cristani. The s-hock dataset: Analyzing crowds at the stadium. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2039–2047, 2015.

Karla Conn, Changchun Liu, Nilanjan Sarkar, Wendy Stone, and Zachary Warren. Affect-sensitive assistive intervention technologies for children with autism: An individual-specific approach. In *RO-MAN 2008-The 17th IEEE International Symposium on Robot and Human Interactive Communication*, pages 442–447. IEEE, 2008.

Mihai Gabriel Constantin, Liviu Daniel Stefan, Bogdan Ionescu, Claire-Helene Demarty, Mats Sjoberg, Markus Schedl, and Guillaume Gravier. Affect in Multimedia: Benchmarking Violent Scenes Detection. *IEEE Transactions on Affective Computing*, pages 1–1, 2020.

Evren Daglarli, Hakan Temeltas, and Murat Yesiloglu. Behavioral task processing for cognitive robots using artificial emotions. *Neurocomputing*, 72(13-15):2835–2844, 2009.

Antonio R Damasio. *Descartes' error: emotion, reason, and the human brain.* G.P. Putnam, 1994.

RJ Davidson, RD Lane, and L Nadel. Cognitive neuroscience of emotion. 2000.

Berardina De Carolis, Stefano Ferilli, and Giuseppe Palestra. Simulating empathic behavior in a social assistive robot. *Multimedia Tools and Applications*, 76(4):5073–5094, 2017.

Kerstin Denecke, Sayan Vaaheesan, and Aaganya Arulnathan. A mental health chatbot for regulating emotions (SERMO)-concept and usability test. *IEEE Transactions on Emerging Topics in Computing*, 9(3):1170–1182, 2020.

Travis DeWolf, Pawel Jaworski, and Chris Eliasmith. Nengo and low-power ai hardware for robust, embedded neurorobotics. *Frontiers in Neurorobotics*, 14:568359, 2020.

A. Dhall, J. Joshi, I. Radwan, and R. Goecke. Finding happiest moments in a social context, booktitle = Proceedings of the 11th Asian Conference on Computer Vision. pages 613—-626, 2012.

A. Dhall, J. Joshi, K. Sikka, R. Goecke, and N. Sebe. The more the merrier: Analysing the affect of a group of people in images. In *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, volume 1, pages 1–8, May 2015.

Abhinav Dhall, Roland Goecke, and Tamás D. Gedeon. Automatic Group Happiness Intensity Analysis. *IEEE Transactions on Affective Computing*, 6:13–26, 2015.

Toby Donaldson, Andrew Park, I Lin, et al. Emotional pathfinding. In *Conference of the Canadian Society for Computational Studies of Intelligence*, pages 31–43. Springer, 2004.

Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain

Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.

Jill A Dosso, Ela Bandari, Aarti Malhotra, Gabriella K Guerra, Jesse Hoey, François Michaud, Tony J Prescott, and Julie M Robillard. User perspectives on emotionally aligned social robots for older adults and persons living with dementia. *Journal of Rehabilitation and Assistive Technologies Engineering*, 9:20556683221108364, 2022. PMID: 35782883.

Jill A Dosso, Ela Bandari, Aarti Malhotra, Jesse Hoey, Francois Michaud, Tony J Prescott, and Julie M. Robillard. Towards emotionally aligned social robots for dementia: perspectives of care partners and persons with dementia. *Alzheimer's & Dementia*, 18(S2):e059261, 2022.

Rob Duell and Jan Treur. A computational analysis of joint decision making processes. In *International Conference on Social Informatics*, pages 292–308. Springer, 2012.

Guillaume Dumas, Jacqueline Nadel, Robert Soussignan, Jacques Martinerie, and Line Garnero. Inter-brain synchronization during social interaction. *PloS one*, 5(8):e12166, 2010.

Chris Eliasmith and Charles Hammond Anderson. Neural Engineering: Computation, Representation, and Dynamics in Neurobiological Systems. *IEEE Transactions on Neural Networks*, 15:528–529, 2004.

Chris Eliasmith, Terrence C. Stewart, Xuan Choo, Trevor Bekolay, Travis DeWolf, Yichuan Tang, and Daniel Rasmussen. A Large-Scale Model of the Functioning Brain. *Science*, 338(6111):1202–1205, 2012.

Chris Eliasmith. *How to build a brain: A neural architecture for biological cognition*. OUP USA, 2013.

Natascha Esau, Lisa Kleinjohann, and Bernd Kleinjohann. Emotional communication with the robot head MEXI. In *2006 9th International Conference on Control, Automation, Robotics and Vision*, pages 1–7. IEEE, 2006.

Usef Faghihi, Pierre Poirier, Daniel Dubois, and Roger Nkambou. How emotional mechanism learn and helps other types of learning in a cognitive agent. In *2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, volume 3, pages 606–611. IEEE, 2008.

Usef Faghihi, Philippe Fournier-Viger, Roger Nkambou, Pierre Poirier, and André Mayers. How emotional mechanism helps episodic learning in a cognitive agent. In *2009 IEEE Symposium on Intelligent Agents*, pages 23–30. IEEE, 2009.

Usef Faghihi, Pierre Poirier, Philippe Fournier-Viger, and Roger Nkambou. Human-like learning in a conscious agent. *Journal of Experimental & Theoretical Artificial Intelligence*, 23(4):497–528, 2011.

Johannes Feldmaier and Klaus Diepold. Path-finding using reinforcement learning and affective states. In *The 23rd IEEE international symposium on robot and human interactive communication*, pages 543–548. IEEE, 2014.

Mallory J Feldman, Erika Siegel, Lisa Feldman Barrett, Karen S Quigley, and Jolie B Wormwood. Affect and social judgment: the roles of physiological reactivity and interoceptive sensitivity. *Affective Science*, 3(2):464–479, 2022.

Maurizio Ficocelli, Junichi Terao, and Goldie Nejat. Promoting interactions between humans and robots using robotic emotional behavior. *IEEE transactions on cybernetics*, 46(12):2911–2923, 2015.

Kate D Fischl, Adam B Cellon, Terrence C Stewart, Timothy K Horiuchi, and Andreas G Andreou. Socio-emotional robot with distributed multi-platform neuromorphic processing:(invited presentation). In *2019 53rd Annual Conference on Information Sciences and Systems (CISS)*, pages 1–6. IEEE, 2019.

Candy PS Fong and Robert S Wyer Jr. Cultural, social, and emotional determinants of decisions under uncertainty. *Organizational Behavior and Human Decision Processes*, 90(2):304–322, 2003.

Clare Francis and David R. Heise. Mean Affective Ratings of 1,500 Concepts by Indiana University Undergraduates in 2002-3. Computer file, Distributed at Affect Control Theory Website, Program Interact bayesact.ca, 2006.

Linda E. Francis, Richard E. Adams, Alexandra Konig, and Jesse Hoey. Identity and the Self in Elderly Adults with Alzheimer's Disease. In *Identities in Everyday Life*. Oxford University Press, 2019.

Nico H Frijda. The laws of emotion. *American psychologist*, 43(5):349, 1988.

Karl Friston. The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11((2)):127–138, 2010.

Sandra Clara Gadanho and John Hallam. Robot learning driven by emotions. *Adaptive behavior*, 9(1):42–64, 2001.

