Humans are not rational; artificial agents are not emotional

Leon S. Sterling[1,*,†], James Marshall[1,†]

[1]Swinburne University of Technology, John Street, Hawthorn, Victoria, 3122, Australia

Abstract

Agents have been used as a unifying metaphor while designing socio-technical systems where humans interact with software components. As socio-technical systems have become consumer-oriented in areas such as health, it has been acknowledged that emotions need to be considered if people are to adopt technology. This paper argues that the common agent metaphor for humans and software is limited. Humans and artificial agents need to be considered and modelled differently with respect to both logical behaviour and emotional goals.

Keywords

Agents, emotions, motivational modelling, design

1. Setting the scene

Multi-agent systems have been championed for over 30 years as an approach to help design and develop intelligent software. The push for agents and multi-agent systems came from two perspectives - distributed intelligence, and software methodology where agents were viewed as a natural successor to objects. The latter perspective is reflected in the name of the workshop WOA 2022, Workshop from Objects to Agents. A prevailing view of agents can be found, for example, in the introductory textbook of Wooldridge. [1]. A view more focussed on Artificial Intelligence is given by Russell and Norvig [2].

Computing is pervasive in today's world and deployed over a range of devices by a multiplicity of users. We are developing and using software systems to interact with both the ever-increasing complexity of the technical world and the growing fluidity of social organizations. How to conceptualize the components of the system? As argued by Sterling and Taveter [3] for example, agents represent a conceptual model for developing software systems that are open, intelligent and adaptive.

Over the past decade the external environment has changed where smart devices are ubiquitous and people use apps in almost all aspects of their lives. There is a consumer culture that is pervasive. Catering to consumer behaviour is a different proposition for developing software

WOA2022: Workshop from Objects to Agents, September 1-2, 2022, Genova, Italy

*Corresponding author

†These authors contributed equally.

lsterling@swin.edu.au (L. S. Sterling); jgmarshall@swin.edu.au (J. Marshall)

https://leonsterling.com (L. S. Sterling)

0000-0002-4122-2465 (L. S. Sterling)

and services. Previously, software was primarily built for individual organisations and used to solve business problems. The boom in smart devices for personal use has increased focus towards the needs and experiences of end-users, allowing technology to support individuals and communities in their everyday lives. In the consumer space, people rapidly make discretionary decisions to use or reject apps. People can quickly reject new technology if it doesn't appeal to their emotional needs. As argued in [4], emotional responses are a key consideration determining whether people adopt technology.

This paper has emerged from grappling with the changing nature of software and service development and renewed engagement with emotions. We argue that the role of the agent metaphor, in particular close identification between humans and artificial agents (for the lack of a better term) needs to be reconsidered. The next section presents a range of background issues, including current thinking about emotions, which has been developing in recent years with the emergence of the theory of constructed emotions, why emotions need to be considered when developing system requirements and what we have learned about modelling software where emotional goals have been expressed. An important consideration for the development of systems is the value of a shared understanding of system goals among diverse team members coming from different disciplines.

Section 3 argues that artificial agents are not emotional in the way that human agents are, and it is misguided to try and equate the two. Section 4 argues that human agents are not logical in the way that artificial agents are modelled, and that should be more explicitly recognised in system implementation. Section 5 argues for design to play a leading role in building socio-technical systems that meet emotional requirements. Section 6 concludes.

2. Underpinnings

Emotions have been considered in some early research on agents. Affective computing had been suggested by several as an approach to consider [5]. Believable agents with emotions were proposed as important to increase uptake of agent systems [6]. A paper presented at an AAAI conference 20 years ago by Scheutz [7] noted that there was no agreement about emotions. Scheutz pointed out that certain authors are quick to label robot (or agent) behaviours as emotions with little justification from psychology research. We agree about the inappropriate labelling of behaviour as emotional. Several terms are conflated as emotions, such as affect, motivation, and basic drives like hunger. In this section we unpick some of the underlying concepts and suggest a way of handling emotional goals in system development.

2.1. The nature of emotions

Emotions have long been recognised as a key part of human experience. Biblical texts use lots of emotion words as an attempt to describe a good life. There is a well established link between emotion, art and design.

In his book The Art of Experience [8], American philosopher John Dewey presents artefacts dating back over 3000 years from the Egyptian, Greek, Minoan, and Roman Empires that evidence the profound importance emotion played in designing not only ancient products but entire civilisations. According to Dewey, art and design represented far more than mere

decoration to people from these ancient civilisations. Rather they reflected the emotions and ideas that are associated with the chief institutions of social life. Dewey argues that emotion is the foundation of aesthetic experience; being the ultimate judgment upon the quality of a civilisation.

