

9

Emotion

What is covered in this chapter:

- The difference between affect, emotions, and mood.
- What roles emotions play in interacting with other humans and robots.
- Basic models of emotions.
- The challenges in emotion processing.

How are you feeling right now? Happy? Bored? A bit self-conscious? Whatever the case may be, it's unlikely that you are feeling absolutely nothing. Various feeling states and their related emotions are a key aspect of our day-to-day experience and of our interactions with other people. Emotions can motivate and modulate behavior and are a necessary component of human cognition and behavior. They can be spread through vicarious experience, such as watching a tense movie, and direct social interaction, such as seeing your best friend happy. Because emotions are such an integral part of human social cognition, they are also an important topic in human–robot interaction (HRI). Social robots are often designed to interpret human emotion, to express emotions, and at times, even to have some form of synthetic emotion driving their behavior. Although emotions are not implemented in each and every social robot, taking emotions into account in the design of a robot can help improve the intuitiveness of the HRI.

This chapter starts with an overview of what researchers mean when they talk about emotions ([Section 9.1](#)), along with the importance of emotions in social interaction ([Section 9.2](#)). In [Section 9.3](#), we turn to how emotions are processed in HRI. [Section 9.4](#) covers the challenges related to robots' understanding, processing, and expressing of emotion during HRI.

9.1 What are emotions, mood, and affect?

From an evolutionary perspective, emotions are necessary for survival because they help individuals respond to environmental factors that either promote or threaten survival (Lang et al., 1997). As such, they prepare the body for behavioral responses, help in decision-making, and facilitate interpersonal interaction. Emotions arise as an appraisal of different situations that people encounter and prepare us for a response (Gross, 2007; Lazarus, 1991). For

example, when another person shoves us out of the way to be first in line, we get angry, and our bodies prepare for a potential conflict: the adrenaline makes us more prone to undertake action, and our expression signals to the other person that he or she crossed a line. Conversely, upon finding out our friend did not invite us to his or her birthday party, sadness hampers quick action, forcing us to reconsider our prior behavior (i.e., what did we do or say that may have offended him or her?) and evokes empathetic responses from others (Bonanno et al., 2008). In this way, emotions can also help us modulate the behaviors of others in an interaction.

Affect is used as a comprehensive term that encompasses the entire spectrum of emotionally laden responses, ranging from quick and subconscious responses caused by an external event to complex moods, such as love, that linger for longer (e.g., Lang et al., 1997; Bonanno et al., 2008; Beedie et al., 2005). Within affect, a distinction is made between emotions and moods (Beedie et al., 2005).

Emotions are usually seen as being caused by an identifiable source, such as an event or seeing emotions in other people. They are often externalized and directed at a specific object or person. For example, you experience happiness when getting a promotion at work, get angry when your phone's battery dies during an important call, or experience a pang of jealousy when a colleague gets a company car and you do not (Beedie et al., 2005). Emotions are also shorter-lived than moods (Gendolla, 2000). *Moods* are more diffuse and internal; often lack a clear cause and object (Ekkekakis, 2013; Russell and Barrett, 1999); and instead are the result of an interaction between environmental, incidental, and cognitive processes—such as the apprehensive mood while waiting a week to hear about the medical test results or the warm feeling of a sunny week spent in the company of friends.

9.1.1 Emotion and interaction

Emotions are not just internal; they are also a universal communication channel that has helped us communicate internal affective states to others and have likely been very important to our survival as a species.

Your emotions provide the outside world with information about your internal affective state, which is helpful to others in two ways. First, emotions convey information about you and your potential future actions. For example, displaying anger and frustration signals to others that you may be preparing for an aggressive response. In addition, emotions can convey information about the environment. An expression of fear may alert others around you of a fast-approaching grizzly bear before you have even found time to scream. In both scenarios, emotion provides an incentive for others to take action. In the case of anger, someone may choose to step down and attempt to suss the situation. In the case of fear, other people will likely scan the environment for a threat (Keltner and Kring, 1998). In this way, the successful communication of emotions promotes survival, enhances social bonds, and minimizes the

chances of social rejection and interpersonal physical aggression (Andersen and Guerrero, 1998).