Kirill Gavrilyuk, Ryan Sanford, Mehrsan Javan, and Cees GM Snoek. Actor-transformers for group activity recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 839–848, 2020.

Rohit Girdhar, Joao Carreira, Carl Doersch, and Andrew Zisserman. Video action transformer network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 244–253, 2019.

João Gluz and Patricia Jaques. A Probabilistic Implementation of Emotional BDI Agents. In *Proceedings of the 6th International Conference on Agents and Artificial Intelligence - Volume 1*, ICAART 2014, page 121–129, Setubal, PRT, 2014. SCITEPRESS - Science and Technology Publications, Lda.

João Carlos Gluz and Patricia Augustin Jaques. A probabilistic approach to represent emotions intensity into BDI agents. In *International Conference on Agents and Artificial Intelligence*, pages 225–242. Springer, 2014.

Erving Goffman et al. *The presentation of self in everyday life*, volume 21. Harmondsworth London, 1978.

Erving Goffman. *The Presentation of Self in Everyday Life*. Anchor Books, USA, 1959.

Erving Goffman. *Behavior in Public Places*. The Free Press, New York, 1963.

Harry F. Gollob. The subject-verb-object approach to social cognition. *Psychological Review*, 81:286–321, 1974.

Gilzamir Gomes, Creto A Vidal, Joaquim B Cavalcante Neto, and Yuri LB Nogueira. An emotional virtual character: A deep learning approach with reinforcement learning. In *2019 21st Symposium on Virtual and Augmented Reality (SVR)*, pages 223–231. IEEE, 2019.

Jonathan Gratch and Stacy Marsella. A domain-independent framework for modeling emotion. *Cognitive Systems Research*, 5(4):269–306, 2004.

Jeffrey A Gray. Three fundamental emotion systems. *The nature of emotion: Fundamental questions*, 14:243–247, 1994.

Xin Guo, Bin Zhu, Luisa F Polanía, Charles Boncelet, and Kenneth E Barner. Group-level emotion recognition using hybrid deep models based on faces, scenes, skeletons and visual attentions. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction*, pages 635–639, 2018.

Daya S Gupta and Hugo Merchant. Understanding the Role of the Time Dimension in the Brain Information Processing. *Frontiers in Psychology*, 8:240, 2017.

Kevin Gurney, Tony J Prescott, and Peter Redgrave. A computational model of action selection in the basal ganglia. I. A new functional anatomy. *Biological cybernetics*, 84:401–410, 2001.

H. Zhang and S. Liu. On-line affective cognitive learning and decision-making for autonomous navigation of mobile robots. In *2009 International Conference on Information and Automation*, pages 1234–1239. IEEE, 2009.

Jonathan Haidt. The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological review*, 108(4):814, 2001.

Qi Han, TA Arentze, HJP Timmermans, Davy Janssens, and D Wets. Modeling individuals' cognitive and affective responses in spatial learning behavior. In *Proceedings of the 13th international conference of Hong Kong Society for Transportation Studies*, pages 13–15, 2008.

Tal Hassner, Yossi Itcher, and Orit Kliper-Gross. Violent flows: Real-time detection of violent crowd behavior. In *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pages 1–6. IEEE, 2012.

David Heise and Neil MacKinnon. *Self, identity, and social institutions.* Springer, 2010.

David R. Heise. Affect control theory: Concepts and model. *The Journal of Mathematical Sociology*, 13(1-2):1–33, 1987.

David R Heise. *Expressive order: Confirming sentiments in social actions.* Springer US, 2007.

David R. Heise. *Surveying Cultures: Discovering Shared Conceptions and Sentiments.* John Wiley & Sons, Inc., 2010.

Joseph Henrich. *The WEIRDest people in the world.* Farrar, Strauss and Giroux, New York, 2020.

Casper Hesp, Ryan Smith, Thomas Parr, Micah Allen, Karl J Friston, and Maxwell JD Ramstead. Deeply felt affect: The emergence of valence in deep active inference. *Neural computation*, 33(2):398–446, 2021.

Jens Hoefinghoff and Josef Pauli. Reversal learning based on somatic markers. In *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, pages 498–504. IEEE, 2013.

Jens Hoefinghoff, Laura Hoffmann, Nicole Kraemer, and Josef Pauli. Deciding like humans do. In *The Twenty-Sixth International FLAIRS Conference*, 2013.

Erik Hoelzl and George Loewenstein. Wearing out your shoes to prevent someone else from stepping into them: Anticipated regret and social takeover in sequential decisions. *Organizational Behavior and Human Decision Processes*, 98(1):15–27, 2005.

Jesse Hoey, Tobias Schröder, and Areej Alhothali. Affect Control Processes: Intelligent Affective Interaction using a Partially Observable Markov Decision Process. *Artificial Intelligence*, 230:134–172, 2016.

Jesse Hoey, Tobias Schröder, Jonathan H Morgan, Kimberly B Rogers, and Meiyappan Nagappan. Affective Dynamics and Control in Group Processes. In *Proceedings of the Group Interaction Frontiers in Technology*, pages 1–7. 2018.

Jesse Hoey, Neil J. MacKinnon, and Tobias Schröder. Denotative and Connotative Management of Uncertainty: A Computational Dual-Process Model. *Judgement and Decision Making*, 16(2), March 2021.

Eric Hogewoning, Joost Broekens, Jeroen Eggermont, and Ernst GP Bovenkamp. Strategies for affect-controlled action-selection in Soar-RL. In *International Work-Conference on the Interplay Between Natural and Artificial Computation*, pages 501–510. Springer, 2007.

Jitka Homolová, Anastasija Černecka, Tatiana V Guy, and Miroslav Kárnỳ. Affective Decision-Making in Ultimatum Game: Responder Case. In *European Conference on Multi-Agent Systems*, pages 127–139. Springer, 2018.

Mark Hoogendoorn, Jan Treur, C Wal, and Arlette van Wissen. Modelling the emergence of group decisions based on mirroring and somatic marking. In *International Conference on Brain Informatics*, pages 29–41. Springer, 2010.

Johan F Hoorn, Matthijs Pontier, and Ghazanfar F Siddiqui. When the User is Instrumental to Robot Goals: First Try–Agent Uses Agent. In *2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, volume 2, pages 296–301. IEEE, 2008.

Keiichi Horio and Atsushi Matsuda. The decision making method added emotion judgment. In *2010 World Automation Congress*, pages 1–6. IEEE, 2010.

Chu-Ren Huang, Ying Chen, and Sophia Yat Mei Lee. Textual Emotion Processing from Event Analysis. In *CIPS-SIGHAN Joint Conference on Chinese Language Processing*, 2010.

X. Huang, A. Dhall, R. Goecke, M. Pietikäinen, and G. Zhao. Multimodal Framework for Analyzing the Affect of a Group of People. *IEEE Transactions on Multimedia*, 20(10):2706–2721, Oct 2018.

Xiao Huang, Wei Wu, Hong Qiao, and Yidao Ji. Brain-inspired motion learning in recurrent neural network with emotion modulation. *IEEE Transactions on Cognitive and Developmental Systems*, 10(4):1153–1164, 2018.

Xiao Huang, Wei Wu, and Hong Qiao. Computational modeling of emotion-motivated decisions for continuous control of mobile robots. *IEEE Transactions on Cognitive and Developmental Systems*, 13(1):31–44, 2020.