It has been argued that emotions are part of our social adaptation strategy. Emotional reactions are formed quickly, and are part of the fast processing that Kahneman discusses in 'Thinking: Fast and Slow' [9]. People needed to respond quickly to people in order to survive as part of a group. Kahneman distinguishes more cognitive thinking as a slower process from faster associative thinking. In 'Behave' [10], Robert Sapolsky distinguishes between cognition and affect and discusses in great detail their effects on behaviour. It is beyond scope to discuss in detail, but the interaction between them is complex.

As early as the 4th century BCE, Aristotle attempted to identify the exact number of core emotions in humans. Described as Aristotle's list of emotion, the philosopher proposed fourteen distinct emotional expressions: fear, confidence, anger, friendship, calm, enmity, shame, shamelessness, pity, kindness, envy, indignation, emulation, and contempt. More than 2000 years after Aristotle created his list of emotions, and despite significant advances in neuroscience and neuroimaging, not much has changed.

A predominant classical view of emotion still claims distinct and universal human emotions as part of our essential core. According to Ekman [11] (Ekman, 2003) there are seven facial expressions that represent universal emotions in the classical theory of emotions. Ekman claimed that the expressions and emotions are common to all humans independent of culture. He opined that are a few more emotions worth exploring and understanding, including contempt. However most of these additional emotional facial expressions can be organised and placed into one of seven emotion families: anger, happiness, sadness, contempt, surprise, fear, disgust.

Ekman's claims have been steadily debunked in recent times. In 'A Human History of Emotion' [12] the author Richard Firth-Godbehere makes a convincing demonstration on how emotions have varied in different cultures throughout the world undermining Ekman's claims of universality. Firth-Godbehere's PhD research was on how disgust has been viewed throughout history with reference to literature. The evidence of change is compelling.

Ekman further claimed that emotions are physical things that are hardwired into the brain as objectively real constructs. Researchers have not been able to establish such a list. There is too much ambiguity on which emotions people are feeling when being measured.

New research into emotion suggests that we have far more agency over creating emotions, both in ourselves and others than previously thought. Developed by Professor of Psychology Lisa Feldman Barrett, the theory disagrees with the classical view of emotions that we are born with emotions built-in [13]. In contrast, the theory of constructed emotions argues that emotions are learned, that we have agency over our emotions and that our emotions are goal-based. Barrett suggests that humans are unique in their ability to create and name (with words) abstract goal-based mental concepts such as 'wealth' and that emotion concepts such as 'desirability' are only goal-based. By reading an instance of an emotion word, we are primed to recognise that emotional concept in others and experience it ourselves. Furthermore, emotional concepts may be categorised and shared. As Barrett explains, *make something up, give it a name, and you've created a concept. Teach your concept to others, and as long as they agree, you've created something real* - an emotion. The more granular a person's emotions are, that is being able

to name different affective states, the better life outcomes they experience, including being generally healthier.

The implications of the theory of constructed emotions are profound. Not only can we create and have agency over our own emotions, but we can also, according to Barrett *be an architect of other people's experiences.* As Barrett explains the theory of constructed emotions, in brief: *Right now, as you read these words, your brain is wired with a powerful conceptual system for emotion. It began purely as an information-gaining system, acquiring knowledge about your world through statistical learning. But words allowed your brain to go beyond the physical regularities that you learned, to invent part of your world, in a collective with other brains. You created powerful, purely mental regularities that helped you control your body budget in order to survive. Some of these mental regularities are emotion concepts, and they function as mental explanations for why your heart thumps in your chest, why your face flushes, and why you feel and act the way you do in certain circumstances. When we share those abstractions with each other, by synchronising our concepts during categorisation, we can perceive each other's emotions and communicate. That, in a nutshell, is the theory of constructed emotion - an explanation for how you experience and perceive emotion effortlessly without the need for emotion fingerprints.*

The main idea of the theory of constructed emotions is that emotions are abstract concepts that we create, share and agree upon. Their purpose is to predict future states. Based on context, emotions guide the interpretation of but are separate from, individual affect or feelings.

Barrett shows that interpreting facial expressions as expressions of emotion depend on context. She shows a small photo of Serena Williams' face that initially looks to be an angry expression. The small photo is a cropped version of a larger photo taken Williams won a tennis final at Wimbledon. In the larger photo the face is seen to be a triumphant roar of joy.

2.2. Why requirements for certain applications need emotion words

Early software engineering focussed on getting functional requirements correct. It was soon realized that concentrating on functionality was not enough. As argued by many, including [14, 15, 16], it was important to also give attention to non-functional requirements. Whether non-functional requirements were met had a huge impact on whether a particular software system would be successful. Note we prefer to use the term quality requirements rather than non-functional requirements as they represent desired qualities [17] of a system.

