

Article

Contextual Emotion Appraisal Based on a Sentential Cognitive System for Robots

Hyunsik Ahn ^{1,*} and Sung Park ²¹ Department of Robot System Engineering, Tongmyong University, Busan 48520, Korea² School of Design, Savannah College of Art and Design, Savannah, GA 31415, USA; spark@scad.edu

* Correspondence: hsahn@tu.ac.kr; Tel.: +82-10-2518-4777

Featured Application: Emotion Appraisal for Human–Robot Interaction.

Abstract: Emotion plays a powerful role in human interaction with robots. In order to express more human-friendly emotions, robots need the capability of contextual appraisal that expresses the emotional relevance of various targets in the spatiotemporal situation. In this paper, an emotional appraisal methodology is proposed to cope with such contexts. Specifically, the Ortony, Clore, and Collins model is abstracted and simplified to approximate an emotional appraisal model in the form of a sentence-based cognitive system. The contextual emotion appraisal is modeled by formulating the emotional relationships among multiple targets and the emotional transition with events and time passing. To verify the proposed robotic system's feasibility, simulations were conducted for scenarios where it interacts with humans manipulating liked or disliked objects on a table. This experiment demonstrated that the robot's emotion could change over time like humans by using a proposed formula for emotional valence, which is moderated by emotion appraisal of occurring events.

**Citation:** Ahn, H.; Park, S.Contextual Emotion Appraisal Based on a Sentential Cognitive System for Robots. *Appl. Sci.* **2021**, *11*, 2027.<https://doi.org/10.3390/app11052027>

Academic Editor: Keun-Chang Kwak

Received: 6 January 2021

Accepted: 21 February 2021

Published: 25 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: emotion appraisal system; contextual emotion appraisal; cognitive robotics; sentential cognitive system; HRI

1. Introduction

As the technology of human–robot interaction (HRI) matures, the idea of robots being companions of humans in their daily lives is becoming popular. Emotional exchange with robots facilitates a human's empathy, engagement, and collaboration. An affective robot involves understanding human emotion and expressing believable emotion to humans, where the latter need to define the robot's emotional states and their expressions. Studies have focused on how emotions should be expressed on a robot's face [1,2] and in its voice [3], and how emotions should be appraised based on recognized information or stimulation [4,5]. One of the most significant studies on robots simulating emotion involves MIT's social robot Kismet [3]. Its cognitive architecture was influenced by Russell's perceptual approach to emotion [6], which included critical emotional constructs, such as arousal (strong versus weak) and valence (positive versus negative) associated with the robot's internal drive (e.g., social needs, fatigue). Kismet can express affect through face, voice, and posture based on an emotional internal representation form by the affective appraisal of external stimuli, such as threatening motion or praising a human's speech.

Emotion is intertwined with context. Specifically, emotion is associated with an object or an individual at a given time. How contexts from past to present are induced to appraise a robot's emotional state is important. A significant amount of research has been conducted in cognitive psychology on this subject. Ortony et al. claimed that emotion is determined by the interpretation of the context that induces emotions as reactions regarding interest toward an event or a subject. They categorized the emotional appraisal process based on 22 contexts [7].

Attempts have been made to apply psychological emotional models to robots. A study appraised emotions by defining a list of objects and actions and analyzed their relationship [4]. Another study appraised emotions using an explanatory process for recognized information [8]. However, it is difficult to consider these studies as full contextual emotion appraisals because they appraise emotions based on objects or events at points in time, rather than in a context extending from past to present.

In this paper, an emotion appraisal system based on the cognitive context of robots is proposed. For computational modeling, a simplified Ortony, Clore, and Collins (OCC) model is adopted, and that is applied to a sentential cognitive system (SCS) to combine emotional valence with cognitive information. The state of human emotions decreases over time to neutral, and some emotions are transferred to different emotions. In this paper, the change in intensity of emotional valence and the transition of emotion are modeled and formulated. For contextual emotion appraisal, the targets of emotion include objects, agents, and events. The robot evaluates the emotion of an agent dealing with objects, and the behavior of the agent becomes an event, which is a target of emotion. An implementation of the proposed approach is conducted with scenarios where a robot interacts with humans manipulating objects on a table.

The contributions of this paper are as follows: (1) a contextual emotion appraisal model abstracting and simplifying the OCC model is implemented in a SCS of a cognitive robot. (2) The suggested emotion appraisal model provides the principal contextual characteristics by formulating the emotional valence related among the targets of objects, agents, and events and introducing emotional transition according to the events interrelated with the targets.

This paper is organized as follows. In Section 2, related work on emotion appraisal studies is described. Section 3 details the contextual emotion appraisal model. Section 4 depicts a SCS based contextual emotion appraisal system. Section 5 provides the implementation of the proposed approach to a service robot, and Section 6 depicts conclusions and future work.

2. Related Work

Studies were conducted to make robots understand and express emotions in various ranges. In particular, as robots start to share their daily lives with humans, research to imitate human emotions (i.e., for robots to express emotions as partners) has been conducted.

For emotional HRI, robots need sophisticated functions to recognize human emotions. They also require formulating emotions attributed to external stimuli or information to express their state of emotion. A human-friendly emotional exchange may only be possible when the robot assesses the emotion considering the temporal, spatial, and event context surrounding the interaction with humans.

Research to understand human emotions by robots involves interdisciplinary studies on multi-modalities to imitate human emotions [9]. Studies of human emotion recognition range from image signal processing to biological signals [10–12]. Among the studies of human emotion recognition, Ikeda et al. used brain waves or heartbeat signals to recognize human emotions [10]. Lincon et al. researched nonverbal communication, such as recognizing and displaying emotions to identify humans to express empathy [11]. Cid et al. studied a person's face to recognize emotions, in addition to voice emotion recognition to enable multi-modal recognition of emotions [12].

In the field of emotion appraisal of robots, there are multiple approaches according to the emotion model, the emotion data types, and experiences from the past to the present.

In regards to the emotion data type, previous studies used stimuli, energies, sentences, targets, or feelings. For stimuli input, Uriel et al. suggested a way to have the robot respond to emotional expressions by touch. Depending on the touch's duration and pressure, a Bayes' rule allows the robot to evaluate and respond to four different emotions [13]. Kim et al. [14] suggested a computational reactive emotion generation model responding to temporally changing responses to stimuli using motivation, habituation, and conditional

models. In the energy model for emotion data, Lee et al. [15] distributed pleasure and arousal in the two-dimensional plane to indicate how emotions are distributed with the concepts of energy, entropy, and constancy for an emotional generation.

Jitviriya et al. used a self-organizing map to express emotions depending on the object. The motivation was determined according to objects, such as color, shape, and distance [16]. However, this method does not represent the context in which objects are used. Samani et al. [17] designed a high level of emotional bonds between humans and robots by applying human affection to robots. The system's advanced artificial intelligence includes three modules: probabilistic love assembly (PLA) based on the psychology of love; artificial endocrine system (AES) based on the physiology of love; and affective state transition (AST) based on emotion.

Another approach is combining emotions with sentences. Park et al. [18] suggested that the robot uses a set of multi-model motions described as a combination of sentence types and emotions to express its behavior. Hakamata et al. suggested a model to create emotional expressions in interactive sentences [19]. This study utilized the emotional weighting vector of the word in conversation and tagged each sentence of the conversation with it as mental status. However, these approaches do not include the maturity and transition of emotions that make a target a medium.

There are some approaches adopting a kind of context using the experience of the past. Zhang et al. [20] proposed an emotional model for robots, using Hidden Markov Model (HMM) techniques to enable past experiences to influence current emotions. This method assessed feelings based on their own emotional history, user impact, and task. Kinoshita et al. [21] proposed an emotion generation model dividing the robot's emotions with the degree of favorability as the short-term impression of a user, and degree of intimacy as the long-term impression, and expressed the accumulated affinities. The proposed model showed that robots respond differently to each user after communicating with multiple users. Kirby et al. [22] presented a generative model of affect that accounts for emotions, moods, and attitudes, including interactions among them. The model attempts to mimic the behavior, particularly with regard to long-term affective human responses. Itoh et al. [23] divided the robot's affective state into mood and emotion, which connected with a conversation between a human and the robot using a dialogist.

Finally, a bulk of research related to the theory of the mind, and the ability to understand other's intentions are worth analyzing in the context of HRI. Adaptive Character of Thought-Rational/Embodied (ACT-R/E), an extension of the prominent cognitive theory of ACT-R, proposed spatial reasoning in a three-dimensional (3D) world to increase the ecological validity required for HRI [24]. The temporal module in ACT-R, which records the events' interval, inspired our appraisal model to represent the temporal context. While SOAR (State, Operator, And Result), the other prominent cognitive module, had been applied to cognitive robots [25], we chose to go by a declarative knowledge representation in the form of a sentential structure for transparency and effective computation, as opposed to SOAR's procedural representation. However, Social Emotional AI (SEAI) is the most recent and impactful development on the subject matter applicable to HRI [26]. The framework captures the intricate interplay between emotion and high-level reasoning (cognition), inspired by Damasio's theory of consciousness [27]. SEAI is capable of abstraction and reasoning by adopting a hybrid reactive architecture.