Eva Hudlicka. Two sides of appraisal: Implementing appraisal and its consequences within a cognitive architecture. In *Proceedings of the AAAI spring symposium: architectures for modeling emotion. AAAI Press*, pages 24–31, 2004.

J. Benjamin Hutchinson and Lisa Feldman Barrett. The Power of Predictions: An Emerging Paradigm for Psychological Research. *Current Directions in Psychological Science*, 28:280 – 291, 2019.

J. F. Hoorn, M. Pontier, and G. F. Siddiqui. Coppélius' concoction: Similarity and complementarity among three affect-related agent models. *Cognitive Systems Research*, 15:33–49, 2012.

Adrien Jauffret, Marwen Belkaid, Nicolas Cuperlier, Philippe Gaussier, and Philippe Tarroux. Frustration as a way toward autonomy and self-improvement in robotic navigation. In *2013 IEEE Third Joint International Conference on Development and Learning and Epigenetic Robotics (ICDL)*, pages 1–7. IEEE, 2013.

Longsheng Jiang and Yue Wang. Respect your emotion: Human-multi-robot teaming based on regret decision model. In *2019 IEEE 15th International Conference on Automation Science and Engineering (CASE)*, pages 936–941. IEEE, 2019.

Yu-Gang Jiang, Guangnan Ye, Shih-Fu Chang, Daniel Ellis, and Alexander C. Loui. Consumer Video Understanding: A Benchmark Database and An Evaluation of Human and Machine Performance. In *Proceedings of ACM International Conference on Multimedia Retrieval (ICMR), oral session*, 2011.

Longsheng Jiang, Dong Chen, Zhaojian Li, and Yue Wang. Risk Representation, Perception, and Propensity in an Integrated Human Lane-Change Decision Model. *IEEE Transactions on Intelligent Transportation Systems*, 23(12):23474–23487, 2022.

Anja Johansson and Pierangelo Dell'Acqua. Affective states in behavior networks. In *Intelligent Computer Graphics 2009*, pages 19–39. Springer, 2009.

Anja Johansson and Pierangelo Dell'Acqua. Pathfinding with emotion maps. In *Intelligent Computer Graphics 2011*, pages 139–155. Springer, 2012.

A Kaklauskas, EK Zavadskas, A Banaitis, I Meidute-Kavaliauskiene, A Liberman, S Dzitac, I Ubarte, A Binkyte, J Cerkauskas, A Kuzminske, et al. A neuro-advertising property video recommendation system. *Technological Forecasting and Social Change*, 131:78–93, 2018.

Chien-Hao Kao, Chih-Chieh Chen, and Yu-Tza Tsai. Model of multi-turn dialogue in emotional chatbot. In *2019 International Conference on Technologies and Applications of Artificial Intelligence (TAAI)*, pages 1–5. IEEE, 2019.

Yuliana Karabelnikova and Alexei V Samsonovich. Virtual partner dance as a paradigm for empirical study of cognitive models of emotional intelligence. *Procedia Computer Science*, 190:414–433, 2021.

Hirokatsu Kataoka, Tenga Wakamiya, Kensho Hara, and Yutaka Satoh. Would mega-scale datasets further enhance spatiotemporal 3D CNNs? *arXiv preprint arXiv:2004.04968*, 2020.

Rahul Katarya and Om Prakash Verma. Recent developments in affective recommender systems. *Physica A: Statistical Mechanics and its Applications*, 461:182–190, 2016.

Janice R Kelly and Sigal G Barsade. Mood and emotions in small groups and work teams. *Organizational behavior and human decision processes*, 86(1):99–130, 2001.

Wasif Khan and Jesse Hoey. How Different Identities Affect Cooperation. In *Proc.of the Humaine Association Conference on Affective Computing and Intelligent Interaction*, San Antonio, TX, 2017. IEEE.

Yong-Duk Kim, Jong-Hwan Kim, and Yong-Jae Kim. Behavior selection and learning for synthetic character. In *Proceedings of the 2004 Congress on Evolutionary Computation (IEEE Cat. No. 04TH8753)*, volume 1, pages 898–903. IEEE, 2004.

Kiril Kiryazov, Robert Lowe, Christian Becker-Asano, and Marco Randazzo. The role of arousal in two-resource problem tasks for humanoid service robots. In *2013 IEEE RO-MAN*, pages 62–69. IEEE, 2013.

Barbara Kitchenham. Procedures for performing systematic reviews. *Keele, UK, Keele University*, 33(2004):1–26, 2004.

Alexandra Konig, Linda E. Francis, Aarti Malhotra, and Jesse Hoey. Defining affective Identities in elderly Nursing Home residents for the design of an emotionally intelligent cognitive assistant. In *Proc. International Conference on Pervasive Computing Technologies for Healthcare*, Cancun, Mexico, 2016.

Alexandra Konig, Linda E. Francis, Jyoti Joshi, Julie M Robillard, and Jesse Hoey. Qualitative study of affective identities in dementia patients for the design of cognitive assistive technologies. *Journal of Rehabilitation and Assistive Technologies Engineering*, 2017.

Ronak Kosti, Jose M Alvarez, Adria Recasens, and Agata Lapedriza. EMOTIC: Emotions in Context dataset. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 61–69, 2017.

Jennifer T Kubota, Mahzarin R Banaji, and Elizabeth A Phelps. The neuroscience of race. *Nature neuroscience*, 15(7):940–948, 2012.

Takashi Kuremoto, Tetsuya Tsurusaki, Kunikazu Kobayashi, Shingo Mabu, and Masanao Obayashi. A model of emotional intelligent agent for cooperative goal exploration. In *International Conference on Intelligent Computing*, pages 21–30. Springer, 2013.

L. Morgado and G. Gaspar. Emotion based adaptive reasoning for resource bounded agents. In *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, pages 921–928, 2005.

L. Morgado and, G. Gaspar. Abstraction level regulation of cognitive processing through emotion-based attention mechanisms. In *International Workshop on Attention in Cognitive Systems*, pages 59–74. Springer, 2007.

C.G. Lange and W. James. *The Emotions*. Hafner Pub. Co., 1967.

Edward J. Lawler, Shane R. Thye, and Jeongkoo Yoon. *Social Commitments in a Depersonalized World*. Russell Sage Foundation, 2009.

Richard S. Lazarus. *Emotion and Adaptation*. Oxford University Press USA, 1991.

Won-Hyong Lee and Jong-Hwan Kim. Hierarchical emotional episodic memory for social human robot collaboration. *Autonomous Robots*, 42(5):1087–1102, 2018.

Christopher P Lee-Johnson and Dale A Carnegie. Mobile robot navigation modulated by artificial emotions. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 40(2):469–480, 2009.

MOHSEN Lesani, AMIR Moghimi, AA Bitaghsir, and CARO Lucas. An emotional decision making help provision approach to distributed fault tolerance in mas. In *Proceedings World Automation Congress, 2004.*, volume 16, pages 299–304. IEEE, 2004.

Shifeng Li, Shi Feng, Daling Wang, Kaisong Song, Yifei Zhang, and Weichao Wang. EmoElicitor: an open domain response generation model with user emotional reaction awareness. In *Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence*, pages 3637–3643, 2021.