9.1.2 Conceptualizing human emotions

Since antiquity, people have given names to the numerous emotions we experience. Aristotle believed there to be 14 different emotions, including anger, love, and mildness. More recently, Ekman listed 15 basic emotions, including pride in achievement, relief, satisfaction, sensory pleasure, and shame (Ekman, 1999). For various reasons, it is impossible to provide a definite list of emotions: for example, they vary between people and cultures, language does not offer a perfect mapping to emotions, and some emotions show overlap. Still, some emotions are likely to be considered more universal than others. Anger, sadness, and happiness are likely candidates for a set of core emotions. Ekman and Friesen (1975), in their seminal work on the facial expression of emotions, listed six basic facial expressions that are recognized across cultures. These facial expressions have often been mistaken for a set of basic emotions we experience, although they were only ever intended to describe a basic set of emotions that we express via our faces and that are recognized by different cultures.

Although many scholars distinguish between basic, or *primary*, emotions and reactive, or *secondary*, emotions, no consensus has been reached yet on which emotions are to be included in the first category and which should be considered secondary (Holm, 1999; Greenberg, 2008), and some scholars argue that basic emotions do not exist at all (see, e.g., Ortony and Turner, 1990). For those who do agree on the existence of basic emotions, primary emotions are considered to be universal across cultures (Stein and Oatley, 1992) and to be quick, gut-level responses (Greenberg, 2008) and include emotions such as amusement, anger, surprise, disgust, and fear. Secondary emotions, on the other hand, are reactive and reflective. They differ across cultures (Kemper, 1987). For example, pride, remorse, and guilt are secondary emotions.

But there have been challenges to the idea of emotions being distinct categories. Russell (1980) argued that emotions are the cognitive interpretations of sensations that are the product of two independent neurophysiological systems, namely, arousal and valence. As such, emotions are spread across a two-dimensional continuum rather than being composed of a set of discrete, independent basic emotions (Posner et al., 2005) (see Figure 9.1). This model has been widely studied and confirmed to hold across different languages and cultures (Russell et al., 1989; Larsen and Diener, 1992). However, a meta-analysis found that although the model makes for a reasonable representation of self-reported affect, not all affective states fall into the expected regions as predicted by the theory, and some cannot even be consistently ascribed to any of the regions, suggesting that assumptions about the nature of some affective states may need to be revised (Remington et al., 2000).

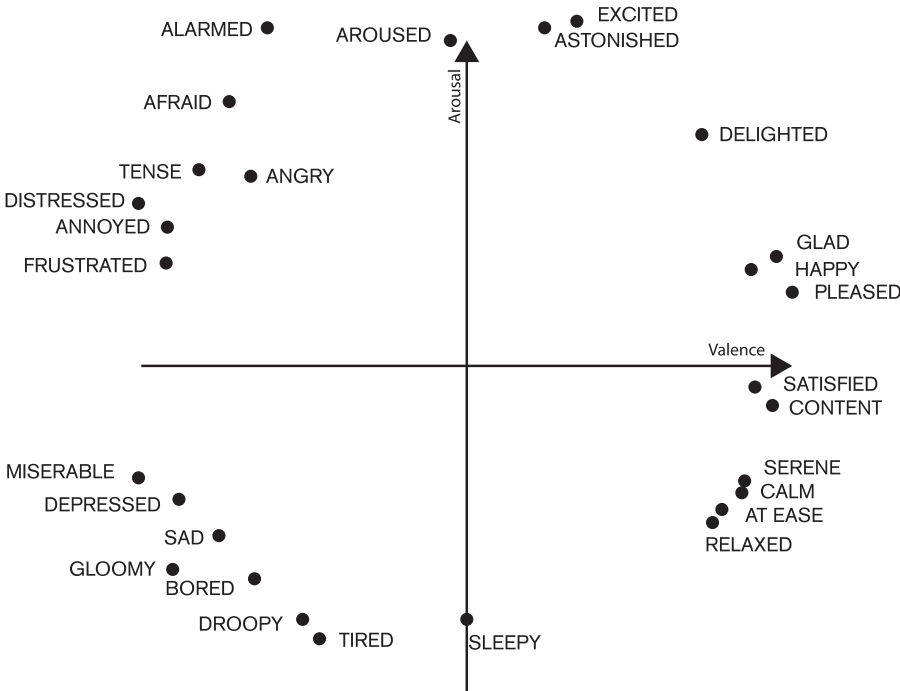


Figure 9.1
Russell's (1980)
circumplex model
of affect.