In summary, most previous studies have emotional evaluation as an emotion that responds spontaneously to different kinds of stimuli. The work of contextual appraisal, using past experiences, uses accumulated emotional values over time or tagging on each dialogue sentence. There is a lack of research to associate contextual events related to objects and agents with temporal and spatial contexts.

3. Contextual Emotion Appraisal Model

In order to let robots and humans communicate by showing emotions, a theory of emotions must be interpreted and modeled from the perspective of a robot's intelligence. In

general, the popular emotion model used in robot emotion research consists of information sensing, emotion appraisal, emotion generation with a personality model, and emotion expression processes. Cognitive information sensed from the external environment is analyzed to extract emotional meanings (i.e., emotion appraisal). Such appraisal is combined with a personality model, by which reactions (e.g., calm, drastic) are mediated by the type of personality. This emotion is expressed through the robot's verbal (linguistic form) or nonverbal (facial expression, gestures, postures, etc.) channels. For more human-friendly emotional interaction, appraisal of emotion needs to be determined in a spatial-temporal context. In this study, such appraisal of contextual emotion is made possible by evaluating emotions on the each event that a robot experienced with the objects and agents using a cognitive system.

The OCC model of emotion contains semantics regarding the appraisal of contextual emotion, as shown in Figure 1 [7]. In this paper, for contextual HRI, emotion appraisal includes not just responses to emotional stimulations, but considers contexts, including spatiotemporal events occurred by agents handling objects. The OCC model is proper as the emotion appraisal model because it contains contextual appraisal for the expression of emotions in which agents, events, and objects are related to each other. The OCC model classifies emotion into 22 types, regarding the aspects of the object, action of the agent, and consequences of the event. The outcome of the agent's action or event is elicited from contextual relationships between the past and present, not merely from what is sensed. In the proposed emotion appraisal model, the cognitive appraisal of emotion, the classification of targets of emotion, and the transition process of emotion, are defined based on a simplified OCC model. This paper focuses on emotion appraisal of robot to agents, objects, their mutual events, and the linguistic descriptive model for being applied to robot cognitive systems. Non-verbal factors of emotion, such as the prosody, gesture, posture, gaze, and facial expression will be studied in future research.

3.1. Emotion Appraisal Model by Simplifying OCC Model

To establish an emotion appraisal model applicable to robots from the OCC model, the following conditions are imposed:

- Primacy condition: although the semantics of emotion can be interpreted variously, the primary meaning of the emotion is used.
- Valence condition: rather than classifying all aspects of emotions, they are represented with the valences of primary emotional states.
- Self-centered condition: only emotions focused on the robot itself are modeled.
- Cognitive condition: only emotions that can be analyzed by the robot's sensory and behavioral information are used.

In this paper, the OCC model is reconstructed with these imposed conditions, applied to the robot's emotion appraisal. Because the OCC model used a dichotomy of positive and negative emotions, the 22 types of emotion are defined from 11 emotional pairs (see Figure 1). The simplified OCC model adopts 10 types of emotion with five emotional pairs (see Table 1 for the emotions), according to the primacy, valence, and self-centered conditions. For example, because admiration-reproach focuses on the other agent and not the robot, this pair is excluded. Gratification-remorse is integrated with joy-distress because the latter is the primary emotion. As an emotional state that might arise in the future, hope-fear is defined as an emotion that encompasses satisfaction-fear-confirmed and relief-disappointment because of the primacy condition, and the subtle differences would be expressed through valence. With this simplification, 10 types of emotion are classified that can be applied to robots.

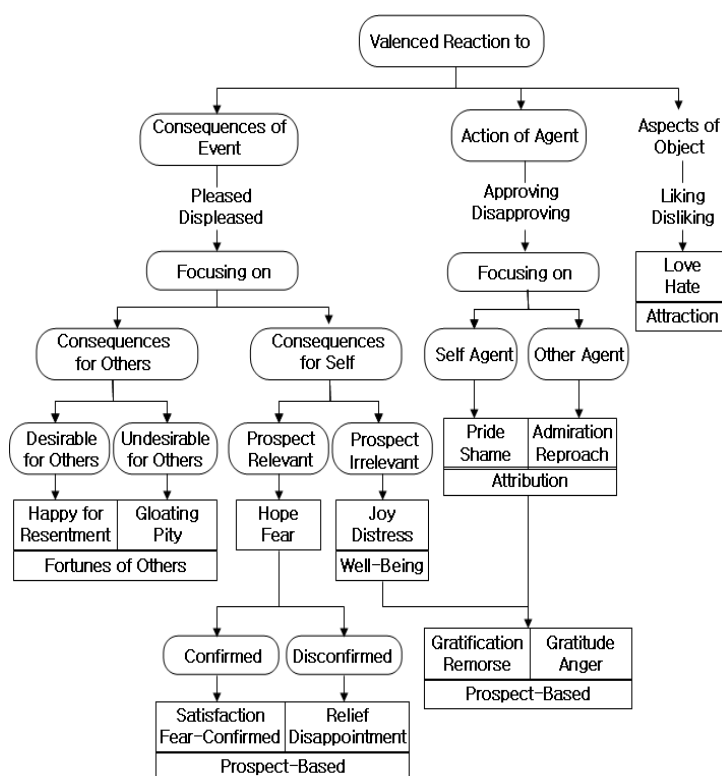


Figure 1. Ortony, Clore, and Collins (OCC) model for emotion appraisal [7].

Table 1. Types and contexts of emotion.

| # | Emotion | Appraisal Context | Dichotomy |
|----|-----------|--|-----------|
| 1 | Love | Aspects of target | Positive |
| 2 | Hate | | Negative |
| 3 | Pride | Action of self | Positive |
| 4 | Shame | | Negative |
| 5 | Gratitude | Action of agent | Positive |
| 6 | Anger | | Negative |
| 7 | Joy | Consequence of event/prospect irrelevant | Positive |
| 8 | Distress | | Negative |
| 9 | Hope | Consequence of event/prospect relevant | Positive |
| 10 | Fear | | Negative |

For a robot to appraise the 10 types of emotion on its own, the appraisal process must specifically involve the events that occur with targets in the given environment. In this paper, a linguistic approach is adopted to model the structure of the emotion appraisal to be related with the targets, objects, agents, and events. To define the emotional context among the targets, sentences that represent the emotions are composed using emotional verbs according to the classification, and the functions of the arguments are analyzed to define the targets of the emotions (see Table 2). Then, wordings conveying emotion are substituted with more general and primary verbs. For example, Thank with Gratitude, Rejoice with Joy, and Boast with Pride. To analyze the sentences composed with emotional verbs (Table 2) in phrase units, syntactic parsing was performed as follows, using Penn Treebank tag groups [28]. In this paper, syntactic parsing is used for analysis of emotion, and only rudimentary syntactic parsing is carried out for interpretation of sentences for expression

of emotion about events and production of sentences in the next chapter. Therefore, we adopt Penn Treebank, which provides relatively light and independent codes.

Table 2. The emotional classes, sentences constructed with emotional verbs, and the targets of emotion.

| # | Emotions | Sentences | Targets of Emotions |
|----|-----------|---|---------------------------------|
| 1 | Love | (S (NP I) (VP love (NP the object))), (S (NP I) (VP love (NP the agent) (PP by (NP the event)))) | (NP the object), (NP the agent) |
| 2 | Hate | (S (NP I) (VP hate (NP the object))), (S (NP I I) (VP hate (NP the agent) (PP by (NP the event)))) | |
| 3 | Pride | (S (NP I) (VP boast (NP the event))) | (NP the event) |
| 4 | Shame | (S (NP I) (VP am (VP shamed (PP for (NP the event)))))) | |
| 5 | Gratitude | (S (NP I) (VP thank (NP the agent) (PP for (NP the event)))) | (NP the agent) |
| 6 | Anger | (S (NP I) (VP anger to (NP the agent) (PP by (NP the event)))) | |
| 7 | Joy | (S (NP I) (VP rejoice (PP at (NP the event)))) | (NP the event) |
| 8 | Distress | (S (NP I) (VP am (VP distressed (PP by (NP the event)))))) | |
| 9 | Hope | (S (NP I) (VP hope (PP for (NP the event)))) | (NP the event) |
| 10 | Fear | (S (NP I) (VP fear (NP the event))) | |

S: sentence, NP: Noun phrase, VP: Verb phrase, PP: Prepositional phrase.