As claimed by Stol and Fitzgerald in a 2018 article [18], awareness of software engineering as a social and multidisciplinary activity has increased. It requires some adjustment to take the social aspect of socio-technical systems seriously. This has led to a consideration of socially oriented requirements [19, 20] as a topic to be considered in its own right. An important aspect of socially oriented requirements are emotions that people may or may not feel when interacting with software. Another aspect is whether software systems embody values. In a talk at the 2019 Requirements Engineering conference, Professor Jon Whittle argued that requirements engineering methods should place greater emphasis on human values in order to build software that aligns with our individual, corporate or societal values. Now more than ever we are seeing a blurring of the lines between social sciences and software engineering. Software developed today incorporates and adapts to our values, attitudes, emotions, behaviours, amongst others. We need to improve our techniques for empirically reasoning about these concepts, and then

ensure they are effectively addressed in design.

Of course, this is not a new phenomenon. A 1946 quote from Lewis Mumford in 'Values for Survival' [21] is apposite. *We must give as much weight to the arousal of the emotions and to the expression of moral and aesthetic values as we now give to science, to invention, to practical organization. One without the other is impotent.*

The emotional aspects of digital products are undeniably important; however, they are often overlooked or ignored as system requirements or goals. There are numerous reasons. Firstly, emotions like desirability are difficult to elicit and measure, as they are subjective. Secondly, unlike the repeatable and logical way computers operate, emotional decisions are often based on, or exhibit logical fallacies that are not easily incorporated into standard system requirements models. Thirdly, system requirements models were created by engineers, who consider requirements as objective [22]. Qualities such as aesthetics and emotion that may map from many goals to many features are figuratively and literally 'fuzzy' to engineers who depict them as clouds and describe them as 'soft goals'. Therefore, they are incompatible with the engineering mindset that seeks cognitive 'closure' and fosters a script of strict 'rational' control of processes [23].

The need for requirements including emotion words is clearly illustrated with a personal anecdote from twelve years ago. Ruth, the mother of the first author, had a fall in the house where she had been living alone for the thirty years since her husband had died and her children had left home. Her children insisted that Ruth needed to use one of the emergency alarm systems that monitor safety at home. With great reluctance and significant protest, Ruth agreed. She quipped "You are making me wear a cowbell." Some time after Ruth passed away, a conversation occurred between the first author and the company responsible for the software in the alarm system. It was pointed out that the system had not been well received by Ruth who felt that it was not emotionally supportive. The software company was stunned as no-one had said anything like that to them previously. They defended themselves by stating that there was no requirement when designing the system that the older person should feel independent and cared for. Perhaps not, but is it clear that there should have been. More about emotional requirements and emergency alarm systems is discussed in [24].

Another project where emotional goals were part of the story was a research project funded by the National Health and Medical Research Council of Australia. The research team performed randomized, clinical trials as to whether electronic sites can give better outcomes for patients with depression than current standard consultations with general medical practitioners [25]. Between 24 and 55 percent of patients attending primary care internationally have depressive symptoms. Currently there is a mismatch between patient needs and the treatment received. Patients with sub-threshold or mild depression, who are likely to recover spontaneously, are often treated with antidepressant medication. On the other hand, patients with severe symptoms often do not receive minimally adequate treatment.

Building on a 10 year study that tracked over 500 people with depressive symptoms to document and describe the nature and course of depression [26], a pre-screening tool was designed and implemented on an iPad to assess potential participants in the study as to whether their mental health symptoms are low-risk, moderate risk or high risk for developing into serious depression. The tool consisted of a number of questions that the medical team consider are a good indication of depression, based on the analysis of data collected over a decade through

the depression study. Two design concepts were developed. The designs included icons relating to each question on the app and an assortment of ways of presenting results. It was striking how the different design concepts engendered different emotional reactions from the focus group participants. One of the design concepts was clearly preferred.

The iPad app was subsequently designed and developed so that interactions with the tool were positive. It needed to be sufficiently appealing to engage people who are unaware that they are at risk of depression and/or have negative attitudes towards the diagnosis of depression. It also needed to deliver the results in a way that does not lead to panic but to discussion with their doctor. The app was piloted through two focus groups with eight participants who were potential end users of the app. Completing the app took participants between two and three minutes. None of the participants had to be instructed in how to use the app. Notably, participants did not report any negative aspect of using the app. All participants explicitly said it was an easy, neutral experience. During piloting, the app was unobtrusive and did not detract from the medical considerations. Participants indicated that the app could aid their self reflection, provide hope, and motivate them to seek treatment.

The research team considered it a success when the focus group said the tool felt authoritative and the patients taking the questionnaire felt more encouraged to take action to support their mental health. Patients reported that they wanted the app to make them feel supported, comfortable and that the information feel relevant and important, and they wanted to see their results. Doctors wanted to have confidence that the app provided reliable information, to have it feel professional and useful for improving depression care, and researchers wanted the app to be scientifically reliable trustworthy and useful for improving depression care. In the second focus group, the look and feel of the app was appreciated. Participants expressed a desire for a positive message tailored to individual results and incorporating a treatment recommendation. In other words, their emotional needs had been catered for in their interaction with the app. One finding was that the system needed to report an outcome about their health. Participants expect a quid pro quo for providing information. More details of the study can be found in [25] including quotes from the participants which indicated such emotional engagement as hope.