Expressing emotions does not just inform others on how you feel—it may actually inform *you* on how you feel. The *facial feedback hypothesis* proposes that facial movement influences the emotions we experience: when participants are forced to adopt a smile (by being told to hold a pencil between their teeth) before reading a comic, they rate the comic as slightly funnier than if they hold the pencil in their hand. If instead they have to hold the pencil between their lips, they find the comic less funny (Strack et al., 1988). In a similar way, administering Botox (which paralyzes facial muscles) has been found to reduce the intensity with which emotions are experienced (Davis et al., 2010).

9.2 When emotions go wrong

The importance of emotions in social interactions becomes especially clear when one partner fails to understand the emotion of the other partner or fails to respond with the proper emotion. Even tiny glitches in providing an adequate emotional response in social interaction can have serious consequences. For example, misinterpreting sarcasm for a genuine response can lead to misunderstandings in the conversation and hurt feelings. The situation becomes more problematic when someone is consistently unable to adequately perceive, express, or respond to affective states.

Problems with emotional responsiveness are one of the defining symptoms of, for example, depression (Joormann and Gotlib, 2010). Although depressed individuals are able to understand the way others are feeling and can express their own emotional state, they have a reduced emotional response to positive stimuli, such as rewards (Pizzagalli et al., 2009), and have recurring negative thoughts about the past, present, and future. As a consequence, a depressed individual's patterns of social interaction often result in social isolation and even more loneliness, feeding into the individual's already frail psychological state.

Furthermore, people might be incapable of recognizing, expressing, and interpreting another person's emotions. For example, people with autism spectrum disorders may find it difficult to correctly interpret displays of emotion (Rutherford and Towns, 2008; Blair, 2005) (see Figure 9.2). This is clearly problematic for everyday social interactions because the affected person cannot intuitively understand the needs of his or her interaction partners and will often respond inappropriately.

Furthermore, people may have trouble expressing their emotional state, for example, when their facial muscles are impaired after a stroke. This makes it hard for their interaction partners to infer their internal states and form an idea of what they mean.

A person's inability to express and interpret emotions comes with serious consequences for the individual's capability to either provide or respond to emotional cues in an appropriate way. This, in turn, impairs the capability to interact with other people effectively and smoothly. Likewise, social interactions with robots may be difficult if the robotic counterpart is unable to express and interpret emotional states.

9.3 Emotions for robots

Emotions are considered an important communication channel in HRI. When a robot expresses emotion, people tend to ascribe a level of social agency to it (Breazeal, 2004a; Novikova and Watts, 2015). Even if a robot has not explicitly been designed to express emotions, users may still interpret the robot's behavior as if it had been motivated by emotional states. A robot that is not programmed to share, understand, or express emotions will thus run into problems when people interpret its behavior as disinterested, cold, or plain rude. Therefore, engineers and designers should consider what emotions the robot's design and behavior convey, whether and how a robot will interpret emotional input, and how it will respond.

9.3.1 Emotion interaction strategies

The most straightforward way of programming emotional responsiveness for social robots may be through mimicry. Mimicking in humans has been shown to create an idea of shared reality: you indicate that you fully understand the other person's situation, which creates closeness (Stel et al., 2008). The exception here might be anger—however good it may feel at first, responding to an

Figure 9.2 Kaspar (2009– present) is a “minimally expressive” robot, built using brackets, servo motors, and a surgical silicon mask. Kaspar is used in autism therapy. (Source: Kerstin Dautenhahn, Ben Robins, Adaptive Systems Research Group, University of Hertfordshire, UK)



angry person by yelling back usually does not facilitate mutual understanding or a resolution of the conflict.

A robot can use mimicry as a simple interaction strategy. It is a relatively simple response because it requires the robot “only” to be capable of recognizing an emotion in the human and then reflecting the emotion back in response. This already poses plenty of challenges, as will be discussed later in this chapter, but at least it cuts out the complicated task of formulating an appropriate response. Moreover, it may be a very basic expectation that humans have toward their interaction partners. Although we may excuse our friends for not knowing how to cheer us up when we are sad, we do expect (and appreciate) that they will respond to our sadness by lowering their brows and heads and becoming more soft-spoken.

One note that has to be made here concerns expectation management. When users perceive the robot to be emotionally responsive, they may extend this observation to expectations about the robot’s compliance with other social norms. For example, a user may expect a robot to remember to ask about a confrontational meeting he was upset about the other night, so when the robot simply wishes him to “have a great day at work!” in the morning, he may be disappointed in the robot’s social skills. Thus, the robot’s emotional responsiveness should match its capability to fulfill other expectations.