Syntactic parsing divides a sentence into multiple phrases. The Penn Treebank parser encompasses verb, noun, and prepositional phrases. Because of research in cognitive linguistics, it is generally accepted that the verb determines the characteristics of the sentence [29]. Therefore, the primary meaning of a sentence is determined by a verb's semantic structure, and a syntactically parsed sentence can be analyzed as an argument structure that defines the characteristics of the sentence. Once the argument structure is determined in association with the emotional verb, the target of emotion contained in the sentence can be defined. In a sentence describing an event, the robot is the subject of the emotion and is depicted as "I." The direct target of emotion is defined in the form of a noun phrase.

Table 3 shows the modeled relationship between the emotions and targets; where circles mean direct relationship, triangles as indirective relationship, and vacant blocks as irrelevance. An object is a passive target that cannot change the environment on its own, and an agent is an active target that can change the environment, such as a human or a robot. An event is defined as something that has occurred involving a robot, an agent, or an object. The target of love–hate (LH) can be an object or an agent. Because pride–shame (PS) is an emotion toward oneself, it is defined as an emotion about an event related to oneself; therefore, other agents cannot be the target of the emotion. The target of gratitude–anger (GA) can be an agent or an event. Present and future events can be targets of joy–distress (JD) and hope–fear (HF), respectively.

Table 3. The contextual emotion appraisal model with targets and source of emotions where circles mean direct relationship, triangles as indirective relationship, and vacant blocks as irrelevance.

| # | Emotions | Targets | | | Sources of Appraisal | | Emotional Valence |
|---|----------------------|---------|-------|-------|----------------------|------------|-------------------|
| | | Object | Agent | Event | A priori | Contextual | |
| 1 | Love–Hate (LH) | ○ | ○ | △ | ○ | ○ | −1.0–1.0 |
| 2 | Pride–Shame (PS) | | | ○ | | ○ | −1.0–1.0 |
| 3 | Gratitude–Anger (GA) | | ○ | △ | | ○ | −1.0–1.0 |
| 4 | Joy–Distress (JD) | | | ○ | | ○ | −1.0–1.0 |
| 5 | Hope–Fear (HF) | | | ○ | | ○ | −1.0–1.0 |

The emotion appraisal has two kinds of sources, *a priori* and contextual sources, as shown in Table 3. In the case of LH targeting objects, a robot needs to evaluate predefined emotional liked or disliked valence of the objects, because the robot does not have the feeling of satisfaction and preference to the objects. For example, because a robot cannot taste an apple, the emotional valence of the robot to an apple needs to have *a priori* definition. On the other hand, the other cases in the 10 emotions targeting agents and events need to evaluate the emotional valences contextually. If a robot feels love–hate toward an agent, the emotion does not abruptly arise, but it is developed over time from the experience with the behavioral events of the target.

In this approach, all of the emotional appraisals are computed with emotional valences with the span of -1.0 to $+1.0$. The negative emotion of emotional dichotomy can be extended to -1.0 according to the intensity of the emotion, and the positive to $+1.0$. At a certain time, the robot's emotional state is the combinational summation of emotions towards targets, agents, and events. To a certain agent at a point, the robot has the emotional state with the form of emotional valences of LH and GA, as shown in the vertical block of the agent of targets in Table 3.

When an emotion is aroused with an event, the state of the emotion sustains for a while and decreases gradually with time passing.

In this paper, an emotional intensity over time, $E(t)$, is suggested using the Gaussian function and sustain and decrease parameters. The Gaussian function with 0 mean is like (1).

$$G_{\sigma}(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp - \left(\frac{t}{\sigma} \right)^2 \quad (1)$$

The emotional intensity $E(t)$ is defined as (2).

$$E(t) = (1 - s)G_{\sigma}(t)/G_{\sigma}(0) + s - ct \quad (2)$$

where σ , the standard deviation in the Gaussian function, is used as the intensity duration, s is the sustainability parameter that retains a little after the main intensity duration, and d is diminution parameter over time span of emotional valence. Figure 2 shows a typical distribution graph of emotion intensity with the time axis with a set of parameters.

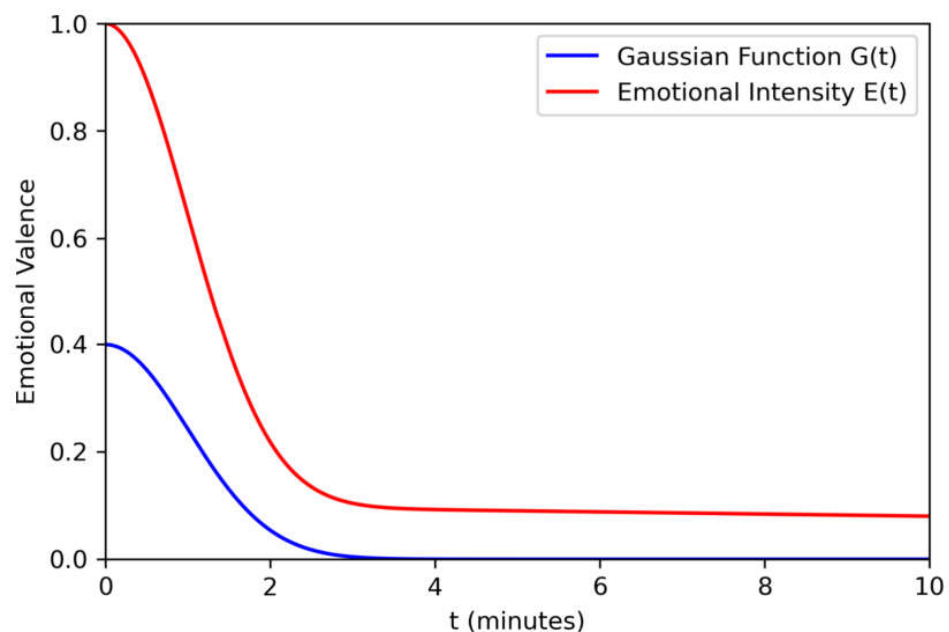


Figure 2. A typical distribution graph of emotional valence with the parameters (σ : 1, s : 0.1, d : 0.002).

In Table 3 are the valences of the emotion pairs distributed within -1.0 to $+1.0$ according to the positive and negative of the emotional states. It means if a positive emotion happens, the negative emotions can cancel each other. Therefore, the valence of an emotion pair is defined at a certain time; it is like (3).

$$V(t) = a E_p(t) + b E_n(t) \quad (3)$$

where $E_p(t)$ is the positive emotion and $E_n(t)$ is negative emotion of the pair, and a and b are an emotional combination rate in the span of 0 to 1, when the first emotion is aroused due to an event.

3.2. Emotional Transition Model

In the proposed emotion appraisal model, the emotional state of a robot is expressed with a combinational set of multiple emotions to targets. In the case of object, the emotional state is described with the valence of LH. However, the emotional state to an agent can be described with a combination of LH and GA, and to an event with PS, JD, and HF. For example, love and gratitude can be felt at the same time, which can be understood as an emotional state being caused by the diversity in the interpretation of the context. Given this, an emotion transition model is proposed to model the emotions' multiplicity and transitional characteristics.

Figure 3 shows the emotion transition model. This model is developed with the types of emotion used in this research and is largely consistent with the mental state transition model proposed by Ren [30]. This model is supported by empirical data from human subjects. Arrow ① indicates that when an agent causing JD events to a robot is recognized, the robot may display GA toward that agent. Arrow ② indicates that LH toward an agent may rise from the GA toward that agent. JD or HF is also produced when a target of LH of objects or agents is identified (③, ⑤) or when such an encounter is expected (④, ⑥). Arrow ⑦ indicates that HF is generated when JD is expected. Arrows ⑧ and ⑨ indicate that JD and HF are produced when PS occurs, or is expected to occur, respectively.

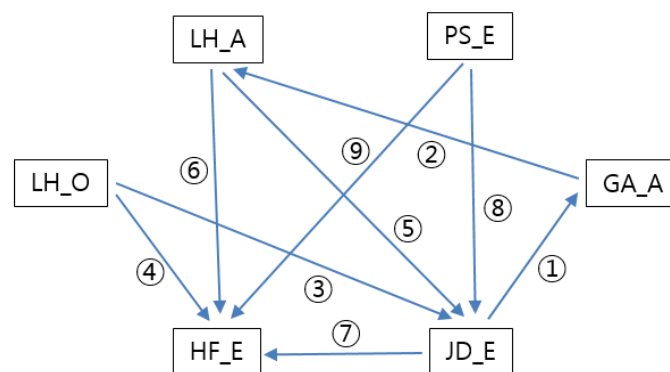


Figure 3. Transition of emotions and their mutual relationships among Love-Hate_Object (LH_O), Love-Hate_Agent (LH_A), Gratitude-Anger_Agent (GA_A), Pride-Shame_Event (PS_E), Joy-Distress_Event (JD_E), and Hope-Fear_Event (HF_E): ① An agent causing JD events, ② LH toward an agent may rise from the GA, ③ target of LH of objects is identified, ④ target of LH of objects is expected, ⑤ target of LH of agents is identified, ⑥ target of LH of agents is expected, and ⑦ HF is generated when JD is expected, ⑧ JD is produced when PS occurs, and ⑨ HF is produced when PS is expected to occur.