Mei Yii Lim, João Dias, Ruth Aylett, and Ana Paiva. Creating adaptive affective autonomous NPCs. *Autonomous Agents and Multi-Agent Systems*, 24(2):287–311, 2012.

Christine Laetitia Lisetti and Piotr Gmytrasiewicz. Can a rational agent afford to be affectless? A formal approach. *Applied Artificial Intelligence*, 16(7-8):577–609, 2002.

Juan Liu and Hiroshi Ando. Emotion eliciting and decision making by psychodynamic appraisal mechanism. In *2008 Conference on Human System Interactions*, pages 645–650. IEEE, 2008.

Changchun Liu, Karla Conn, Nilanjan Sarkar, and Wendy Stone. Online affect detection and robot behavior adaptation for intervention of children with autism. *IEEE transactions on robotics*, 24(4):883–896, 2008.

Difei Liu, Shen Zhuo Liu, Xiaoming Liu, Chong Qing Zhang, Aosika Li, Chenggong Jin, Yijun Chen, Hangwei Wang, and Xiaochu Zhang. Interactive Brain Activity: Review and Progress on EEG-Based Hyperscanning in Social Interactions. *Frontiers in Psychology*, 9:1862, 2018.

Michael Loizou, Thomas Hartley, Sarah Slater, Robert Newman, and Lucia Pannese. Emotions for intelligent agents in crisis management simulations: a survey. In *2012 17th International Conference on Computer Games (CGAMES)*, pages 213–219. IEEE, 2012.

Cheng-Xiang Lu, Zhi-Yuan Sun, Zhong-Zhi Shi, and Bao-Xiang Cao. Using emotions as intrinsic motivation to accelerate classic reinforcement learning. In *2016 International Conference on Information System and Artificial Intelligence (ISAI)*, pages 332–337. IEEE, 2016.

Fei Lu, Wenjia Si, and Guohui Tian. Autonomous cognition and correction system of robot service based on emotional information and case-based reasoning. In *2021 33rd Chinese Control and Decision Conference (CCDC)*, pages 3235–3240. IEEE, 2021.

M. Belkaid, N. Cuperlier, and P. Gaussier. Emotional modulation of peripersonal space as a way to represent reachable and comfort areas. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 353–359. IEEE, 2015.

M. Belkaid, N. Cuperlier and P. Gaussier. Autonomous cognitive robots need emotional modulations: Introducing the eMODUL model. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 49(1):206–215, 2018.

Neil MacKinnon and Dawn Robinson. Back to the Future: 25 Years of Research in Affect Control Theory. *Advances in Group Processes*, 31:139–173, 2014.

Neil Joseph MacKinnon. *Symbolic interactionism as affect control.* Symbolic interactionism as affect control. Suny Press, 1994. Pages: xvi, 245.

Aarti Malhotra and Jesse Hoey. Emotions in Socio-cultural Interactive AI Agents. In *2021 9th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*, pages 1–6, 2021.

Aarti Malhotra, Terrence C Stewart, and Jesse Hoey. A Biologically-Inspired Neural Implementation of Affect Control Theory. In *International Conference on Cognitive Modelling, Toronto*, 2020.

Aarti Malhotra, Garima Sharma, Rishikesh Kumar, Abhinav Dhall, and Jesse Hoey. Social Event Context and Affect Prediction in Group Videos. In *2023 11th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*, pages 1–8, 2023.

Aarti Malhotra, Jesse Hoey, Rebecca Hutchinson, and Gabby Chan. Computational Models of Emotions in Decision-Making Agents: A Systematic Review. *IEEE Transactions on Affective Computing*, Under Review.

Marcos Maroto-Gómez, María Malfaz, Álvaro Castro-González, and Miguel Ángel Salichs. A motivational model based on artificial biological functions for the intelligent decision-making of social robots. *Memetic Computing*, pages 1–21, 2023.

Andreas Marpaung and Avelino Gonzalez. Can an affect-sensitive system afford to be context independent? In *Modeling and Using Context: 10th International and Interdisciplinary Conference, CONTEXT 2017, Paris, France, June 20-23, 2017, Proceedings 10*, pages 454–467. Springer, 2017.

Juan Martınez-Miranda and Arantza Aldea. Emotions in human and artificial intelligence. *Computers in Human Behavior*, 21(2):323–341, 2005.

Aleix M Martinez. Context may reveal how you feel. *Proceedings of the National Academy of Sciences*, 116(15):7169–7171, 2019.

Mehdi Mashayekhi, Nirav Ajmeri, George F List, and Munindar P Singh. Prosocial norm emergence in multi-agent systems. *ACM Transactions on Autonomous and Adaptive Systems (TAAS)*, 17(1-2):1–24, 2022.

Atsushi Matsuda, Hideaki Misawa, and Keiichi Horio. Decision making based on reinforcement learning and emotion learning for social behavior. In *2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011)*, pages 2714–2719. IEEE, 2011.

Riccardo Mattivi, Jasper Uijlings, Francesco GB De Natale, and Nicu Sebe. Exploitation of time constraints for (sub-) event recognition. In *Proceedings of the 2011 joint ACM workshop on Modeling and representing events*, pages 7–12, 2011.

George Herbert Mead. *Mind, Self and Society.* University of Chicago Press, 1934.

Ying Mei. Emotion-driven attention of the robotic manipulator for action selection. In *2016 35th Chinese Control Conference (CCC)*, pages 7173–7178. IEEE, 2016.

Nesreen Mejbri, Fathi Essalmi, Mohamed Jemni, and Bader A Alyoubi. Trends in the use of affective computing in e-learning environments. *Education and Information Technologies*, pages 1–23, 2022.

Trisha Mittal, Puneet Mathur, Aniket Bera, and Dinesh Manocha. Affect2mm: Affective analysis of multimedia content using emotion causality. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5661–5671, 2021.

Jan Mizgajski and Mikołaj Morzy. Affective recommender systems in online news industry: how emotions influence reading choices. *User Modeling and User-Adapted Interaction*, 29(2):345–379, 2019.

Tetsuji Mochida, Akio Ishiguro, Tadeshi Aoki, and Yoshiki Uchikawa. Behavior arbitration for autonomous mobile robots using emotion mechanisms. In *Proceedings 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems. Human Robot Interaction and Cooperative Robots*, volume 3, pages 516–521. IEEE, 1995.

Thomas M Moerland, Joost Broekens, and Catholijn M Jonker. Emotion in reinforcement learning agents and robots: a survey. *Machine Learning*, 107:443–480, 2018.

Luís Morgado and Graça Gaspar. Adaptation and decision-making driven by emotional memories. In *Portuguese Conference on Artificial Intelligence*, pages 102–114. Springer, 2005.

Maher Ben Moussa and Nadia Magnenat-Thalmann. Toward socially responsible agents: integrating attachment and learning in emotional decision-making. *Computer Animation and Virtual Worlds*, 24(3-4):327–334, 2013.

N. Esau, L. Kleinjohann, and B. Kleinjohann. Integration of emotional reactions on human facial expressions into the robot head MEXI. In *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 534–541. IEEE, 2007.

Ali Nasir. Markov decision process for emotional behavior of socially assistive robot. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 1198–1203. IEEE, 2018.