9.3.2 *Artificial perception of emotions*

Robots need to register a wide variety of emotional cues, some explicit and some subtle, before being capable of emotional interaction. For instance, if we want to create a robot that responds emotionally when someone displays aggressive behavior, such as throwing an item at it, we need to integrate technologies for human behavior recognition and object recognition.

More specifically, we may want to create a robot that responds to human emotions. There are many studies on affect recognition (Gunes et al., 2011; Zeng et al., 2009). The most typical approach to recognizing or classifying emotions is to use computer vision to extract emotions from facial cues. Provided with a data set of human (frontal) faces with correctly labeled emotions, machine-learning systems, such as those using deep-learning techniques (LeCun et al., 2015), can extract features from the image to recognize a range of facial emotions. A famous example of this is smile recognition, which is broadly implemented in digital cameras nowadays. Affect recognition may also imply the interpretation of other visual cues, such as walking patterns, alleviating the need for a clear view of the user’s face (Venture et al., 2014).

Many consumer-market digital cameras have a smile-detection feature. If a group poses in front of the camera, it will only take a shot when all the people in the frame smile. This technology partly replaces the timer function, which could never guarantee that everybody would look at the camera and smile at the time of the picture being taken.

Next to visual cues, human speech is perhaps the second-most-important channel to extract emotion from. In particular, prosody, the patterns of stress and intonation in spoken language, can be used to read the emotional state of the speaker. For instance, when people are happy, they tend to talk with a higher pitch. When sad, they tend to speak slowly and with a lower pitch. Researchers have developed pattern-recognition techniques (i.e., machine learning) to infer human emotions from speech (El Ayadi et al., 2011; Han et al., 2014).

Finally, a robot can sense human affect from other modalities. For instance, human skin conductance changes in response to an individual's affective state. A prominent example of the use of skin conductance as a measure is the polygraph or lie detector. However, skin-conductance sensors have been tried in HRI, with only limited success (Bethel et al., 2007).

9.3.3 Expressing emotions with robots

Typically, people design robots that convey emotions through facial expressions. The most common approach here is to mimic the way in which people display emotions. This is a good example of how the study of human behaviors can be used for designing robot behaviors. The facial expression of emotions has been well documented (Hjortsjo, 1969). Ekman's Facial Action Coding System (FACS), in which human facial muscles are grouped as action units (AUs), describes emotions as combinations of action units (Ekman and Friesen, 1978). For instance, when a person displays a happy face (i.e., smiling), the muscles involved are the *orbicularis oculi* and *pars orbitalis*, which raise the cheek (AU6), and the *zygomaticus major*, which raises the corners of the mouth (AU12).

Using a simplified equivalent of human facial muscles, researchers have developed robots that are capable of conveying emotions through facial expressions. For instance, a robotic face with soft rubber skin and 19 pneumatic actuators was developed by Hashimoto et al. (2013). This robot uses AUs to express facial emotions. For example, it activates actuators corresponding to AU6 and A12 to express happiness. There are many other robots designed to express emotion that rely on a simplified interpretation of human facial cues, including Kismet (Breazeal and Scassellati, 1999), Eddie (Sosnowski et al., 2006), iCat (van Breemen et al., 2005), and eMuu (Bartneck, 2002), among others (see Figure 9.3).

Robots can also express emotion through various humanlike modalities, such as body movements and prosody. But even non-anthropomorphic robots can express affect, by means of adjusting their navigational trajectories. For instance, research on a cleaning robot (Saerbeck and Bartneck, 2010) and a flying robot (Sharma et al., 2013) showed that they could display affect by adapting particular motion patterns. Some other ways in which non-anthropomorphic robots can express affect include speed of motion, body posture, sound, color, and orientation (see Figure 9.4) to the person they are interacting with (Bethel and Murphy, 2008).