The emotion transition model explains how a single event may cause several emotional states based on the relationships among emotions. The current emotional valence is refreshed as a recipient with the previous valence and the summation of the transition of

other emotional valences with weight factors. In this paper, the transition of the emotion is defined as (4).

$$V_r(t) = V_r(t-1) + \sum_{k=1}^n h_k V_k(t) \text{ when } |V_k(t)| > T_k \quad (4)$$

where $V_r(t)$ is the current valence of emotion as a recipient, $V_k(t)$ is a sender of the valence, and h_k is weight factor of transition with the span of 0 to 1. Each $V_k(t)$ has a threshold value (T_k). T_k is used as a level of emotional intensity, which is needed to be over for the transition.

4. An Emotion Appraisal System on a Sentential Cognitive System

In this section, the SCS that bridges the cognitive information acquired from the sensing system to the emotion appraisal system is described.

4.1. An SCS for Emotion Appraisal

An emotion appraisal system based on the SCS is shown in Figure 4. The SCS consists of four functional blocks; a memory, an event manager, a perception, and a behavior [31]. Differently from previous SCSs, it has an emotion appraisal module (dotted line) in the event manager to appraise emotion in context. One of the merits of the proposed model is that the emotion appraisal is fulfilled based on the information of the SCS. The other advantage is that an emotion can be generated with context of cognitive information. In the SCS, an event is defined as newly recognized cognitive information different from the current state in the perception and behavior modules. The visual module, listening module, and sensing module receive information from the external environment. The utterance module and motion module exert action toward the environment. In the event manager, there is an event interpreter to represent an event to a sentence and to appraise the emotion of it, an event generator to produce such behaviors as utterance and motion, and schematic imagery to allocate objects virtually for spatial inference.

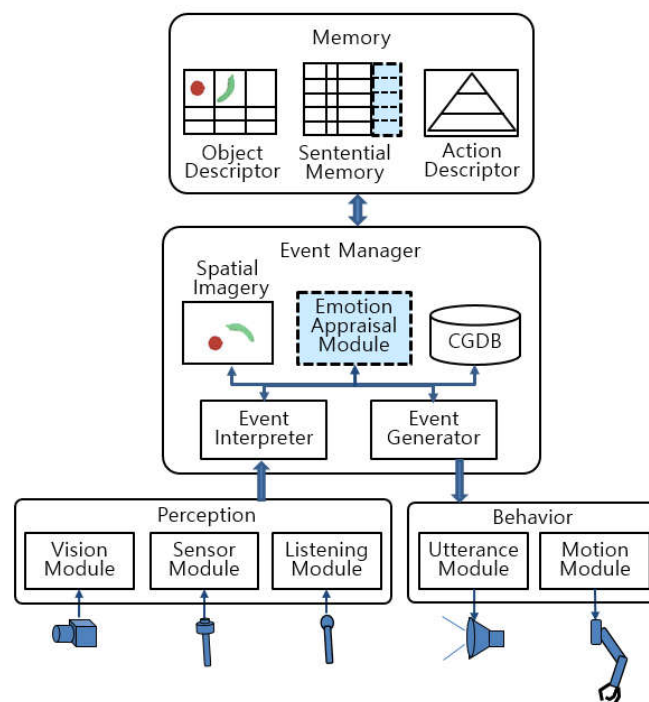


Figure 4. Schematic of an emotion appraisal system based on sentential cognitive system (SCS) that produces modular sentences with emotion valences from all events, and stores them in the sentential and auxiliary memories.

When the events happen to the robot, the cognitive information is interpreted as a sentence where the valence of emotion are tagged in the sentence. The memory of the system consists of the sentential memory, the object descriptor, and the action descriptor. The sentential memory stores interpreted sentences and their valences of emotion in chronological order (dotted line in the sentential memory). Auxiliary storage modules are used to represent the events effectively. The object descriptor stores the appearances of objects and agents with their cognitive properties and valences of emotion. The action descriptor stores behavioral information linking a verbal motion to a motion program function. The cognitive action module linked with sentential form of description was implemented and tested in the previous study [32]

4.2. Contextual Emotion Appraisal Based on the SCS

In this section, the process by which a robot recognizes an event and appraises emotions contextually by the emotion appraisal algorithm of the cognitive system is described. The proposed contextual emotion model consists of emotion pairs, according to the context shown in Table 3. Cases appraised with positive and negative emotions are defined as contrasting interpretations in the same context. The valence of emotion is quantified with real numbers between -1.0 and $+1.0$; a positive number denotes a positive emotion, 0 denotes a neutral emotion, and a negative number denotes a negative emotion. Emotions are appraised according to their classification as follows.

(1) Love–Hate (LH)

LH toward an object is developed from the likes/dislikes information defined a priori. Love develops when a liked sensation is detected by the sensor or when a liked object or agent is recognized. The emotion of love can also develop contextually toward an agent. For example, when an agent gives an object that the robot likes, gratitude develops, which leads to an emotional transition to love.

(2) Pride–Shame (PS)

PS is a contextually developed emotional state toward the robot itself. For example, a robot develops pride when it succeeds in completing an intended action, and shame when it fails.

(3) Gratitude–Anger (GA)

GA is produced by an agent executing a certain event. Gratitude is developed from a context in which an agent's action provides something that the robot likes. Anger occurs when an agent provides something that the robot dislikes.

(4) Joy–Distress (JD)

JD is defined as an emotion produced when the consequence of an event turns out to be favorable or unfavorable for the robot.

(5) Hope–Fear (HF)

HF is defined as an emotion produced when the future consequence of an event is expected to be favorable or unfavorable for the robot.

Table 4 shows multiple events and emotional valences affected by them. In the vision module, an object (O_i) appeared, and a sentence was created, as well as the emotional valence of the agent (A_i) who caused the event that affects the object. The objects have an a priori emotional valence of LH (l). The agents have the values of emotions LH (l) and GA (g), and the events have the values of emotions of PS (p), JD (j), and HF (h). When a vision event occurs, the vision module interprets the event and generates a sentence (S_1) with the cognitive information of the object. The interpreter produces a sentence with the syntactic parsing that tags the cognitive information to the components of the sentence. O_i is an object and x, y, z, θ are its position and pose. The emotional valence of each target has (l), (l, g) and (p, j, h) in between -1.0 and 1.0 . In the case of S_5 and S_6 , even if they are events, the event of question and answer does not make any change of emotional valence.

Table 4. Event and emotional valence. Each event occurs in a module and the objects, agents, and events have their own emotional valence according to (1)–(4). (Mod.: Modules, Arg_1: 1st Argument, Adj.: Adjunction, O_i : i th object, A_j : j th agent, E_k : k th event; S1: 1st sentence; l, g, p, j , and h : emotional valences of LH, GA, PS, JD, and HF respectively; V: vision module, L: listening module, U: utterance module).

| # | Mod. | Sentences of Events | | | | | | | Emotional Valences | | |
|----------------|------|---------------------|-------------|----------------|-------------|-------------------------|------|---------------------|--------------------|-------------------|-----------------------|
| | | Type | Verb | Arg_1 | Arg_2 | Space | Time | Adj. | O_i (LH) | A_j (LH, GA) | E_k (PS, JD, HF) |
| S ₁ | V | Declarative | appeared | (NP O_i) | - | (PP at (NP x, y, z)) | - | (PP by (NP A_j)) | (l) | (l, g) | (p, j, h) |
| S ₂ | V | Declarative | disappeared | (NP O_i) | - | - | - | (PP by (NP A_j)) | (l) | (l, g) | (p, j, h) |
| S ₃ | V | Declarative | appeared | (NP A_j) | - | (PP at (NP the front)) | - | - | - | (l, g) | (p, j, h) |
| S ₄ | V | Declarative | disappeared | (NP A_j) | - | - | - | - | - | (l, g) | (p, j, h) |
| S ₅ | U | What | is | (NP your name) | - | - | - | - | - | - | - |
| S ₆ | L | Declarative | am | (NP I) | (NP A_j) | - | - | - | - | - | - |

5. Experiments and Results

5.1. Implementation

The proposed emotion appraisal system was implemented using a sentence-based cognitive system, and experiments were conducted using scenarios to test the feasibility of the system. The experimental scenarios involve analyzing robot's emotional state when an agent gives or takes away an object that the robot likes or dislikes. Figure 5 shows the robot, Tongmyong University roBOt (TUBO), used in the experiment that is embedded with the emotion appraisal system.