Recent research in using electronic health records for self-managing health has shown that patients wanted to feel empowered, in control and resilient, while maintaining meaningful connections with family and carers [27]. Current solutions fail to adequately address these emotional goals. Citizens have been confronted with a platform which they refuse to trust with their personal data.

Many different types of emotions are important to consider when designing technology. For example, commonly recognised western emotions such as joy can be a core goal of many software systems. A major motivation of computer games or software systems that utilise gamification is to create a fun and delightful experience. In the agent literature, emotions have focussed on the believability of the agent. In the heyday of e-commerce innovation, there was discussion of stickiness of sites and whether the agents increased engagement. Of course the aspects of stickiness have been refined to a fine art by social media applications such as Facebook and YouTube.

In our research, we address the need for emotion words with motivational models. In brief, a motivational model is a high level diagram describing an overall socio-technical system, a software product or a company. Motivational models are the evolution of goal models as

described in [3]. The intent underlying the creation of a motivational model is to capture a shared understanding of the system, its goals and stakeholders. The motivational model diagram captures what the system to be is being designed to do, what qualities the system should have such as security and privacy, and how interactions with the system should feel for key stakeholders. What the system should do, effectively functional requirements, and how the system should be, effectively quality requirements, are reasonably standard. Including goals for emotional responses at the highest level of the project is relatively new.

Representing emotional goals at a high level of abstraction presents a challenge. They are inherently ambiguous and subjective, challenging to address in design, and difficult to evaluate. An example of an emotional goal is for people to feel empowered while interacting with software. The emotion of feeling empowered is a property of a person and not of software which partially explains why emotions have not traditionally been taken into account. Yet emotional goals are essential as people tend to reject software that does not adequately support the way they wish to feel during interactions.

User emotions may be characterised as soft goals [16] but are only indirectly represented using traditional software engineering methods and are typically ignored even if raised. Explicit emotional goals are discussed by Marshall [28] as separate from other quality goals to both promote and represent emotional expectations of users and other stakeholders. It can be argued that emotional goals are different from traditional quality goals such as performance and reliability because they are properties of the user rather than the system.

An example of an emotional goal from our case study of emergency alarm systems [24] is that the relatives of older people want to feel reassured that their relative is safe. Relatives feeling reassured is hard to consider being a property of the system. Measuring such feelings requires us to measure the emotional state of people (e.g. representative users) unlike other qualities; for example, performance, which can be measured directly, or even usability, which can often be measured using well-defined metrics and qualitative observation from system logs. We advocate that emotional goals be regarded as first-class objects in software engineering and be treated separately from other quality goals.

Let us consider emotions. Do existing software engineering techniques effectively translate emotional goals and requirements into design? We contend that requirements relating to emotions differ from traditional functional and non-functional requirements. Emotional goals, such as the goal of feeling empowered while interacting with software, is a property of a person and not of software. Emotional goals are inherently ambiguous, subjective, difficult to elicit, difficult to represent, difficult to address in design, and difficult to evaluate. Existing artifacts that capture soft goals include use cases, personas, scenarios or cultural probes. However, these alone are still insufficient when designing for technology embedded within complex social situations.

Talking about how software needs to cater to the emotional needs of its users from is powerful from a psychological perspective. Indeed our involvement in several projects came about because the project leads were concerned about engaging with users' emotions, a sentiment not commonly expressed. Exactly defining emotions, however, is not necessary in our experience. Emotions are not clear-cut and hence emotional requirements are ambiguous. The ambiguity may lead them to be sidelined. We believe that emotions are important to address and need to be integrated with the elicitation of functional and other quality requirements.

There is occasional confusion between qualities and emotions, compounded by the fact that the same word can appear both as a quality goal and an emotional goal. An example might be from an electronic banking app which both needs to be secure and feel secure. They generate different requirements. Consider the quality of security for the electronic banking app which would need to be realised by appropriate encryption algorithms. On the other hand, the system needs to feel secure by looking professional and not asking for inappropriate information. Encryption algorithms and professional look and feel are two separate issues to address in the software. In practice confusion between qualities and emotions does not cause problems in the elicitation process in our experience.