Rony Novianto and Mary-Anne Williams. Innate and learned emotion network. In *Biologically Inspired Cognitive Architectures 2011*, pages 263–268. IOS Press, 2011.

Suman Ojha, Jonathan Vitale, and Mary-Anne Williams. Computational emotion models: a thematic review. *International Journal of Social Robotics*, 13:1253–1279, 2021.

Charles E. Osgood, G. J. Suci, and Percy H. Tannenbaum. *The Measurement of Meaning*. University of Illinois Press, Urbana, 1957.

Charles Egerton Osgood, George J Suci, and Percy H Tannenbaum. *The measurement of meaning*. University of Illinois press, 1957.

Charles E. Osgood, William H. May, and Murray S. Miron. *Cross-cultural universals of affective meaning.* Cross-cultural universals of affective meaning. University of Illinois Press, 1975. Pages: xix, 486.

Matthew J Page, Joanne E McKenzie, Patrick M Bossuyt, Isabelle Boutron, Tammy C Hoffmann, Cynthia D Mulrow, Larissa Shamseer, Jennifer M Tetzlaff, Elie A Akl, Sue E Brennan, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Systematic reviews*, 10(1):1–11, 2021.

Ana Paiva, Iolanda Leite, Hana Boukricha, and Ipke Wachsmuth. Empathy in virtual agents and robots: A survey. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 7(3):1–40, 2017.

Alexandros Papangelis, Fillia Makedon, and Robert Gatchel. Assessing and Monitoring Post-Traumatic Stress Disorder through Natural Interaction with an Adaptive Dialogue System. *Journal of Applied Biobehavioral Research*, 19(3):192–215, 2014.

Paola Pennisi, Alessandro Tonacci, Gennaro Tartarisco, Lucia Billeci, Liliana Ruta, Sebastiano Gangemi, and Giovanni Pioggia. Autism and social robotics: A systematic review. *Autism Research*, 9(2):165–183, 2016.

Luís Moniz Pereira, Tom Lenaerts, Luis A Martinez-Vaquero, and The Anh Han. Social Manifestation of Guilt Leads to Stable Cooperation in Multi-Agent Systems. In *AAMAS*, pages 1422–1430, 2017.

Elizabeth A Phelps, Karolina M Lempert, and Peter Sokol-Hessner. Emotion and decision making: multiple modulatory neural circuits. *Annual review of neuroscience*, 37:263–287, 2014.

César F Pimentel and Maria R Cravo. Don't think too much! —Artificial somatic markers for action selection. In *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, pages 1–8. IEEE, 2009.

Marco Polignano, Fedelucio Narducci, Marco de Gemmis, and Giovanni Semeraro. Towards emotion-aware recommender systems: an affective coherence model based on emotion-driven behaviors. *Expert Systems with Applications*, 170:114382, 2021.

Matthijs Pontier and Ghazanfar Farooq Siddiqui. An affective agent playing Tic-Tac-Toe as part of a healing environment. In *International Conference on Principles and Practice of Multi-Agent Systems*, pages 33–47. Springer, 2009.

Matthijs A Pontier, Guy Widdershoven, and Johan F Hoorn. Moral coppélia-combining ratio with affect in ethical reasoning. In *Ibero-American Conference on Artificial Intelligence*, pages 442–451. Springer, 2012.

Brittany Postnikoff. Robot social engineering. Master's thesis, University of Waterloo, 2020.

William T. Powers. *Behavior: The control of perception*. Aldine publishing co., Chicago, 1973.

Joanne Pransky. AIBO–the No. 1 selling service robot. *Industrial robot: An international journal*, 2001.

Rendhir R Prasad and Ciza Thomas. Ontology-Based Emotional Decision-Making in Self-Evolving Defensive Agents. *IEEE Access*, 10:108749–108759, 2022.

Tony J Prescott, Ben Mitchinson, and Sebastian Conran. Miro: An animal-like companion robot with a biomimetic brain-based control system. In *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, pages 50–51, 2017.

Dharani Punithan and Byoung-Tak Zhang. Lifelong Learning with the Feedback-loop between Emotions and Actions via Internal Reward. *Procedia computer science*, 145:423–427, 2018.

Xiaotong Qie, Jinghua Wu, Yan Li, and Yi Sun. A stage model for agent-based emotional persuasion with an adaptive target: From a social exchange perspective. *Information Sciences*, 610:90–113, 2022.

Kha Gia Quach, Ngan Le, Khoa Luu, Chi Nhan Duong, Ibsa Jalata, and Kaushik Roy. Non-volume preserving-based feature fusion approach to group-level expression recognition on crowd videos. *arXiv preprint arXiv:1811.11849*, 2018.

Niklas Rach, Klaus Weber, Yuchi Yang, Stefan Ultes, Elisabeth André, and Wolfgang Minker. EVA 2.0: Emotional and rational multimodal argumentation between virtual agents. *it-Information Technology*, 63(1):17–30, 2021.

Khamael Raqim Raheem and Israa Hadi Ali. Survey: Affective recommender systems techniques. In *IOP Conference Series: Materials Science and Engineering*, volume 928, page 032042. IOP Publishing, 2020.

Alexandr Rassadin, Alexey Gruzdev, and Andrey Savchenko. Group-level emotion recognition using transfer learning from face identification. In *ACM International Conference on Multimodal Interaction*, pages 544–548, 2017.

Timo Reuter, Symeon Papadopoulos, Giorgos Petkos, Vasileios Mezaris, Yiannis Kompatsiaris, Philipp Cimiano, Christopher De Vries, and Shlomo Geva. Social event detection at mediaeval 2013: Challenges, datasets, and evaluation. In *Proceedings of the MediaEval 2013 Multimedia Benchmark Workshop Barcelona, Spain, October 18-19, 2013*, 2013.

James K Rilling and Alan G Sanfey. The neuroscience of social decision-making. *Annual review of psychology*, 62:23–48, 2011.

JA Rincon, Fernando de la Prieta, Damiano Zanardini, Vicente Julian, and Carlos Carrascosa. Influencing over people with a social emotional model. *Neurocomputing*, 231:47–54, 2017.

Dawn T Robinson, Lynn Smith-Lovin, and Allison K Wisecup. Affect control theory. In *Handbook of the sociology of emotions*, pages 179–202. Springer, 2006.

Luis-Felipe Rodriguez, Felix Ramos, and Yingxu Wang. Cognitive computational models of emotions. In *IEEE 10th International Conference on Cognitive Informatics and Cognitive Computing (ICCI-CC'11)*, pages 75–84. IEEE, 2011.

Shaghayegh Roohi, Jari Takatalo, Christian Guckelsberger, and Perttu Hämäläinen. Review of intrinsic motivation in simulation-based game testing. In *Proceedings of the 2018 chi conference on human factors in computing systems*, pages 1–13, 2018.

Ira Roseman. Emotional behaviors, emotivational goals, emotion strategies: Multiple levels of organization integrate variable and consistent responses. *Emotion Review*, 3(4):434–443, 10 2011.

Fabian Rubilar, María-José Escobar, and Tomás Arredondo. Bio-inspired architecture for a reactive-deliberative robot controller. In *2014 International Joint Conference on Neural Networks (IJCNN)*, pages 2027–2035. IEEE, 2014.

