Knowledge Representation for Conceptual, Motivational, and Affective Processes in Natural Language Communication

Seng-Beng Ho[1], Zhaoxia Wang[2], Boon-Kiat Quek[1], Erik Cambria[3]

[1]Institute of High Performance Computing, Singapore
[2]Singapore Management University
[3]Nanyang Technological University, Singapore
hosb@ihpc.a-star.edu.sg, zxwang@smu.edu.sg,
quekbk@ihpc.a-star.edu.sg, cambria.ntu.edu.sg

Abstract. Natural language communication is an intricate and complex process. The speaker usually begins with an intention and motivation of what is to be communicated, and what outcomes are expected from the communication, while taking into consideration the listener's mental model to concoct an appropriate sentence. Likewise, the listener has to interpret the speaker's message, and respond accordingly, also with the speaker's mental model in mind. Doing this successfully entails the appropriate representation of the conceptual, motivational, and affective processes that underlie language generation and understanding. Whereas big-data approaches in language processing (such as chatbots and machine translation) have performed well, achieving natural language based communication in human-robot collaboration is non-trivial, and requires a deeper representation of the conceptual, motivational, and affective processes involved in conveying precise instructions to robots. This paper capitalizes on the UGALRS (Unified General Autonomous and Language Reasoning System) framework and the CD+ (Conceptual Dependency Plus) representational scheme to demonstrate how social communication through language can be supported by a knowledge representational scheme that handles conceptual, motivational, and affective processes in a deep and generalizable way. Through an illustrative set of concepts, motivations, and emotions, we show how these aspects are integrated into a general framework for knowledge representation and processing that could serve the purpose of natural language communication for an intelligent system.

Keywords. natural language communication, natural language understanding, knowledge representation, motivational processes, affective processes

1 Introduction

Current research in AI typically regards natural language understanding and natural language generation as separate processes. In many language processing systems, the focus is on surface level comprehension without a deep representation of meaning [1]–[3]. Despite the absence of deep meaning representation in these systems, conversational systems like chatbots and machine translation have achieved significant commercial success [4]. However, there are certain applications of natural language understanding that require a deeper or grounded level of representations. For instance, when using natural language instructions to guide a robot in performing certain actions, the robotic system needs to comprehend the instructions at a level where the true meaning is represented, understood, and translated into actions and behaviors. Additionally, human language communication frequently involves motivational and affective processes. For language generation, there must first be underlying motivations influenced by ongoing emotional states, before constructing and producing an utterance.

In order for language understanding to occur accurately, the listener's internal language understanding processes must include a model of the utterer's motivational and emotional states. This allows the appropriate comprehension of spoken sentences and enables the listener to respond correctly. Consider the case of sentiment analysis, which focuses on detecting emotions conveyed in sentences. While extensively researched [5]–[11], there has yet to be an integrated natural language communication framework that connects the processes of language generation and understanding in a systematic manner while incorporating a principled treatment of motivational and emotional processes, which are closely interwoven with language generation and understanding [1], [2]. This paper aims to address these issues by utilizing the UGALRS architectural framework proposed by Ho[12].

The remainder of this paper is organized as follows: Section II discusses motivations and background; Section III introduces the basic architectural and representational constructs; Section IV elaborates on the core of the representations for language communication; finally, Section V proposes concluding remarks.

2 Motivation and Background

Natural language communication is a two-way process of generation and understanding. An utterance from a person presupposes a listener in mind and is motivated by a communicative goal and intention to achieve that goal. Then, depending on the context and emotional state of the speaker and the intended effect on the listener, an appropriate sentence is formed to communicate the intended message to the listener [13]. To do this successfully, the speaker would need a good model or representation of the mental and emotional states of the listener.

While current big data approaches to language processing, such as GPT-3 [4] are successful in many applications, they do not explicitly model these internal processes. The consequence is that certain explanations and more complex responses are not possible. For example, suppose Person A asks Person B, "Why do you say that to him?" [II.1] And Person B replies, "I want him to feel hurt." [II.2] Person A might counter suggest, "Well, I would suggest a better way to do that, which is to…" [III.3]. Utterance II.2 offers an explanation of why Person B said something, and Person B, having a model of her own mental processes including the motivation behind the earlier sentence Person A is asking her about, will be able to provide Person A with an explanation of what "causes" her to say the hurtful thing earlier.