To conclude this subsection we again quote from Barrett [13]. *I hope by now you appreciate the drama that is going on here. Emotion words are not about emotional facts in the world that are stored like static files in your brain. They reflect the varied emotional meanings you construct from mere physical signals in the world using your emotion knowledge. You acquired that knowledge, in part, from the collective knowledge contained in the brains of those who cared for you, talked to you, and helped you to create your social world. Emotions are not reactions to the world; they are your constructions of the world.*

2.3. Interdisciplinary teams and the need for a shared understanding

Software development teams are increasingly multi-disciplinary due to the expansion of application domains. The increase in disciplines places greater demands on communication as less knowledge can be assumed. Both technical and non-technical team members are able - and often expected - to contribute to ongoing discussions and decisions about requirements. Agile practices have become widespread. They require lightweight models encapsulating requirements and designs to promote shared understanding amongst the team members. The process for acquiring requirements has shifted. Rather than having a requirements engineer interview stakeholders with a business need, more collaborative processes are used. Design thinking has become more prominent, with participatory and co-design approaches being used to help elicit requirements. Co-design is important for consumer applications as they are more able to explore emotional requirements which can relate to diverse social, cultural, organisational and political situations around software use. Examples are wellness apps and services for the homeless. Externally, more software projects are coordinating with diverse stakeholder groups. For eGovernment and eHealth solutions, multiple distinct organisations are needed to provide input on the software design. Achieving consensus is difficult as people are no longer forced to use systems as they once were with business software. Wider engagement of stakeholders, both internal and external, has made software development more challenging.

In many instances it is impossible to control all of the factors which used to be part of development. For example, if developing an app for homelessness which needs to accommodate government policies, the developer cannot control how the government policies will be introduced or managed or how agencies will interpret them. Another complicating factor is that many fundamental requirements are not known in advance, but instead emerge over time through discussion and deliberative processes.

In grappling with the changing environment and demands, we have seen a need for improved support for discussion, deliberative dialogue and sense-making. Discussing emotions increases

engagement. One response to our research has been that discussing emotions is not a responsibility of the software engineer. Rather, ensuring that software engenders appropriate emotional responses will be taken into account by the user experience team. User experience does need to take emotional elements into consideration. However it is our view that emotions need to be considered by the whole team, with the high-level goals being clearly articulated and with the understanding that unless these goals are met, no goals are met. For example if the goal is to create software that people use, and people don't use the software, all other goals, such as how well the software is programmed, are redundant and worthless.

A positive emotional experience cannot always be easily added if it has not been included in the design. From requirements methods, user stories can address the emotional needs as can personas, but they will not be considered unless they are explicitly goals of the project. Everyone needs to be on the same page. An important component of ensuring everyone is aligned is awareness of whether the emotional goals are being addressed. It is important that the whole project team shares an understanding of the project which leads to the next subsection. As reported in [29], software developers failed to understand the need for engagement in one version of the software. The software needed to be discarded and a new software team assembled. It was too late to have UX designers salvage software that did not promote the needed engagement from students.

A key part of our process of developing models for requirements gathering is do/be/feel elicitation. [30]. As a group the stakeholders produce lists of words that the software or service is striving to be. The effect is to share understandings of words during the elicitation. Further there are echoes of Barrett's theory of constructing an emotion together.

3. Artificial agents are not emotional

As inexperienced parents, the first author and his wife struggled with responses to comfort their eldest daughter Danya when she cried in the first few months of her life. Unfortunately, it was a common experience. Also unfortunately, it was distressing when Danya continued to cry.

A surprising strategy that we happened upon was to hold up a toy rubber hammer in front of Danya. The hammer was orange and squeaked if it hit someone on the head or any hard surface. The hammer also had a smiling face painted on it. When seeing the face on the hammer, Danya would start laughing, and many times that enabled her to settle. Danya's reaction was undoubtedly emotional, and was reinforced over repeated use.

To contrast from the next section, Danya's behaviour was not rational. Incidentally the first author's other two children did not react in the same way. There was something individual and contextual about the response.

Now consider a Nao robot. Nao robots have vision capabilities. They could be programmed to laugh when they see a face painted on an orange rubber hammer. We contend that a Nao robot laughing on recognising the face on the hammer would not be an emotional reaction. The laughing is of no benefit to the Nao. Holding up a hammer in front of a Nao would not be a method for it to deal with discomfort in the world around it. Nor would it be remembered as an element of comfort.

Robots elicit emotions, but that is because of people's inherent tendency to anthromorphise

and treat objects as social beings. Even robots that do not resemble humans, such as Roomba, ElliQ, Starship robot, and AIBO can bring about strong emotional reactions in humans. For example, on the 10th of January 2017 an incident was reported in Estonia, where a drunk man attacked a Starship delivery robot, kicking it several times by leg. Having robots elicit emotions is a case of design.

Nao robots are interesting to consider in the context of artificial agents. They have been very well designed and can express some personality and mimic emotions in ways which can be helpful. A Nao robot was used successfully in a project with the Royal Children's Hospital in Melbourne to encourage children to perform their physiotherapy exercises [31]. The robot was programmed to be encouraging and make jokes so that the child would be more motivated to do exercises. Throughout the project there were many occasions when the Nao robot would perform a choreographed dance to the Korean pop music hit 'Gangnam Style.' The robot performance would invariably cause the people watching to smile. Frankly the robot looked cute and ridiculous when performing.