Timothy Rumbell, John Barnden, Susan Denham, and Thomas Wennekers. Emotions in autonomous agents: comparative analysis of mechanisms and functions. *Autonomous Agents and Multi-Agent Systems*, 25:1–45, 2012.

Miguel A Salichs and Maria Malfaz. A New Approach to Modeling Emotions and Their Use on a Decision-Making System for Artificial Agents. *IEEE Transactions on Affective Computing*, 1(3):56–68, 2012.

Abdolhossein Sarrafzadeh, Samuel Alexander, Farhad Dadgostar, Chao Fan, and Abbas Bigdeli. “How do you know that I don’t understand?” A look at the future of intelligent tutoring systems. *Computers in Human Behavior*, 24(4):1342–1363, 2008.

Klaus R. Scherer. Emotions are emergent processes: they require a dynamic computational architecture. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 364 1535:3459–74, 2009.

Paul W Schermerhorn and Matthias Scheutz. The Utility of Affect in the Selection of Actions and Goals Under Real-World Constraints. In *IC-AI*, pages 948–853, 2009.

Tobias Schröder, Jesse Hoey, and Kimberly B Rogers. Modeling dynamic identities and uncertainty in social interactions: Bayesian affect control theory. *American Sociological Review*, 81(4):828–855, 2016.

N Schwarz and GL Clore. How Do I Feel About It? Information Functions of Affective States. *Affect, Cognition and Social Behavior (1988: Hogrefe International, Toronto)*, pages 44–62.

Norbert Schwarz and Gerald L. Clore. Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of personality and social psychology*, 45(3):513–523, 1983.

Norbert Schwarz. *Feelings as information: Informational and motivational functions of affective states.* The Guilford Press, 1990.

Pedro Sequeira, Francisco S Melo, and Ana Paiva. Learning by appraising: an emotion-based approach to intrinsic reward design. *Adaptive Behavior*, 22(5):330–349, 2014.

Gilad Sharir, Asaf Noy, and Lihi Zelnik-Manor. An Image is Worth 16x16 Words, What is a Video Worth? *arXiv preprint arXiv:2103.13915*, 2021.

Garima Sharma, Shreya Ghosh, and Abhinav Dhall. Automatic Group Level Affect and Cohesion Prediction in Videos. In *2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*, pages 161–167, 2019.

Garima Sharma, Abhinav Dhall, and Jianfei Cai. Audio-visual Automatic Group Affect Analysis. *IEEE Transactions on Affective Computing*, 2021.

Manika Sharma, Raman Mittal, Ambuj Bharati, Deepika Saxena, and Ashutosh Kumar Singh. Emotional information-based hybrid recommendation system. In *Soft Computing for Problem Solving: Proceedings of the SocProS 2022*, pages 249–267. Springer, 2023.

Alexei Sharpanskykh and Jan Treur. Adaptive modelling of social decision making by agents integrating simulated behaviour and perception chains. In *International Conference on Computational Collective Intelligence*, pages 284–295. Springer, 2010.

Alexei Sharpanskykh and Jan Treur. An adaptive agent model for affective social decision making. *Biologically Inspired Cognitive Architectures*, 5:72–81, 2013.

Alexei Sharpanskykh and Kashif Zia. Understanding the role of emotions in group dynamics in emergency situations. In *Transactions on computational collective intelligence XV*, pages 28–48. Springer, 2014.

Henrique Siqueira, Alexander Sutherland, Pablo Barros, Mattias Kerzel, Sven Magg, and Stefan Wermter. Disambiguating affective stimulus associations for robot perception and dialogue. In *2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids)*, pages 1–9. IEEE, 2018.

Jan E Stets and Richard T Serpe. *Identities in everyday life*. Oxford University Press, USA, 2019.

Terrence C Stewart, Xuan Choo, and Chris Eliasmith. Dynamic Behaviour of a Spiking Model of Action Selection in the Basal Ganglia. In *Proceedings of the 10th International Conference on Cognitive Modeling*, pages 235–40, 2010.

Terrence C. Stewart, Trevor Bekolay, and Chris Eliasmith. Learning to select actions with spiking neurons in the basal ganglia. *Frontiers in Decision Neuroscience*, 6:2, 2012.

Latha Subramainan and Moamin A Mahmoud. Academic Emotions Review: Types, Triggers, Reactions, and Computational Models. In *2020 8th International Conference on Information Technology and Multimedia (ICIMU)*, pages 223–230. IEEE, 2020.

Lianzhi Tan, Kaipeng Zhang, Kai Wang, Xiaoxing Zeng, Xiaojiang Peng, and Yu Qiao. Group Emotion Recognition with Individual Facial Emotion CNNs and Global Image Based CNNs. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, ICMI '17, pages 549–552, New York, NY, USA, 2017. ACM.

Sai Prithvisingh Taurah, Jeshta Bhoyedhur, and Roopesh Kevin Sungkur. Emotion-Based Adaptive Learning Systems. In *International Conference on Machine Learning for Networking*, pages 273–286. Springer, 2019.

Peggy A Thoits. The sociology of emotions. *Annual review of sociology*, 15(1):317–342, 1989.

Thi Ngoc Trang Tran, Alexander Felfernig, and Nava Tintarev. Humanized recommender systems: State-of-the-art and research issues. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 11(2):1–41, 2021.

Abhishek Tripathi, TS Ashwin, and Ram Mohana Reddy Guddeti. Emoware: A context-aware framework for personalized video recommendation using affective video sequences. *IEEE Access*, 7:51185–51200, 2019.

Diana D Tsankova. Emotionally influenced coordination of behaviors for autonomous mobile robots. In *Proceedings First International IEEE Symposium Intelligent Systems*, volume 1, pages 92–97. IEEE, 2002.

Diana D Tsankova. Emotional intervention on an action selection mechanism based on artificial immune networks for navigation of autonomous agents. *Adaptive Behavior*, 17(2):135–152, 2009.

Sz-Ting Tzeng, Nirav Ajmeri, and Munindar P Singh. Noe: Norm Emergence and Robustness Based on Emotions in Multiagent Systems. In *International Workshop on Coordination, Organizations, Institutions, Norms, and Ethics for Governance of Multi-Agent Systems*, pages 62–77. Springer, 2021.

Wilco W van Dijk and Joop Van der Pligt. The impact of probability and magnitude of outcome on disappointment and elation. *Organizational Behavior and Human Decision Processes*, 69(3):277–284, 1997.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

Yufei Wang, Zhe Lin, Xiaohui Shen, Radomir Mech, Gavin Miller, and Garrison W Cottrell. Event-specific image importance. In *proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4810–4819, 2016.

Xiaofang Wang, Xuehan Xiong, Maxim Neumann, AJ Piergiovanni, Michael S Ryoo, Anelia Angelova, Kris M Kitani, and Wei Hua. Attentionnas: Spatiotemporal attention cell search for video classification. In *European Conference on Computer Vision*, pages 449–465. Springer, 2020.

Hui Wang, Jinying Lin, Zhen Ma, Yurii Vasylkiv, Heike Brock, Keisuke Nakamura, Randy Gomez, Bo He, and Guangliang Li. Affective Behavior Learning for Social Robot Haru with Implicit Evaluative Feedback. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3881–3888. IEEE, 2022.