Person A, who presumably also has an internal model of herself, Person B, and perhaps also the person whom Person B wants to hurt, will then be able to infer the various mental causalities involved and propose a different way to manage those emotions that Person B is experiencing. Whether her suggestion will be a malicious one intended to go along with and please Person B or a benign one to placate the situation for the betterment of all involved will depend on the background of their conversation and other intentions of Person A to start with. Processes such as these (i.e., attributions of intention, causality, emotions) could presumably occur when humans communicate with robots to instruct them to carry out certain tasks, or when robots are communicating with each other in collaborating to perform certain tasks. Thus, a robotic or AI system could benefit from a fuller model of the internal processes of language generation, communication, and understanding, such that humans could in turn attribute mental states and processes to them.

Language communication is inherently a form of social interaction, and autonomous systems capable of interacting and collaborating with other agents through language understanding and generation, are effectively functioning as social agents. This suggests that it might be worthwhile examining existing work involving internal operating architectures for autonomous systems or social agent, as a source of inspiration for building such agents. Among various challenges, issues relating to grounding language representations would need to be addressed. In this regard, Ho [12] has introduced the Unified General Autonomous and Language Reasoning System (UGALRS)—an architecture that not only outlines the various essential operating components of autonomous systems, but also provides a framework for grounded meaning representations. This framework, known as CD+ (an enhanced version of the conceptual dependency (CD) theory [22]–[24]),operates within the UGALRS framework. For the purposes of this paper, we will adopt UGALRS and CD+ as our chosen representational framework.

As mentioned earlier, the role of motivational and affective processes in language communication is often overlooked in linguistic research [14]–[17] computational linguistics, and natural language processing [1], [2], [18]. However, these processes play significant roles in language generation and understanding. In this regard, the UGALRS plus CD+ framework offers the necessary representational constructs to address motivational and affective processes effectively. Given that motivational and affective processes are essential components in the functioning of a social agent involved in communication and collaboration, UGALRS could serve as an architecture suited for social agent as well.

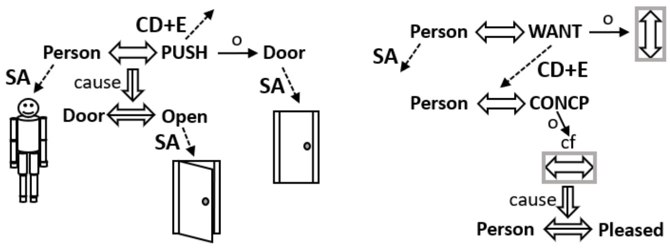
Quek et al. [19]–[21] have developed an architecture that integrates motivation and emotion into the functioning of an autonomous system. While primarily directed at typical robotic actions such as navigation and exploration, it aligns well with our work where these “actions” correspond to language generation within the communication process. Given the breadth of conceptual, motivational, and affective processes, for the purposes of this paper we will focus our discussion to a limited subset of each aspect, as our goal is to articulate and elucidate the intricate connections between these components within a general framework, which could be expanded upon in future research.

3 Basic Architectural and Representational Constructs

In this section we explain the construction of the basic architectural and representational constructs in terms of the UGALRS architecture and the associated CD+ representational scheme.

3.1 The CD+ Concept Representational Framework

Ho [12] developed a general representational framework that could encompass a large variety of, if not all, representations of concepts. A key idea posited in [12] is that many often-encountered concepts are functional in nature. This framework, known as CD+, is an extension of Schank's conceptual dependency (CD) representational framework [22]–[24]. The representations used in CD+ are cognitive, causal, and grounded.



(a) (b)

Fig. 1. Examples of CD+ representations. (a) "Person pushes the door open." (b) "Person wants something."

Figure 1(a) shows the representation of the sentence, or conceptualization, "*Person pushes the door open*." The horizontal double arrow links the subject, Person, to the object (o), Door. This is termed a "conceptualization". In CD+, the two main constructs being enhanced over CD are the Structure Anchor (SA) and the CD+ Elaboration (CD+E) [12]. SA is a detailed structural or analogical representation of the object the symbol refers to, which provides grounding for the symbol. In this case, the concept of a Person, could encompass representational details of its bodily structure, specified to a level of detail that includes every point, limbs, and joints, as shown.