Nao robots are best used as a way of engaging humans in activities [32]. However they have been proposed as a means to help autistic children recognise emotions. It is important to appreciate that the robots are not feeling the body sensations that people do. We do not believe it is helpful to describe the robots in emotional terms. Any identification of concepts is by the programmer/designer, not by the robot. To relate it to constructed emotion theory, the robot can use an emotion word but is not participating in the same emotional experience as the human.

To summarise, when considering emotions in the context of artificial agents, we are not conceiving of the agent as having feelings or emotions. The emotions may be engendered in the person interacting with the artificial agent. It is a design challenge to ensure that the person experiences the feelings associated with the emotion words.

4. Humans are not rational

It has been common to describe cognitive thinking as the pinnacle of human thought, more advanced than decision making determined by feelings and emotions. More recent thinking from neuroscience has been challenging that perspective. Kahneman describes two versions of thinking in his well-cited book 'Thinking Fast and Slow' [9]). The new perspective has emotional thinking and cognitive thinking more in partnership.

A field that was quick to describe people as rational decision makers is economics. However pioneering researchers Kahneman and Tversky showed that we have biases. We behave differently depending on how information is presented. Many examples are given in [9]). The updated perspective of behavioural economics has been nicely described in 'Predictably Irrational' [33].

That people do not behave rationally has been largely ignored by intelligent agent research. A common perspective for agent research is the BDI architecture, purportedly abstracted from human beliefs, desires and intentions. Though it is unclear that there is any support from the psychology literature.

As discussed by Barrett among others, the brain is good at making predictions. The brain has understanding of likely situations and reacts to changed circumstances. The reaction is through

fast associations, not always through cognitive reasoning. Babies and young children learn through reacting and adjusting to external stimuli quickly. Later cognitive reasoning builds on top of such brain behaviour.

A recent popular science book is 'the Extended Mind' by Annie Murphy Paul [34]. The book contends that there is more than just the brain involved in decision making. There is non conscious information acquisition happening throughout the body. A related fact is that our introspection about how we make decisions is not reliable. We are good at telling stories, and kidding ourselves.

Another point made in 'The Extended Mind' is that we are affected by our environment, whether we are moving or stationary and other factors. We tend to identify with our brain being a computer making decisions. In the conclusions Paul writes: *we should seek to productively alter our own state when engaging in mental labor. We've repeatedly confronted the limits of the brain-as-computer analogy, ... When fed a chunk of information, a computer processes it in the same way on each occasion - whether it's been at work for five minutes or five hours, whether it's located in a fluorescent-lit office or positioned next to a sunny window, whether it's near other computers or is the only computer in the room. This is how computers operate, but the same doesn't hold for human beings. The way we're able to think about information is dramatically affected by the state we're in when we encounter it.*

It is perhaps an over simplification. My laptop performs differently often depending on load and network issues. But researchers maintain that variability is a problem. The book 'Noise' [35] argues that we should remove human bias in favour of rational decision making. But being human entails subjectivity and designing socio-technical systems needs to take human biases into account. Of course we don't want biased artificial agents. So human agents and artificial agents need different consideration when designing a socio-technical system.

The masters research of Nicole Ronald at the University of Melbourne involved the modelling of pedestrians. Nicole had come from an engineering approach to model pedestrian flow and wondered whether modelling pedestrians as BDI agents with desires to shop or browse while walking would give a better simulation result. The research is reported in [36]. It turned out BDI agents and the system architecture did not easily capture what was being modelled. The reasoning behaviour of BDI agents did not adjust easily to modelling real people walking. Agent-based simulations can be valuable but only in simulating part of human behaviour.

5. A role for design in meeting emotional goals

What are the implications of the discussion above? It is important to include emotion words in developing software, services and other products. It is a challenge to ensure the software or service engenders the desired emotion. However it is a challenge designers are trained to undertake. We note that the first step in the popular five stop process for design thinking is to empathise with your client, something not usually done for software development.

Professional designers articulate emotional goals as higher-level objectives, and try to align with the desires, needs and emotions of users. They are conveyed in brand values, marketing material and are used to inform key design decisions. Hitting the right emotional tone is part of empathising with the customer and user — a key step in design thinking and broadly implicit in

all professional design approaches. However designing for emotions is more subtle, as it also depends on factors such as environment, context, and social influence. The construction of emotions requires both the push of design intention and a pull of empathy, in a feedback loop that constructs a collective intention.

Emotions play a key role in design. Famously, Apple designers ask how do we want people to feel? [37]. In addition Hartmut Esslinger, a designer of Apple products claims as described in [38] that 'Form follows Emotion.' How emotions are used in digital media design capstone projects has been presented in [28].