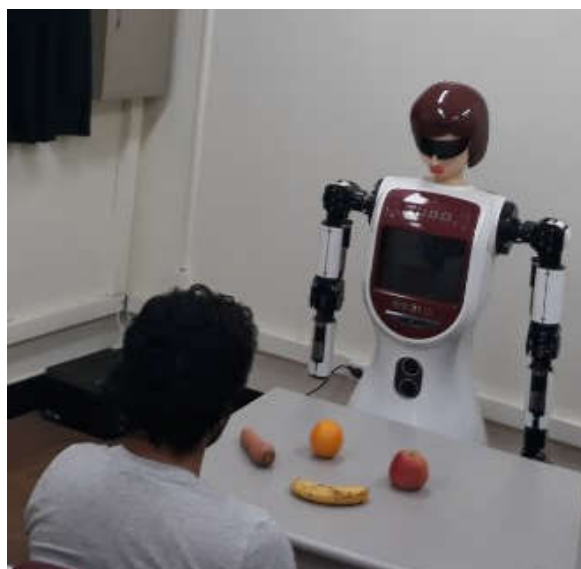


Figure 5. Emotion appraisal system implemented in a robot (TUBO) and an agent.

The emotion appraisal system was implemented with an IBM PC in the Visual C++ environment and implemented by programming the modules depicted in Figure 4. Link Grammar Parser's Penn Treebank was used as the syntactic parser [33], and automatic parsing was enabled so that events could be represented in sentences, and emotions related to such sentences could be appraised [31]. An RGB-D camera, Microsoft Kinect, attached at the head of the robot, was used for the vision module of SCS [34].

As explained in Table 3, emotions toward a target are classified into 10 types, consisting of five pairs of contrasting meanings. Objects are appraised with LH, agents with LH and GA, and events with PS, JD, and HF. A single emotion is assigned to one event, and transitions could take place among the five types of emotional state, as shown in Figure 3.

When visual events happened, the vision module of SCS-acquired RGB-D images captured by Kinect were processed to recognize the positions and poses of objects on a table. To recognize the objects in the RGB image of the input data, you only look once (YOLO), a convolution neural network (CNN), was adopted [35]. The vision module first found the bounding box (BBX) and label of the object. OpenGL libraries were used to get x, y, z coordinates of the cloud points using perspective transformation of depth data.

Figure 6 shows an RGB-D image and 3D recognition of objects in the vision module. Figure 6a shows object recognition results using Yolo from the RGB image, and Figure 6b shows the 3D view of table, which was transformed from the depth data with RGB texture mapping. In the 3D view, the center position of the object is identified on the planar coordinate system. The labels and positions of the objects are stored in the object descriptor.

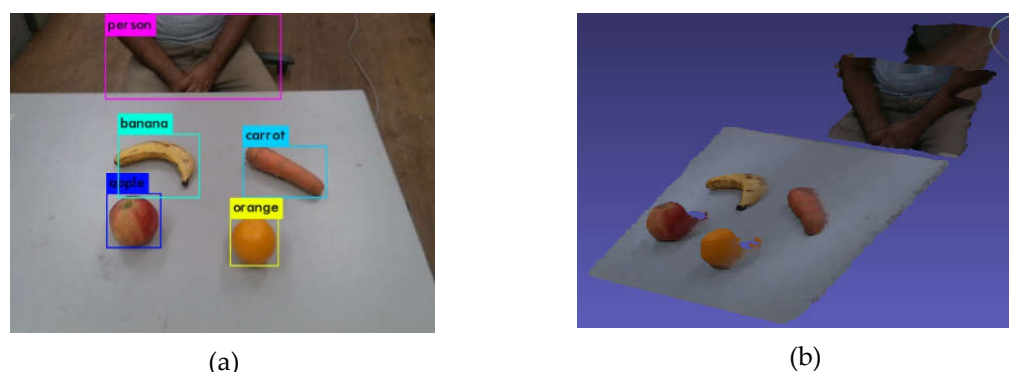


Figure 6. (a) Example of object recognition from RGB image, and (b) three-dimensional (3D) image transformed from depth data to the table surface plane.

5.2. Experimental Results

The scenario of the experiments involved testing the state of emotions when agents (human) brought or took fruits that the robot liked or disliked on a table. When the scene of the objects was changed in the presence of an agent, the robot considered that the agent had taken an action and regarded it as an event. Every event was represented as a sentence, and emotion appraisal was performed simultaneously. The proposed emotion appraisal system was tested with vision, listening, and utterance modules.

The emotional valences of a priori emotions to objects were predetermined, as shown in Table 5. It assumes that the robot has positive *a priori* LH emotion to apples (0.5) and oranges (0.3), but negative to bananas (−0.2) and carrots (−0.3). Just as humans have different emotional characteristics, the emotional valences of the fruits and the characteristics of the robot's emotions were predetermined by users. Table 6 shows the parameters of variation of emotional intensity $E(t)$ over time depicted in (2); LH of objects (LH_O), LH of agents (LH_A), GA of agents (GA_A), JD of events (JD_E), and HF of events (HF_E). These parameters are the result of randomly specifying the characteristics of the robot's emotions. Figure 7 shows the variation of emotional intensity over time with initial values of 1 by applying the parameters of the targets of Table 6. In the case of LH_O, the intensity of emotional valence is 1 and sustains all of the time, which means a priori valence of the emotion that a robot can have a fixed like or dislike to the fruits. The graph shows the parameters—that the emotion of LH_A has long duration of the intensity and sustains the emotion with higher intensity than LH_O and JD_E. The intensity of HF_E is defined with a short duration and long sustaining time with high values because the emotion of hope and fear could maintain by the time of expected events.

Table 5. The emotional valences of a priori LH emotion to objects in the experiments.

| Objects | <i>A Priori</i> LH_O Valences |
|---------|-------------------------------|
| apples | 0.5 |
| orange | 0.3 |
| bananas | −0.2 |
| carrot | −0.3 |

Table 6. The parameters of an emotional intensity over time $E(t)$ of (1) and (2); LH of objects (LH_O), LH of agents (LH_A), GA of agents (GA_A), and JD of events (JD_E) in the experiments.

| Targets | Parameters of $E(t)$ of (1) and (2) | | |
|---------|-------------------------------------|-----|-------|
| | σ | s | d |
| LH_O | 1 | 1 | 0 |
| LH_A | 2 | 0.5 | 0.002 |
| GA_A | 1.6 | 0.1 | 0.002 |
| JD_E | 1.9 | 0.2 | 0.002 |
| HF_E | 1 | 0.8 | 0 |

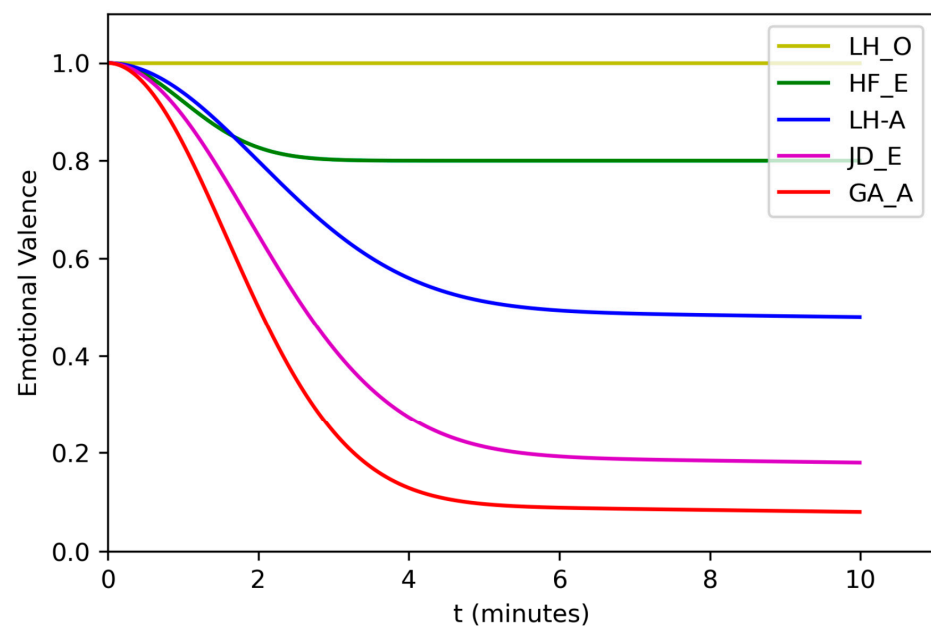


Figure 7. The variation of emotional intensity over time applying the parameters of the targets. In the case of LH_O, the intensity of emotional valence is 1 all of the time, which means a priori valence of the emotion to the fruits. The parameters for the emotion of LH_A has a long duration of the intensity and sustains the emotion with higher intensity than GA_A and JD_E. The intensity of HF_E has a short duration and long sustain with high values because the emotions of hope and fear could maintain high intensity for the future positive or negative events.