Figure 9.3
Emotions expressed through mechanical facial expressions. Left: eMuu (2001). Middle: iCat (2005–2012). Right: Flobi (2010). (Source: Left and middle, Christoph Bartneck; right, University of Bielefeld)



Figure 9.4
Non-anthropomorphic robots can express emotion through their behavior or through the addition of expressive features, such as lights. Anki, the producer of Cozmo (2016–2019), describes its robot as “[having] his own lively personality, driven by powerful A.I., and brought to life with complex facial expressions, a host of emotions and his own emotive language and soundtrack.” (Source: Anki)

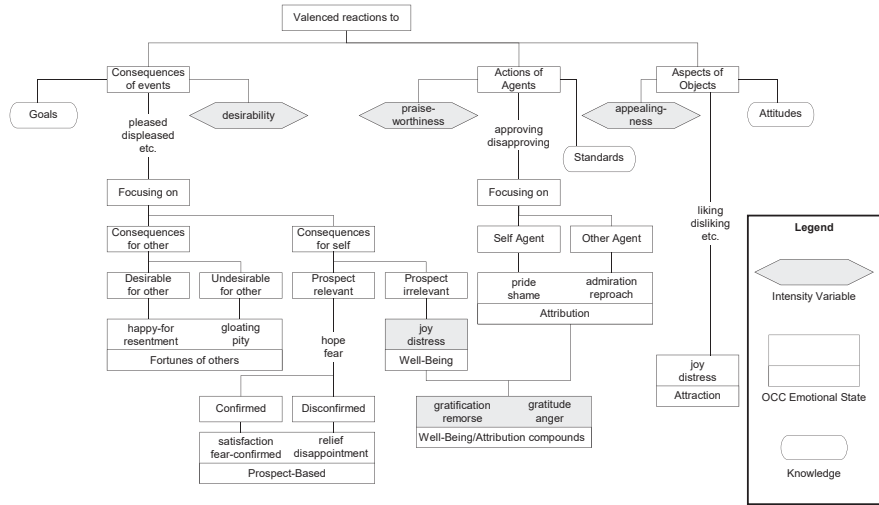
9.3.4 Emotion models

Psychologists have attempted to capture human emotions in formal models (Plutchik and Conte, 1997; Scherer, 1984). The benefit of this approach is that it views emotions as a numerical representation, which in turn lends itself well to representing emotion in computers and robots. These models also put different emotional categories in relation to each other, for example, by defining happiness as the polar opposite of sadness or by defining a distance function between emotions.

Emotion models are not only used to capture the emotional state of the user but can also be used to represent the emotional state of the robot itself and subsequently drive the behavior of the robot. For example, a robot with an almost empty battery can act tired and announce it needs a rest. Once it has reached the charger, it needs to update its internal emotional state to happy. Expressing this emotional state allows the user to have access to the robot’s internal state and will enrich the interaction.

A classic emotion model that has been used in some robots is the OCC model, named after its authors’ initials (Ortony et al., 1988). This model specifies 22 emotion categories based on valenced reactions to situations, such as events and acts of agents (including oneself), or as reactions to attractive or unattractive objects (see Figure 9.5). It also offers a structure for the variables,

Figure 9.5 The OCC model of emotions.



such as the likelihood of an event or the familiarity of an object, which determines the intensity of the emotion types. It contains a sufficient level of complexity and detail to cover most situations an emotional robot might have to deal with.

Needless to say, many robots do not possess the ability to express all 22 emotions. Even if they could, implementing 22 different emotions can be challenging; hence, many robot designers prefer to reduce the number of categories. Often, a decision is made to implement only Ekman’s six basic facial emotional expressions. These are reliably recognized, even across cultures Ekman (1992). However, a robot that only expresses six emotions makes for a quite limited interaction experience.

Perhaps more popular than the OCC model are the models that represent emotion as a point in a multidimensional space. Russell’s two-dimensional (2D) space of arousal and valence (see Figure 9.1) captures a wide range of emotions on a 2D plane and is one of the simplest emotion models that still has sufficient expressive power for HRI (Russell, 1980). The original 2D circumplex model, however, places “angry” and “afraid” side by side, whereas most people would argue that these are vastly different emotions. Later versions thus added a third axis, leading to the framework by Mehrabian (1980; see also Mehrabian and Russell 1974). This framework captures emotions in a three-dimensional (3D) continuous space, with the dimensions consisting of pleasure (P), arousal (A), and dominance (D) (see Figure 9.6). The PAD space model has been used on many social robots, including Kismet, to model the user’s and the robot’s emotional state (Breazeal, 2003).

9.4 Challenges in affective HRI

Despite considerable efforts in the perception, representation, and expression of emotion in virtual agents and robots, there are still a number of open challenges.