An interesting example of a different way of developing a product is the drawing program Kid Pix which engaged children by design. It was developed in 1988 by the artist and programmer Craig Hickman, in partnership with his three-year-old son Ben. Kid Pix broke rules of usability, such as having seemingly redundant tools. For example, Kid Pix had eight different erasers, including a drain and an exploding firecracker because Craig discovered that erasing drawings was part of the fun. Kid Pix is a seminal example of emotive meaningful software. A quote from musician Brian Eno typifies reactions to the program *I keep showing people Kid Pix as the epitome of what I mean: how it produces total delight in almost everyone almost straightaway is a miracle of Design.* [39].

Referring to emotions happens despite the lack of consensus in exactly what emotions are. Some believe in a hierarchy of emotions, building from basic emotions such as fear, anger or joy. Others believe that emotions are constructed concepts developed through life experience. We advocate for being able to address emotion goals as software requirements.

6. Conclusions

We regard emotional goals as being consistent with the theory of constructed emotion as articulated and popularised by Lisa Feldman Barrett [13]. Running a do/be/feel elicitation as part of motivational modelling which is reinforced through design is a way of constructing an emotion concept. The emotion word is agreed to by all stakeholders through a motivational model.

Our approach to emotional goals differs from affective computing. In the latter, the system is concerned with detecting and expressing emotions. In our methods, emotions are constructed terms for aspirational goals that are intended for the user to feel when interacting with a system.

Artificial agents and robots can be part of a socio-technical system. But they do not feel or experience emotions in the same way that people do. The use of emotion words in requirements is to help people engage with products.

Further the behaviour of people and artificial agents are not the same. We advocate logical rules for artificial agents, but appreciate people are not purely logical in behaviour. It is a challenge for the designer of a socio-technical systems to model the various agents appropriately.

Acknowledgments

The authors would like to acknowledge discussions with members of the Future Self and Design Living Lab at Swinburne University of Technology. The research was partly funded by the ARC

DP200102955, ”Maturing design-led innovation processes with motivational models”.

References

[1] M. Wooldridge, An Introduction to Multiagent Systems, 2 ed., Wiley, Chichester, UK, 2009.

[2] S. J. Russell, P. Norvig, Artificial Intelligence: a modern approach, 3 ed., Pearson, 2009.

[3] L. Sterling, K. Taveter, The art of agent-oriented modeling, MIT Press, 2009.

[4] D. A. Mendoza, D. T. Miller, D. S. Pedell, P. Sterling, et al., The role of users' emotions and associated quality goals on appropriation of systems: two case studies (2013).

[5] R. W. Picard, Affective Computing, MIT Press, Cambridge, MA, USA, 1997.

[6] J. Bates, The role of emotion in believable agents, Commun. ACM 37 (1994) 122–125. URL: https://doi.org/10.1145/176789.176803. doi:10.1145/176789.176803.

[7] M. Scheutz, Agents with or without emotions? (2002).

[8] J. Dewey, Art as experience, George Allen and Unwin Ltd, 1934.

[9] D. Kahneman, Thinking, fast and slow, Farrar, Straus and Giroux, New York, 2011. URL: https://www.amazon.de/Thinking-Fast-Slow-Daniel-Kahneman/dp/0374275637/ref=wl_it_dp_o_pdT1_nS_nC?ie=UTF8&colid=151193SNGKJT9&coliid=I3OCESLZCVDFL7.

[10] R. Sapolsky, Behave: The Biology of Humans at Our Best and Worst, Penguin Press, 2017.

[11] P. Ekman, Emotions revealed : recognizing faces and feelings to improve communication and emotional life, New York : Times Books, 2003.

[12] R. Firth-Godbehere, A Human History of Emotion: How the Way We Feel Built the World We Know, HarperCollins Publishers, 2021. URL: https://books.google.com.au/books?id=xx8mEAAAQBAJ.

[13] L. F. Barrett, The theory of constructed emotion: an active inference account of interoception and categorization, Social cognitive and affective neuroscience 12 (2017) 1–23.

[14] L. Chung, B. A. Nixon, E. Yu, J. Mylopoulos, Non-functional requirements in software engineering, volume 5, Springer Science & Business Media, 2012.

[15] I. Ramos, D. M. Berry, Is emotion relevant to requirements engineering?, Requirements Engineering 10 (2005) 238–242.

[16] S. Thew, A. Sutcliffe, Investigating the role of 'soft issues' in the re process, in: 2008 16th IEEE International Requirements Engineering Conference, 2008, pp. 63–66. doi:10.1109/RE.2008.35.

[17] L. Bass, P. Clements, R. Kazman, Software architecture in practice, Addison-Wesley Professional, 2003.

[18] K.-J. Stol, B. Fitzgerald, The abc of software engineering research, ACM Trans. Softw. Eng. Methodol. 27 (2018). URL: https://doi.org/10.1145/3241743. doi:10.1145/3241743.