Figure 8 shows the results of object recognition of the series of events obtained using YOLO in the vision module, indicating the BBX and labels of objects. An agent was labeled as a person, each fruit was labeled, and the center positions of them were obtained from calculated coordinates by perspective transformation of the depth data.

Table 7. The sentences of the events and emotional valences stored in the sentential memory. (V: vision module, L: listening module, U: utterance module).

| # | Time (hh:mm:ss) | Module | Sentences of Events | Emotional Valences of Targets | | |
|-----------------|--------------------|--------|---|-------------------------------|----------------------------|------------------------|
| | | | | A ₁ (LH, GA) | A ₂ (LH, GA) | Events (PS, JD, HF) |
| S ₁ | 10:13:02 | V | (S (NP A new agent) (VP appeared (PP at (NP the front)))) | (0.0, 0.0) | (0.0, 0.0) | (0.0, 0.0, 0.0) |
| S ₂ | 10:13:05 | U | (S What (S (VP is (NP your name)))) | (0.0, 0.0) | (0.0, 0.0) | (0.0, 0.0, 0.0) |
| S ₃ | 10:13:09 | L | (S (NP I) (VP am (NP John))) | (0.0, 0.0) | (0.0, 0.0) | (0.0, 0.0, 0.0) |
| S ₄ | 10:13:25 | V | (S (NP An apple) (VP appeared (PP at (NP (173, 296)))))) | (0.0, 0.41) | (0.0, 0.0) | (0.0, 0.45, 0.0) |
| S ₅ | 10:13:28 | U | (S (VP Thank (NP you))) | (0.0, 0.41) | (0.0, 0.0) | (0.0, 0.45, 0.0) |
| S ₆ | 10:14:06 | V | (S (NP An orange) (VP appeared (PP at (NP (343, 162)))))) | (0.49, 0.62) | (0.0, 0.0) | (0.0, 1.0, 0.0) |
| S ₇ | 10:14:10 | U | (S (VP Thank (NP you))) | (0.49, 0.61) | (0.0, 0.0) | (0.0, 1.0, 0.0) |
| S ₈ | 10:14:14 | U | (S (NP I) (VP love (NP you))) | (0.49, 0.60) | (0.0, 0.0) | (0.0, 1.0, 0.0) |
| S ₉ | 10:14:52 | L | (S (VP see (NP you) (NP tomorrow))) | (0.47, 0.50) | (0.0, 0.0) | (0.0, 0.90, 0.38) |
| S ₁₀ | 10:14:57 | V | (S (NP John) (VP disappeared)) | (0.47, 0.48) | (0.0, 0.0) | (0.0, 0.97, 0.38) |
| S ₁₁ | 10:16:03 | V | (S (NP A new agent) (VP appeared (PP at (NP the front)))) | (0.40, 0.26) | (0.0, 0.0) | (0.0, 0.57, 0.34) |
| S ₁₂ | 10:16:06 | U | (S What (S (VP is (NP your name)))) | (0.40, 0.25) | (0.0, 0.0) | (0.0, 0.56, 0.34) |
| S ₁₃ | 10:16:11 | L | (S (NP I) (VP am (NP Tom))) | (0.39, 0.24) | (0.0, 0.0) | (0.0, 0.54, 0.34) |
| S ₁₄ | 10:17:07 | V | (S (NP A banana) (VP appeared (PP at (NP (635, 393)))))) | (0.32, 0.12) | (0.0, −0.16) | (0.0, 0.19, 0.31) |
| S ₁₅ | 10:17:23 | V | (S (NP An apple) (VP disappeared)) | (0.31, 0.10) | (0.0, −0.57) | (0.0, −0.29, 0.31) |
| S ₁₆ | 10:17:26 | U | (S (NP I) (VP am (ADJP angry))) | (0.31, 0.10) | (0.0, −0.56) | (0.0, −0.30, 0.31) |
| S ₁₇ | 10:17:32 | V | (S (NP An orange) (VP disappeared)) | (0.30, 0.09) | (−0.64, −0.80) | (0.0, −1.0, 0.31) |
| S ₁₈ | 10:17:35 | U | (S (NP I) (VP am (ADJP angry))) | (0.30, 0.09) | (−0.64, −0.80) | (0.0, −1.0, 0.31) |
| S ₁₉ | 10:17:38 | U | (S (NP I) (VP hate (NP you))) | (0.30, 0.09) | (−0.64, −0.80) | (0.0, −1.0, 0.31) |
| S ₂₀ | 10:19:02 | V | (S (NP A carrot) (VP appeared (PP at (NP (201, 278)))))) | (0.25, 0.06) | (−0.57, −0.67) | (0.0, −0.90, 0.30) |
| S ₂₁ | 10:19:05 | U | (S (NP I) (VP am (ADJP angry))) | (0.25, 0.06) | (−0.57, −0.65) | (0.0, −0.88, 0.30) |
| S ₂₂ | 10:20:38 | L | (S (VP see (NP you) (NP tomorrow))) | (0.24, 0.06) | (−0.49, −0.27) | (0.0, −0.42, −0.09) |
| S ₂₃ | 10:20:46 | L | (S (NP Tom) (VP disappeared)) | (0.24, 0.06) | (−0.48, −0.25) | (0.0, −0.38, −0.09) |

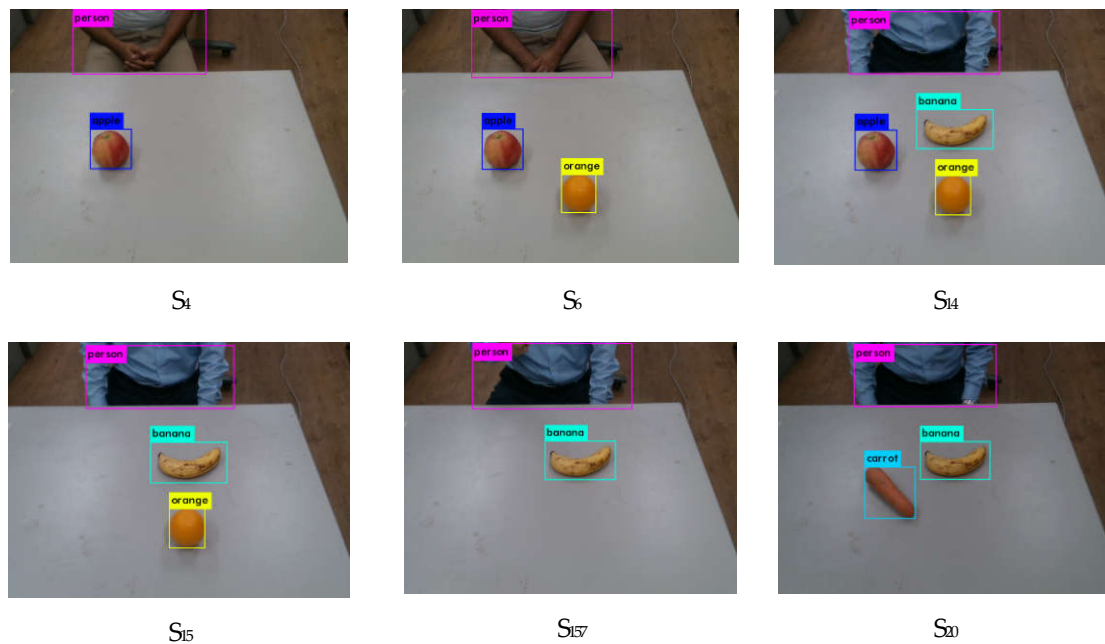


Figure 8. The scenes of the experiments and the object recognition results of each event (captions are the sentence numbers in Table 7). In the case of S_4 and S_6 , a user (A_1) as an agent did the positive action to the robot, but the other user (A_2) did the negative in the events of S_{14} to S_{20} . Table 7 shows the sentential memory of each event with the emotional valence, which was obtained from the emotional appraisal system. Events that occurred with agents and objects in front of the robot were stored in the sentential memory with chronological order. In the case of S_1 – S_3 and S_{11} – S_{13} , new agents, A_1 and A_2 , appeared, and the robot had a neutral emotional valence to them at first.

Figure 9 shows the variations of valences of emotions over time for the targets of two agents. When the agents brought in fruits that the robot liked or disliked, the event produced the valence of JD_E by the transition of LH_O (Figure 3,③). When the valence of JD_E was over the threshold value, there occurred a transition toward GA_A (Figure 3,①). The valence of GA_A activated LH_A by a transition (Figure 3,②) with the same way.