A convenient way to implement this representational model could be using the kind of analogical representation used in computer graphics in the form of a high density point cloud consisting of points corresponding to every point of the object involved, or some vectorial representations that represent the loci of these points. The various parts on the object involved (say, in the case of Person, could include the various body parts such as the head, torso, and limbs) may be movable relative to one another, and would have to be captured in the model. CD+E is used to elaborate on the symbols that represent certain actions, such as in this case, PUSH. PUSH involves a number of sub-steps, such as "*first place palm flat on Door near center of Door*," "*then exert strength in the direction perpendicular to Door's surface*," etc.

Here, we use English sentences to describe these steps for the ease of explanation, but each of these steps is in turn representable in CD+ form. This could carry on as we traverse a hierarchy of details until an eventual, "ground level," is reached, in which the concepts used are "ground level concepts." In [12], a comprehensive set of ground level concepts is presented, serving as a common ground for many concepts. In the two English sentences above, every concept used in their representations has to be clearly defined and further elaborated in the form of CD+E, or by using ground level concepts (which could be a SA). For instance, in the sentence "*first place palm near center of Door*," all symbols such as "*first*," "*place*," "*palm*," "*near*," "*center*," "*of*," and "*Door*" have to be explicitly defined via the CD+ framework [12].

The sentence "*Person pushes the door open*" connotes some causality that is not explicit in the sentence. The implicit concept involved is "*Person pushes the door and it causes the door to open*." This causality is shown inFigure 1(a), as a vertical arrow with a line down in the middle. The horizontal double arrow with a line across the middle represents the "state" of something, in this case, the state of Door is Open after being pushed. The concept of Open also needs to be grounded, in this case in the form of an SA. Figure 1(b) shows a more complex conceptualization that involves the concept of WANT. In a sentence such as "*Person wants X*," where X could be an object (say, "*ice cream*") or it could be another conceptualization (say, "*the house to be demolished*"), the implicit causality is that "*if X is obtained or can be realized, Person will be pleased*." (The *object* of WANT is indicated with a link labelled with an "o") "*Pleased"* is a fundamental ground level concept capturing the emotional state of the person involved. It is considered a basic emotional state.

In Figure 1(b), the vertical double arrow with a gray box around it represents any conceptualization (such as "*House be Demolished*"), and when Person "WANT" that, the concept of WANT is elaborated by CD+E into a causal connection between, say, "*House be Demolished*" (the same conceptualization of the object of WANT) and Person being in the state of "*Pleased"*. The "c" above the horizontal gray box represents the conditional ("if"), and "f" represents the future tense – that is, "*if the house is demolished, it will cause Person to be Pleased*." Now, this entire causation is in turn a conceptualization created in Person's internal mental processes. Therefore, this entire conceptualization is the object of what Person conceptualizes - CONCP (CONCePtualize). There are other more complex CD+ constructs discussed in [12] but these examples would suffice for our subsequent discussion.

3.2 The UGALRS Architecture

As mentioned above, CD+ representations operate within a general autonomous system (or social agent) architecture, the UGALRS, for the representation of concepts [12]. Figure 2 shows part of the full UGALRS architecture that focuses on the language aspects. The focus of our attention is on the LANGUAGE COMMUNICATION REASONING CORE (LCRC) but it is by no means the only module that is important. The reason why this module occupies a larger space in this figure and has its details – the sub-modules – illustrated is that the CD+ representations that we will be using for illustrating the concepts involved in language communication reference these submodules in LCRC. The submodules in LCRC are PROBLEM SOLVING (PS), SIMULATION (SM), BUFFER(BF), and CONTROL(CT) modules. In the full UGALRS [12] there is a corresponding module, REASONING CORE (REAC) to LCRC, whose major input and output are the PERCEPTION and ACTION systems respectively (called the SENSORY AND ACTION CORE – SAAC) that are non-language related, but involve the prototypical perception and action processes. The basic idea behind UGALRS views the language communication process as similar to the usual "perception and action" processes related to vision and robotic limb action, but in the sphere of language, such that "language understanding" in this context corresponds to "perception," while "language generation" is the analogous "action." These are done through the LANGUAGE SENSORY AND ACTION CORE (LSAC) here. To initiate an utterance, an intelligent autonomous system (IAS) would begin with some motivation, thus, the MOTIVATION CORE (MOTC). (On the REAC side, the MOTC would drive its reasoning to either understand the perceived information or to generate physical actions.)