[19] G. Baxter, I. Sommerville, Socio-technical systems: From design methods to systems engineering, Interacting with Computers 23 (2010) 4–17. URL: https://doi.org/10.1016/j.intcom.2010.07.003. doi:10.1016/j.intcom.2010.07.003. arXiv:https://academic.oup.com/iwc/article-pdf/23/1/4/2038336/iwc23-0004.pdf.

[20] S. Pedell, T. Miller, F. Vetere, L. Sterling, S. Howard, Socially-oriented requirements engineering: software engineering meets ethnography, in: Perspectives on Culture and Agent-based Simulations, Springer, 2014, pp. 191–210.

[21] L. Mumford, Values for Survival: Essays, Addresses, and Letters on Politics and Education, Essay index reprint series, Books for Libraries Press, 1972. URL: https://books.google.com.au/books?id=vlgZNQAACAAJ.

[22] P. Hodges, S. Ruecker, C. Scaletsky, J. Rivera, R. Faller, A. Geppert, Four criteria for design theories, She Ji 3 (2017) 65–74. doi:10.1016/j.sheji.2017.02.003, publisher Copyright: © 2017 Tongji University and Tongji University Press Copyright: Copyright 2019 Elsevier B.V., All rights reserved.

[23] D. Gambetta, S. Hertog, Why are there so many engineers among islamic radicals?, European Journal of Sociology/Archives Européennes de Sociologie 50 (2009) 201–230.

[24] T. Miller, S. Pedell, A. A. Lopez-Lorca, A. Mendoza, L. Sterling, A. Keirnan, Emotion-led modelling for people-oriented requirements engineering: the case study of emergency systems, Journal of Systems and Software 105 (2015) 54–71.

[25] C. Wachtler, A. Coe, S. Davidson, S. Fletcher, A. Mendoza, L. Sterling, J. Gunn, et al., Development of a mobile clinical prediction tool to estimate future depression severity and guide treatment in primary care: user-centered design, JMIR mHealth and uHealth 6 (2018) e9502.

[26] J. Gunn, colleagues, Diamond study on depression, 2015. URL: http://purl.org/au-research/grants/nhmrc/1002908.

[27] K. Taveter, L. Sterling, S. Pedell, R. Burrows, E. M. Taveter, A method for eliciting and representing emotional requirements: Two case studies in e-healthcare, in: 2019 IEEE 27th International Requirements Engineering Conference Workshops (REW), IEEE, 2019, pp. 100–105.

[28] J. Marshall, Agent-based modelling of emotional goals in digital media design projects, in: Innovative Methods, User-Friendly Tools, Coding, and Design Approaches in People-Oriented Programming, IGI Global, 2018, pp. 262–284.

[29] L. Sterling, S. Pedell, G. Oates, Using Motivational Modelling With an App Designed to Increase Student Performance and Retention, in: D. Glick, A. Cohen, C. Chang (Eds.), Early Warning Systems and Targeted Interventions for Student Success in Online Courses, IGI Global, Hershey, PA, USA, 2020, pp. 161–176. URL: http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-7998-5074-8.ch008. doi:10.4018/978-1-7998-5074-8.ch008.

[30] A. Lopez-Lorca, R. Burrows, L. Sterling, Teaching motivational models in agile requirements engineering, in: Proceedings of the Requirements in Education and Training workshop at RE'18, 2018.

[31] F. Martí Carrillo, J. Butchart, S. Knight, A. Scheinberg, L. Wise, L. Sterling, C. McCarthy, Adapting a general-purpose social robot for paediatric rehabilitation through in situ design, ACM Transactions on Human-Robot Interaction (THRI) 7 (2018) 1–30.

[32] S. Pedell, K. Constantin, D. Muñoz, L. Sterling, Designing meaningful, beneficial and positive human robot interactions with older adults for increased wellbeing during care activities, in: Handbook of Artificial Intelligence in Healthcare, Springer, 2022, pp. 85–108.

[33] D. Ariely, Predictably Irrational: The Hidden Forces That Shape Our Decisions, Harper-Collins Publishers, 2008.

[34] A. Paul, The Extended Mind: The Power of Thinking Outside the Brain, Houghton Mifflin Harcourt Publishing Company, 2021. URL: https://books.google.com.au/books?id=Dk-_

DwAAQBAJ.
[35] D. Kahneman, O. Sibony, C. R. Sunstein, Noise: a flaw in human judgment, Little, Brown Spark, 2021.
[36] N. Ronald, L. Sterling, M. Kirley, et al., An agent-based approach to modelling pedestrian behaviour, International journal of Simulation 8 (2007) 25–38.
[37] Apple, Apple manifesto, 2013. URL: http://www.apple.com/designed-by-apple.
[38] O. Demirblek, B. Sener, Product design, semantics and emotional response, Ergonomics 46 (2003) 1346–1360.
[39] B. Eno, A Year With Swollen Appendices, Faber and Faber, 1996.