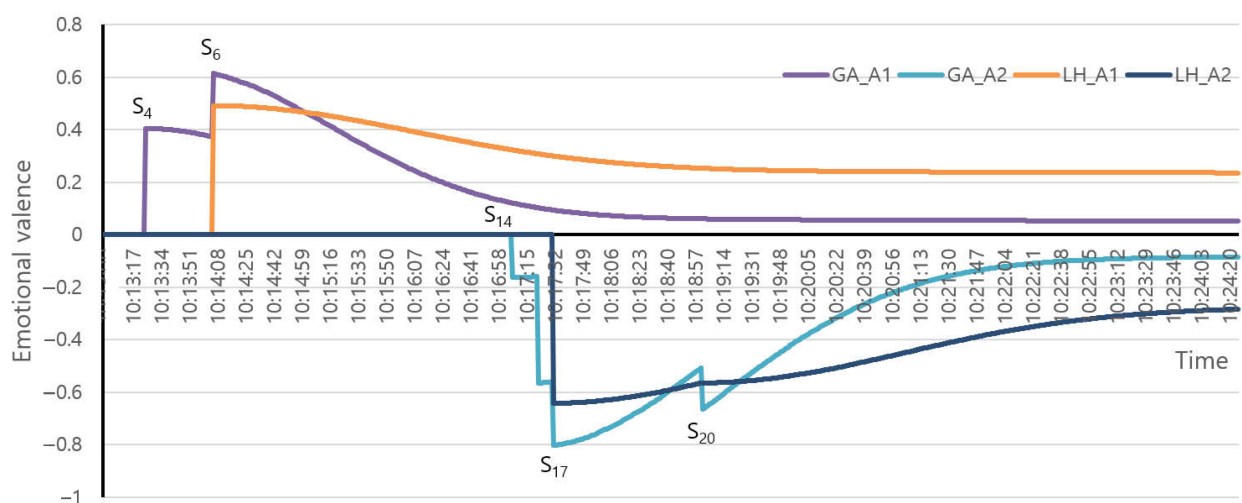


Figure 9. The variation of emotion appraisal toward the agents (A_1 and A_2).

When an apple (a priori LH_O valence: 0.5) appeared by agent A_1 , the event produced valence of JD_E by the transition of LH_O with the weight factor (h_k : 0.9). When the valence of JD_E was over a threshold value (T_k : 0.6), GA_ A_1 was activated by the transition of

JD_E. GA_A₁ was sustained for the time being and decreased with the graphical form of GA_A in Figure 7. Because GA_A₁ was over a threshold value (0.5) with a new event of bringing a fruit, the robot remarked “Thank you John” to the agent A₁, like (S₅, S₇). With the appearance of an orange (a priori LH valence: 0.3), the valence of GA_A₁ was over the threshold value (T_k : 0.6), which made an emotional transition to LH_A₁ with a weight factor (h_k : 0.8) of (4). Because LH_A₁ was over the threshold value (0.5) with a new event, the robot uttered “I love you John” to the agent A₁ (S₈).

When an agent A₂ brought fruits having negative a priori LH_O valences, or took away the liked fruits, the GA_A₂ had negative values and expressed negative utterances, such as “I am angry Tom” (S₁₆, S₁₈, S₂₁). The negative valence of GA_A₂ also made an emotional transition to LH_A₂ with the same T_k (0.5) (S₁₇) and the utterance of “I hate you Tom” (S₁₉). In Figure 9, we can see that the valence of GA_A₂ decreased to neutral in a short time whereas the valence of LH_A₂ sustained longer by adopting the parameters of $E(t)$ in Table 6, like the graphical characteristics of each emotion as shown in Figure 7. In the case of PS_E, it was developed by the event the robot did itself. For example, if the robot fails to pick up an apple when it is normal to pick it up, it will produce PS_E valence. In this experiment, this kind of event did not happen, and all of the valences of PS_E were 0.0.

Figure 10 shows the variations of emotion appraisal toward all events over time. There were multiple events, such as the appearance and disappearance of fruits, and the love and hate events to agents. The JD_E of the robot was appraised by the combinational transitions of LH_O (Figure 3,③) and LH_A (Figure 3,⑤) according to Equation (4). The transition from LH_O was computed by summation of the positive and negative events with the emotional combination rate (0.9) when the robot encountered the fruits, and the transition from LH_A was calculated with the integration of positive and negative LH_A. The JD_E was positive at first, but converted to negative at the time of 10:17:32 (S₁₅).

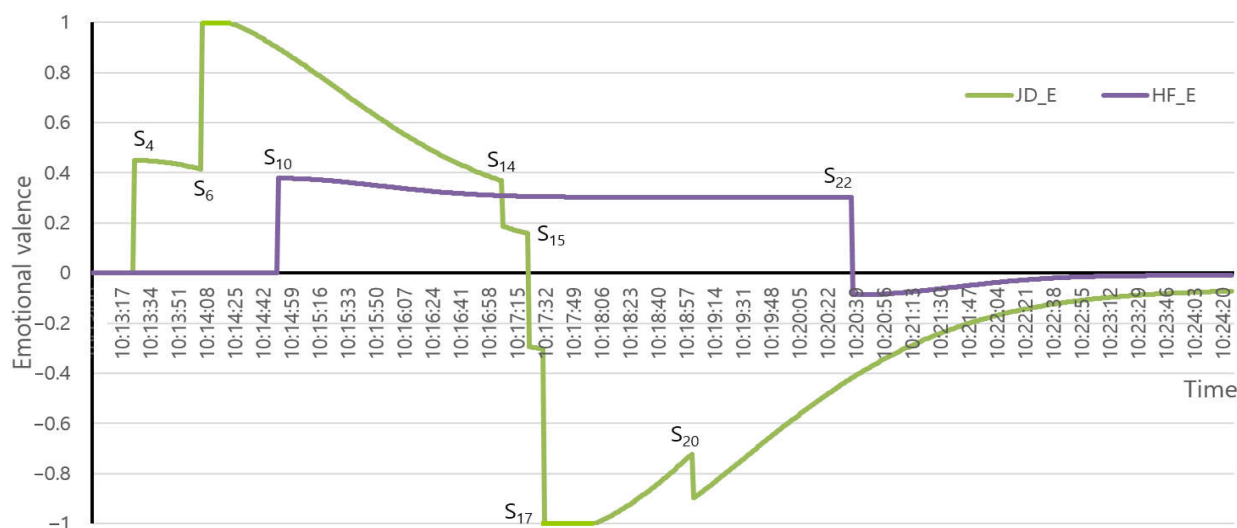


Figure 10. The variation of emotion appraisal toward the events.

When it was predicted that the target of H_A1 would appear in the future with “See you tomorrow (S₉),” the transition from LH_A (Figure 3,⑥) to HF_E happened at first with T_k (0.5) and h_k (0.8). When the robot heard the same sentence from A₂, the valence of HF_E switched to negative and gradually shifted to neutral value.

These experiments demonstrated contextual emotion appraisal in multiple points of view. First, the proposed contextual appraisal system showed that it was possible to evaluate emotions to the target of agent with events related to liked or disliked objects as well as compute a certain situation of a target, which means that the proposed model can support contextual emotion appraisal. Secondly, the emotional transition model can

not only make a certain emotion sustain for the time being and decrease to neutral state, but also develop it to the other emotions by combination of the current emotions. These transition characteristics of the model can be a principal aspect of contextual emotion appraisal. Due to these merits, the proposed emotion appraisal system built on the SCS of a robot could be of practical use in human-friendly emotional interactions between human and robots.

In the experiment, empirical emotions about the agent (LH, GA) and the events (JD, HF) from the transition of emotion were learned from the events, and changed like human emotion over time. Previous studies on emotion appraisal used the user's impact or user's favorability and intimacy [20,21]. For more human-like HRI, simple sensing or abstract expressions, such as intimacy or impact, have limitations in reality. The emotional model of a robot needs to have the capability of evaluating the meaning of other human actions related to objects to realize the robot's feelings toward humans. It is of significance that such an advanced appraisal model was presented in this paper.

6. Conclusions

In this paper, a contextual emotion appraisal system using a sentence-based cognitive system that can be applied to HRI was proposed. Ten types of emotion for context-based emotion appraisal were classified by redefining the psychological OCC model. Syntactic analysis of sentences, including emotional verbs, showed the characteristics of emotions and their targets, such as object, agents, and events. The contextual appraisal was modeled by formulating the emotional relationships among targets, including objects, agents, and events, and the emotional transition over time. The emotional valences of targets were evaluated with respect to each event, which caused emotional context, tagged on the parsed sentence expressing the event, and stored in the sentential memory of SCS.

The emotion appraisal system showed the variation of a robot's emotions when agents gave or took away objects that the robot liked or disliked, which means that the proposed model can support contextual emotion appraisal related to the target of objects, agents, and events. The proposed emotional transition model showed that it could cause other emotions from the current emotions with threshold values. Due to these merits, the proposed emotion appraisal system built on the SCS of a robot could be of practical use in human-friendly emotional interaction between humans and robots.