Donghyeon Won, Zachary C Steinert-Threlkeld, and Jungseock Joo. Protest activity detection and perceived violence estimation from social media images. In *Proceedings of the 25th ACM international conference on Multimedia*, pages 786–794, 2017.

Jinghua Wu, Hongyu Chen, Yan Li, and Yunhui Liu. A behavioral assessment model for emotional persuasion driven by agent-based decision-making. *Expert Systems with Applications*, 204:117556, 2022.

Shengkun Xie, Anna T Lawniczak, and Chong Gan. Modeling and Analysis of Autonomous Agents' Decisions in Learning to Cross a Cellular Automaton-Based Highway. *Computation*, 7(3):53, 2019.

Yuanjun Xiong, Kai Zhu, Dahua Lin, and Xiaoou Tang. Recognize complex events from static images by fusing deep channels. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1600–1609, 2015.

Mei Ying and Liu Zhentao. An emotion-driven attention model for service robot. In *2016 12th World Congress on Intelligent Control and Automation (WCICA)*, pages 1526–1531. IEEE, 2016.

Nutchanon Yongsatianchot and Stacy Marsella. A computational model of coping and decision making in high-stress, uncertain situations: an application to hurricane evacuation decisions. *IEEE Transactions on Affective Computing*, 2022.

Chao Yu, Minjie Zhang, and Fenghui Ren. Emotional multiagent reinforcement learning in social dilemmas. In *International Conference on Principles and Practice of Multi-Agent Systems*, pages 372–387. Springer, 2013.

Chao Yu, Minjie Zhang, Fenghui Ren, and Guozhen Tan. Emotional multiagent reinforcement learning in spatial social dilemmas. *IEEE transactions on neural networks and learning systems*, 26(12):3083–3096, 2015.

Hongye Zhai, Jian Wang, Jialong Xie, Weijie Lin, and Botao Zhang. A local behavior selection method of mobile robots based on a novel "environment-emotion" model. In *2022 41st Chinese Control Conference (CCC)*, pages 3645–3650. IEEE, 2022.

Huidi Zhang and Shirong Liu. Design of autonomous navigation system based on affective cognitive learning and decision-making. In *2009 IEEE international conference on robotics and biomimetics (ROBIO)*, pages 2491–2496. IEEE, 2009.

Yutong Zhang and Zhigang Zeng. Neuromorphic Circuit Implementation of Operant Conditioning Based on Emotion Generation and Modulation. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2023.

Huidi Zhang, Shirong Liu, and Simon X Yang. A hybrid robot navigation approach based on partial planning and emotion-based behavior coordination. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1183–1188. IEEE, 2006.

Weidong Zhou and Richard Coggins. Computational models of the amygdala and the orbitofrontal cortex: A hierarchical reinforcement learning system for robotic control. In *Australian Joint Conference on Artificial Intelligence*, pages 419–430. Springer, 2002.

Han Zhou, Hongpeng Yin, Hengyi Zheng, and Yanxia Li. A survey on multi-modal social event detection. *Knowledge-Based Systems*, 195:105695, 2020.

APPENDICES

Appendix A

Systematic Review Search Queries

We provide search queries used for systematic review of computational models of emotion and decision-making or behavior. Query details and search results from four databases (Scopus, PsycINFO, ACM and IEEE) are given in sections A.1, A.4, A.3 and A.2 respectively. Total results obtained from searches in October 2021 are 2167 (Scopus) + 894 (PsycINFO) + 830 (ACM) + 1469 (IEEE)= 5360. The same queries were used to update the search in June 2023.

A.1 Scopus

Table A.1: Scopus queries and results (October 20th, 2021, 11am)

Search	Query	Field	Results
#1	"Affect" OR "affective" OR emotion* OR feel* OR sentiment*	Title, abstract, keywords	2,982,985
#2	anticipat* OR behavior* OR behaviour* OR select* OR adapt* OR decision* OR "decision mak*" OR decid* OR react* OR reason*	Title, abstract, keywords	20,310,461
#3	#1 W/5 #2	Title, abstract, keywords	223,803
#4	"adaptive agent*" OR "artificial agent*" OR "intelligent agent*" OR "AI agent*" OR "multi agent*" OR multiagent* OR "affective agent*" OR "cognitive agent*" OR "social agent*" OR "virtual agent*" OR "autonomous agent*" OR "sentient agent*" OR "recommendation agent*" OR "recommendation system*" OR "reccommender system*"	Title, abstract, keywords	163,865
#5	"artificial intelligence" OR "artificially intelligent" OR "computational intelligence" OR "artificial affect" OR "affective comput*" OR "fuzzy system*"	Title, abstract, keywords	457,714
#6	agent* OR "computational model*" OR "computational method*" OR robot*	Title, abstract, keywords	6,268,782
#7	#5 AND #6	Title, abstract, keywords	60,316
#8	#4 OR #7	Title, abstract, keywords	208,126
#9	#3 AND #8	Title, abstract, keywords	2211
#10	#9 AND limit to English		2167

A.2 ACM Digital Library

Table A.2: ACM queries and results (October 27th, 2021, 5pm)

Search	Query (using the Edit Query box)	Results
#1	Title:("Affect" OR "affective" OR emotion* OR feel* OR sentiment*) OR Abstract:("Affect" OR "affective" OR emotion* OR feel* OR sentiment*) OR Keyword:("Affect" OR "affective" OR emotion* OR feel* OR sentiment*)	33,481
#2	Title:(anticipat* OR behavior* OR behaviour* OR select* OR adapt* OR decision* OR decid* OR react* OR reason*) OR Abstract:(anticipat* OR behavior* OR behaviour* OR select* OR adapt* OR decision* OR decid* OR react* OR reason*) OR Keyword:(anticipat* OR behavior* OR behaviour* OR select* OR adapt* OR decision* OR decid* OR react* OR reason*)	180,069
#3	Title:("adaptive agent" OR "artificial agent" OR "intelligent agent" OR "AI agent" OR "multi agent" OR multiagent OR "affective agent" OR "cognitive agent" OR "social agent" OR "virtual agent" OR "adaptive agents" OR "artificial agents" OR "intelligent agents" OR "AI agents" OR "multi agents" OR multiagents OR "affective agents" OR "cognitive agents" OR "social agents" OR "virtual agents" OR "sentient systems" OR "autonomous agent" OR "recommendation system" OR "recommender system" OR "autonomous agents" OR "recommendation systems" OR "recommender systems" OR "expert systems" OR "expert system") OR Abstract:("adaptive agent" OR "artificial agent" OR "intelligent agent" OR "AI agent" OR "multi agent" OR multiagent OR "affective agent" OR "cognitive agent" OR "social agent" OR "virtual agent" OR "adaptive agents" OR "artificial agents" OR "intelligent agents" OR "AI agents" OR "multi agents" OR multiagents OR "affective agents" OR "cognitive agents" OR "social agents" OR "virtual agents" OR "sentient systems" OR "autonomous agent" OR "recommendation system" OR "recommender system" OR "autonomous agents" OR "recommendation systems" OR "recommender systems" OR "expert systems" OR "expert system") OR Keyword:("adaptive agent" OR "artificial agent" OR "intelligent agent" OR "AI agent" OR "multi agent" OR multiagent OR "affective agent" OR "cognitive agent" OR "social agent" OR "virtual agent" OR "adaptive agents" OR "artificial agents" OR "intelligent agents" OR "AI agents" OR "multi agents" OR multiagents OR "affective agents" OR "cognitive agents" OR "social agents" OR "virtual agents" OR "sentient systems" OR "autonomous agent" OR "recommendation system" OR "recommender system" OR "autonomous agents" OR "recommendation systems" OR "recommender systems" OR "expert systems" OR "expert system")	14,188
#4	Title:("artificial intelligence" OR "artificially intelligent" OR "computational intelligence" OR "affective computing") OR Abstract:("artificial intelligence" OR "artificially intelligent" OR "computational intelligence" OR "affective computing") OR Keyword:("artificial intelligence" OR "artificially intelligent" OR "computational intelligence" OR "affective computing")	6,180
#5	Title:(agent* OR robot* OR "computational model" OR "computational models" OR "computational modeling" OR "computational modelling") OR Abstract:(agent* OR robot* OR "computational model" OR "computational models" OR "computational modeling" OR "computational modelling") OR Keyword:(agent* OR robot* OR "computational model" OR "computational models" OR "computational modeling" OR "computational modelling")	48,594
#6	#4 AND #5	1,357
#7	#3 OR #6	15,181
#1	#1 AND #2 AND #7	830