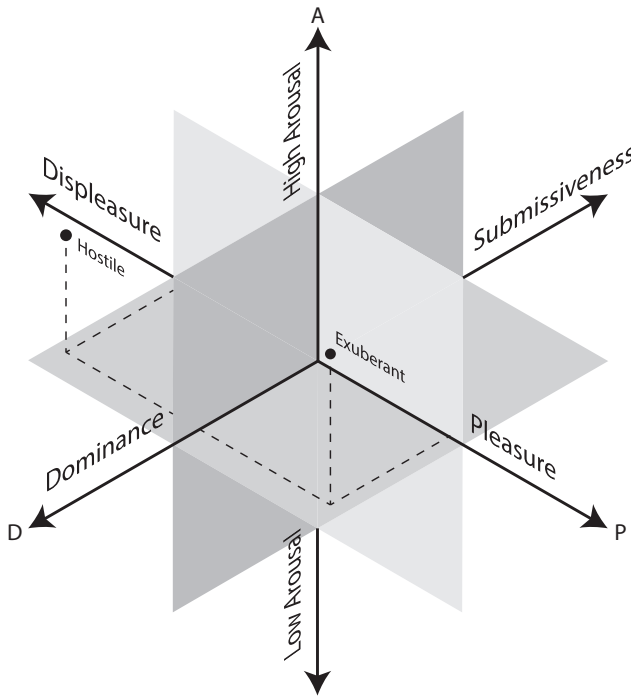


Figure 9.6 The PAD emotion model. An emotion is represented as a point in a 3D space, with axes representing pleasure (P), dominance (D), and arousal (A).

It is virtually impossible to correctly read emotions from facial information alone (see [Figure 9.7](#)). Given that people struggle to correctly read emotions from still facial images, robots will certainly have trouble with this as well. The addition of more information—such as the context of the interaction, animated rather than still expressions of emotion, and body language—allows us to increase the recognition rate, both by people and by algorithms.

Another problem in emotion recognition by computers is that almost all algorithms are trained on emotions that have been acted out by actors. As such, these emotions are exaggerated and bear little resemblance to the emotions we experience and express in daily life. This also means that most emotion-recognition software is only able to correctly recognize emotions that are displayed with a certain exaggerated intensity. Because of this, their use in real-world applications is still limited (Pantic et al., 2007), and the recognition accuracy of subtle emotional expressions drops dramatically (Bartneck and Reichenbach, 2005). Another problem is that most emotion-recognition software returns probabilities for only the six basic emotional expressions proposed by Ekman, or a point in a 2D or 3D emotion space. This is perhaps a rather limited view of emotion and misses many of the emotions we experience in real life, such as pride, embarrassment, guilt, or annoyance.

Another aspect of emotion recognition that poses difficulty for robots is recognizing emotions across a wide variety of people. Although we may all be expressing a number of universal emotions, we do not all do it with the same

Figure 9.7 Can you tell if the tennis player just scored or lost a point? A study showed that people struggled to correctly read strong emotions from the static faces alone, but they could, however, when only seeing the body posture (Aviezer et al., 2012). (Source: Steven Pisano)



intensity, in the same type of context, or with the same meaning. Interpreting the emotional status of a person, therefore, requires a sensitivity to his or her individual affective quirks. Humans become adept at this through long years of interacting with each other but also through long-term experience with individuals. That is why you might be able to tell that your partner is laughing out of annoyance rather than happiness, whereas new acquaintances may not be able to do so. Robots still decode emotions largely based on momentary snapshots of a person's countenance, and they do not develop more long-term models of affect, emotion, and mood for their interaction partners.

Finally, a robot's emotional responsiveness can fool potential end users into thinking the robot would actually experience genuine emotions. A robot merely expressing a certain emotion does not replace the actual, visceral experience of an emotional state. The robot merely displays emotional states in response to a computational model. Affective cognition, in which a full socioemotional repertoire is expressed and recognized for different users and contexts, still remains elusive.

9.5 Conclusion

Emotions are an important aspect of social interaction. In addition to intrapersonal functions such as evaluation of the situation and a motivation for action, they also serve an important interpersonal function because they inform others around us about our current mental state and (by extension) what kind of behavior they can expect from us. As such, in order to get a smooth interaction between a human and a robot, the robot will need to be able to both recognize the emotions displayed by the human and generate emotions for itself to help inform the human user on its internal state.