The experiment showed that contextual emotion appraisal could be achieved by evaluating the variations of a robot's emotions when agents give or take away an object that the robot likes or dislikes. For the proposed model to be applied in practical HRI, the cognitive capability of perception and behavior modules need to be improved, including 3D motion recognition and ontological classification of all the targets. There has been growing interest in deep-learning approaches, which could be adopted for the advancement of this field of research. Moreover, an integrated cognitive model that includes emotion appraisal from verbal and nonverbal cues could increase the validity of a human-friendly HRI.

Author Contributions: H.A.: Conceptualization, methodology, software, validation, formal analysis, original draft preparation, and project administration; S.P.: investigation, resources, data curation, and review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the Basic Science Research Program through the National Research Foundation of Korea (grant number 2012-0003116).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hashimoto, M.; Yokogawa, C.; Sadoyama, T. Development and control of a face robot imitating human muscular structures. In Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, Beijing, China, 9–15 October 2006; pp. 1855–1860.
2. Lee, H.-S.; Park, J.-W.; Chung, M.-J. A linear affective space model based on the facial expressions for mascot-type robots. In Proceedings of the 2006 SICE-ICASE International Joint Conference, Busan, Korea, 18–21 October 2006; pp. 5367–5372.
3. Breazeal, C. Emotive qualities in lip synchronized robot speech. *Adv. Robot.* **2003**, *17*, 97–113. [[CrossRef](#)]
4. Park, G.-Y.; Lee, S.-I.; Kwon, W.-Y.; Park, J.-B. Neurocognitive affective system for an emotive robot. In Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robot and Systems, Beijing, China, 9–15 October 2006; pp. 2595–2600.
5. Arkin, R.C.; Fujita, M.; Takagi, T.; Hasegawa, R. An ethological and emotional basis for human-robot interaction. *Robot. Auton. Syst.* **2003**, *43*, 191–201. [[CrossRef](#)]
6. Russell, J. Evidence for a three-factor theory of emotions. *J. Res. Personal.* **1977**, *11*, 273–294. [[CrossRef](#)]
7. Ortony, A.; Clore, G.L.; Collins, A. *The Cognitive Structure of Emotion*; Cambridge University Press: Cambridge, UK, 1988.
8. Oh, S.; Gratch, J.; Woo, W. Explanatory style for socially interactive agents. In Proceedings of the 2nd International Conference on Affective Computing and Intelligent Interaction, Lisbon, Portugal, 12–14 September 2007; pp. 534–545.
9. Marsella, S.; Gratch, J.; Petta, P. Computational models of emotion. In *Blueprint for Affective Computing: A Source Book*, 1st ed.; Scherer, K.R., Bänziger, T., Roesch, E.B., Eds.; Oxford University Press: Oxford, UK, 2010.
10. Ikeda, Y.; Horie, R.; Sugaya, M. Estimating emotion with biological information for robot interaction. *Procedia Comput. Sci.* **2017**, *112*, 1589–1600. [[CrossRef](#)]
11. Rincon, J.A.; Martin, A.; Costa, A.G.A.S.; Novais, P.; Julian, V.; Carrascosa, C. EmIR: An Emotional Intelligent Robot Assistant. In Proceedings of the AfCAI, Valencia, Spain, 19–20 April 2018.
12. Cid, F.; Manso, L.J.; Nunez, P. A novel multimodal emotion recognition approach for affective human robot interaction. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Hamburg, Germany, 28 September–2 October 2015; pp. 1–9.
13. Martinez-Hernandez, U.; Prescott, T.J. Expressive touch: Control of robot emotional expression by touch. In Proceedings of the 25th IEEE International Symposium on Robot and Human Interactive Communication, New York, NY, USA, 26–31 August 2016; pp. 974–979.
14. Kim, H.-R.; Koo, S.-Y.; Kwon, D.-S. Designing reactive emotion generation model for interactive robots. In Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, Taipei, Taiwan, 18–22 October 2010; pp. 18–22.
15. Lee, W.H.; Park, J.-W.; Kim, W.H.; Kim, J.C.; Chung, M.J. Robot's emotion generation model for transition and diversity using energy, entropy, and homeostasis concepts. In Proceedings of the 2010 IEEE International Conference on Robotics and Biomimetics, Tianjin, China, 14–18 December 2010; pp. 555–560.
16. Jitviriy, W.; Hayashi, E. Design of emotion generation model and action selection for robots using a selforganizing map. In Proceedings of the 11th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Nakhon Ratchasima, Thailand, 14–17 May 2014; pp. 1–6.
17. Samani, H.A.; Saadatian, E. A multidisciplinary artificial intelligence model of an affective robot. *Int. J. Adv. Robot. Syst.* **2012**, *9*, 1–11. [[CrossRef](#)]
18. Park, J.-C.; Song, H.; Koo, S.; Kim, Y.-M.; Kwon, D.-S. Robot's behavior expressions according to the sentence types and emotions with modification by personality. In Proceedings of the 2010 IEEE Workshop on Advanced Robotics and its Social Impacts, Seoul, Korea, 26–28 October 2010; pp. 105–110.
19. Hakamata, A.; Ren, F.; Tsuchiya, S. Human emotion model based on discourse sentence for expression generation of conversation agent. In Proceedings of the 2008 International Conference on Natural Language Processing and Knowledge Engineering, Beijing, China, 19–22 October 2008.
20. Zhang, X.; Alves, S.; Nejat, G.; Benhabib, B. A robot emotion model with history. In Proceedings of the 2017 IEEE International Symposium on Robotics and Intelligent Sensors, Ottawa, ON, Canada, 5–7 October 2017; pp. 230–235.
21. Kinoshita, S.; Takenouchi, H.; Tokumaru, M. An emotion-generation model for a robot that reacts after considering the dialogist. In Proceedings of the 2014 International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), Palawan, Philippines, 12–16 November 2014; pp. 1–6.
22. Kirby, R.; Forlizzi, J.; Simmons, R. Affective social robots. *Robot. Auton. Syst.* **2010**, *58*, 322–332. [[CrossRef](#)]
23. Itoh, C.; Kato, S.; Itoh, H. Mood-transition-based emotion generation model for the robot's personality. In Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics, San Antonio, TX, USA, 11–14 October 2009; pp. 2878–2883.
24. Trafton, J.; Hiatt, L.M.; Harrison, A.M.; Tamborello, F.P.; Khemlani, S.S.; Schultz, A.C. ACT-R/E: An Embodied Cognitive Architecture for Human-Robot Interaction. *J. Hum. Robot Interact.* **2013**, *2*, 30–55. [[CrossRef](#)]
25. Laird, J.; Kinkade, K.R.; Mohan, S.; Xu, J.Z. Cognitive Robotics using the SOAR Cognitive Architecture. In Proceedings of the 2012 AAAI Workshop, CogRob@AAAI 2012, Toronto, ON, Canada, 22–23 July 2012.
26. Cominelli, L.; Mazzei, D.; Rossi, D. SEAI: Social Emotional Artificial Intelligence Based on Damasio's Theory of Mind. *Frontiers in Robotics and AI* **2018**, *5*, 6. [[CrossRef](#)] [[PubMed](#)]
27. Damasio, A. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*; Spektrum Der Wiss: Heidelberg, Germany, 2000; p. 104.

28. Marcus, M.P.; Santorini, B.; Marcinkiewicz, M.A. Building a large annotated corpus of English: The Penn Treebank. *Comput. Linguist.* **1993**, *19*, 313–330.
29. Garnsey, S.M.; Pearlmutter, N.J.; Myers, E.; Lotocky, M.A. The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *J. Mem. Lang.* **1997**, *37*, 58–93. [[CrossRef](#)]
30. Ren, F. Affective Information processing and recognizing human emotion. *Electron. Notes Theor. Comput. Sci.* **2009**, *225*, 39–50. [[CrossRef](#)]
31. Ahn, H. A sentential cognitive system of robots for conversational human–robot interaction. *J. Intell. Fuzzy Syst.* **2018**, *35*, 6047–6059. [[CrossRef](#)]
32. Ahn, H.; Ko, H. Natural-language-based robot action control using a hierarchical behavior model. *IEIE Trans. Smart Process. Comput.* **2012**, *1*, 192–200.
33. Link Grammar. Available online: <http://www.abisource.com/projects/link-grammar> (accessed on 1 July 2018).
34. Kinect for Windows. Available online: <http://developer.microsoft.com/en-us/windows/kinect> (accessed on 1 July 2018).
35. Redmon, J.; Farhadi, A. YOLO9000: Better, Faster, Stronger, CVPR. 2017. Available online: <https://pjreddie.com/publications/> (accessed on 1 September 2019).