A.3 PsycINFO

Table A.3: PsycINFO queries and results (October 20th, 2021, 1:30pm)

Search	Query	Field	Results
#1	"Affect" OR "affective" OR emotion* OR feel* OR sentiment*	Keywords	
#2	"Affect" OR "affective" OR emotion* OR feel* OR sentiment*	Title	
#3	"Affect" OR "affective" OR emotion* OR feel* OR sentiment*	Abstract	
#4	{Emotions} OR {Contempt} OR {Desire} OR {Emotional Content} OR {Emotional Disturbances} OR {Emotional Health} OR {Emotional Processing} OR {Emotional Regulation} OR {Emotional States} OR {Emotional Style} OR {Emotional Support} OR {Expressed Emotion} OR {Forgiveness} OR {Negative Emotions} OR {Positive Emotions}	Index Terms	
#5	#1 OR #2 OR #3 OR #4		769,948
#6	anticipat* OR behavior* OR behaviour* OR select* OR adapt* OR decision* OR decid* OR react* OR reason* OR "decision science*" OR action*	Keywords	
#7	anticipat* OR behavior* OR behaviour* OR select* OR adapt* OR decision* OR decid* OR react* OR reason* OR "decision science*" OR action*	Title	
#8	anticipat* OR behavior* OR behaviour* OR select* OR adapt* OR decision* OR decid* OR react* OR reason* OR "decision science*" OR action*	Abstract	
#9	Behavior OR Decision Making OR Choice Behavior OR Approach Behavior OR Behavior Change OR Performance OR Planned Behavior OR Reasoned Action OR Responses OR Behavior Analysis	Index Terms	
#10	#6 OR #7 OR #8 OR #9		2,192,290
#11	"artificial intelligence" OR "artificially intelligent" OR "computational intelligence" OR "artificial affect" OR "affective comput*" OR "fuzzy system*" OR "dynamic computational model*" OR "drift diffusion model*" OR "predictive model*" OR "predictive process*" OR "iterative reprocessing" OR "Bayesian statistics" OR "Bayesian model*"	Keywords	
#12	"artificial intelligence" OR "artificially intelligent" OR "computational intelligence" OR "artificial affect" OR "affective comput*" OR "fuzzy system*" OR "dynamic computational model*" OR "drift diffusion model*" OR "predictive model*" OR "predictive process*" OR "iterative reprocessing" OR "Bayesian statistics" OR "Bayesian model*"	Title	
#13	"artificial intelligence" OR "artificially intelligent" OR "computational intelligence" OR "artificial affect" OR "affective comput*" OR "fuzzy system*" OR "dynamic computational model*" OR "drift diffusion model*" OR "predictive model*" OR "predictive process*" OR "iterative reprocessing" OR "Bayesian statistics" OR "Bayesian model*"	Abstract	
#14	{Affective Computing} OR {Artificial Intelligence} OR {Expert Systems} OR {Artificial Neural Networks} OR {Cognitive Computing} OR {Knowledge Engineering} OR {Knowledge Representation} OR {Machine Learning} OR {Fuzzy Logic} OR {Intelligent Agents}	Index Terms	
#15	#11 OR #12 OR #13 OR #14		36,852
#16	agent* OR robot* OR "computational model*" OR "computational method*"	Keywords	
#17	agent* OR robot* OR "computational model*" OR "computational method*"	Title	
#18	agent* OR robot* OR "computational model*" OR "computational method*"	Abstract	
#19	{Computational Modeling} OR {Human Machine Systems} OR {Human Machine Systems Design} OR {Human Robot Interaction} OR {Computer Simulation} OR {Computational Neuroscience}	Index Terms	
#20	#16 OR #17 OR #18 OR #19		99,418
#21	#15 AND #20		8,309
#22	"adaptive agent*" OR "artificial agent*" OR "intelligent agent*" OR "AI agent*" OR "multi agent*" OR multiagent* OR "affective agent*" OR "cognitive agent*" OR "social agent*" OR "virtual agent*" OR "sentient agent*" OR "autonomous agent*" OR "sentient agent*" OR "recommendation agent*" OR "recommendation system*" OR "recommender system*"	Keywords	
#23	"adaptive agent*" OR "artificial agent*" OR "intelligent agent*" OR "AI agent*" OR "multi agent*" OR multiagent* OR "affective agent*" OR "cognitive agent*" OR "social agent*" OR "virtual agent*" OR "sentient agent*" OR "autonomous agent*" OR "sentient agent*" OR "recommendation agent*" OR "recommendation system*" OR "recommender system*"	Title	
#24	"adaptive agent*" OR "artificial agent*" OR "intelligent agent*" OR "AI agent*" OR "multi agent*" OR multiagent* OR "affective agent*" OR "cognitive agent*" OR "social agent*" OR "virtual agent*" OR "sentient agent*" OR "autonomous agent*" OR "sentient agent*" OR "recommendation agent*" OR "recommendation system*" OR "recommender system*"	Abstract	
#25	{Social Robotics}	Index Terms	
#26	#22 OR #23 OR #24 OR #25		5,466
#27	#21 OR #26		11,629
#28	#5 AND #10 AND #27		894

A.4 IEEE Xplore Digital Library

Table A.4: IEEE Xplore Digital Library queries and results (October 20th, 2021, 11:30am)

Search	Query	Field	Results
#1	"Affect" OR "affective" OR feeling OR feel OR emotion* OR sentiment OR sentimental	All metadata	98,522
#2	Anticipat* OR behavior OR behavioral OR select* OR adapt* OR decision OR "decision making" OR "decision maker" OR "decision makers" OR decide OR react* OR reason*	All metadata	1,361,090
#3	#1 NEAR/10 #2	All metadata	11,954
#4	adaptive OR artificial OR intelligent OR intelligence OR AI OR affective OR cognitive OR social OR virtual OR recommender OR recommendation OR sentient OR autonomous	All metadata	1,218,823
#5	agent OR robot OR "computational model*" OR "computational method" OR "computational methods"	All metadata	455,779
#6	#4 NEAR/10 #5	All metadata	140,943
#7	#3 AND #7	All metadata	1469