Questions for you to think about:

- Come up with a list of 10 emotions, and then try to display them nonverbally to a friend. Can your friend guess which emotion you are showing?
- Let's role-play: To understand how emotions are involved in our daily interaction, imagine being incapable of both experiencing and processing any information involving emotion. Then, set out to have a chat with a friend (consider telling the friend beforehand about your experiment). Try not to respond to whatever emotion your talking partner displays, and try not to show any emotional feedback. What happens?
- Are there tasks for which a robot should or shouldn't have emotion? Is it a good idea to implement emotion in a self-driving car, for example? If not, what are the potential problems?

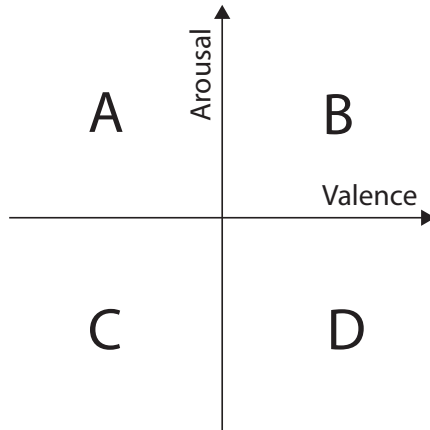


Figure 9.8 Emotion quadrants.

9.6 Exercises

The answers to these questions are available in the Appendix.

*** Exercise 9.1 Emotion quadrants** Associate the emotions with the correct quadrant as shown in [Figure 9.8](#).

1. Afraid: _____
2. Angry: _____
3. Astonished: _____
4. Bored: _____
5. Calm: _____
6. Content: _____
7. Delighted: _____
8. Depressed: _____
9. Frustrated: _____
10. Happy: _____
11. Relaxed: _____
12. Tired: _____

*** Exercise 9.2 Ekman and Friesen emotions** Ekman and Friesen proposed a set of six emotions. What was their purpose? Select one option from the following list:

1. To define a set of basic emotions
2. To define a set of negative emotions
3. To describe a list of facial expressions that are recognized across cultures
4. To describe the smallest shared set of emotions we all experience

**** Exercise 9.3 OCC model** The OCC model of emotions distinguished valenced reactions to what? Select one or more options from the following list:

1. Consequences of events
2. The robot's own emotional state
3. The emotions of the human user

4. Aspects of objects
5. Actions of agents
6. Aspects of agents

*** **Exercise 9.4 Robots with soul** Watch this video, and then answer the question that follows.

Guy Hoffman, “Robots with Soul” www.ted.com/talks/guy_hoffman_robots_with_soul

1. Hoffman uses principles from animation to improve interaction between humans and robots. After watching Hoffman’s TED talk, describe at least two potential benefits and two potential limitations of using animation principles in HRI. Do not just name the benefits and limitations, but explain why you see them as such in terms of the kinds of effects they can have on the success and quality of the HRI.

Future reading:

- Bartneck, Christoph, and Lyons, Michael J. Facial expression analysis, modeling and synthesis: Overcoming the limitations of artificial intelligence with the art of the soluble. In Vallverdu, Jordi, and Casacuberta, David, editors, *Handbook of Research on Synthetic Emotions and Sociable Robotics: New Applications in Affective Computing and Artificial Intelligence*, pages 33–53. IGI Global, Hershey, PA, 2009. URL <http://bartneck.de/publications/2009/facialExpressionAnalysisModelingSynthesisAI/bartneckLyonsEmotionBook2009.pdf>
- Breazeal, Cynthia. Social interactions in HRI: The robot view. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 34(2):181–186, 2004b. doi: 10.1109/TSMCC.2004.826268. URL <https://doi.org/10.1109/TSMCC.2004.826268>
- Calvo, Rafael A., D’Mello, Sidney, Gratch, Jonathan, and Kappas, Arvid. *The Oxford Handbook of Affective Computing*. Oxford University Press, New York, 2015. ISBN 978-0199942237. URL <http://worldcat.org/oclc/1008985555>
- Picard, Rosalind W. *Affective Computing*. MIT Press, Cambridge, MA, Cambridge, MA, 1997. ISBN 978-0262661157. URL <https://mitpress.mit.edu/books/affective-computing>
- Trapp, Robert, Petta, Paolo, and Payr, Sabine. *Emotions in Humans and Artifacts*. MIT Press, Cambridge, MA, 2003. ISBN 978-0262201421. URL <https://mitpress.mit.edu/books/emotions-humans-and-artifacts>