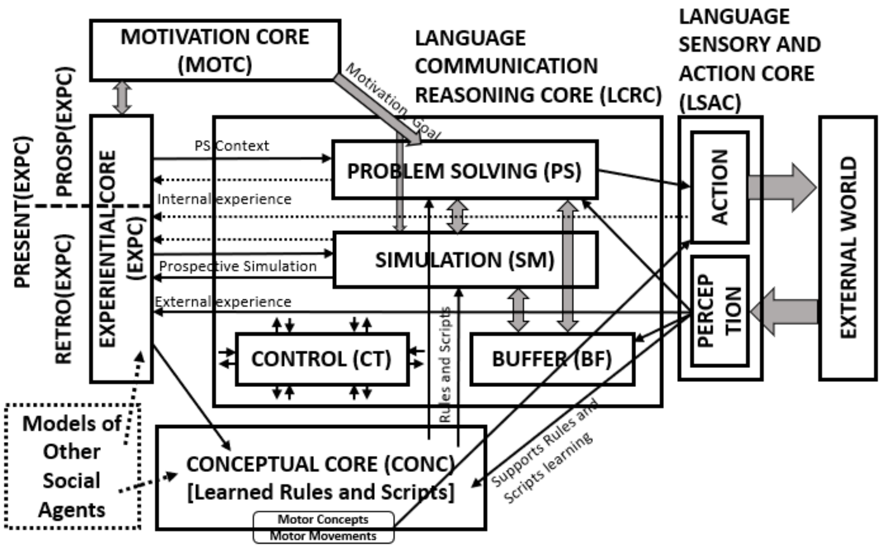


Fig. 2. The UGALRS architecture for an Intelligent Autonomous System (IAS), a robot, or a social agent. Based on [12]. See text for explanation.

PS directs a problem-solving process to concoct an appropriate sentence to prospectively satisfy the motivation involved (like concocting an appropriate physical action through PS on the REAC side). SM is used to anticipate, based on some earlier learned language communication rules, what language responses from the intended recipient of the current utterance are expected. BF contains the concepts that are currently being operated on. Finally, CT directs and control all other modules.

The EXPERIENTIAL CORE (EXPC) records all experiences, linguistic or otherwise. Therefore, there is a path from the PERCEPTION module to EXPC. EXPC also records internal experiences such as what takes place in PS, SM, or what actions (utterances) are (were) emitted. EXPC also provides the context for PS, and directs prospective simulation in SM and receives its results. EXPC can be divided into 3 portions – the present (PRESENT(EXPC)), the past (RETRO(EXPC)) and the future (PROSP(EXPC)). RETRO(EXPC) and PROSP(EXPC) can be used to ground the concepts of past and future as illustrated in [12].

The CONCEPTUAL CORE (CONC) stores knowledge as rules and scripts represented in the form of CD+. Scripts are long causal sequences of events that pertain to certain knowledge complexes, as articulated in [24] and explained in the context of CD+ in [12]. In Figure 2, it is shown that "Models of Other Social Agents" are stored in EXPC and CONC, in EXPC in the form of un-generalized instances, and in CONC, in the form of generalized knowledge. The dotted box of "Models of Other Social Agents" is not a functioning module of UGALRS, but serves to indicate the knowledge involved and where it is located in UGALRS. The sources of external knowledge for EXPC and CONC go through the PERCEPTION module in LSAC.

4 Representations for Language Communication

Armed with the devices and constructs provided by UGALRS and CD+, in this section we illustrate their uses in representing the complex and intricate processes involved in language communication. It will be seen that between just a few sentences, many processes take place in the internal reasoning and problem-solving modules of the utterer, be it human or robotic (i.e., an IAS or a robotic social agent). These processes involve not only the conceptual, but also the motivational and affective, that can be represented by CD+ within UGALRS.

In the following section, we will represent the internal conceptual, motivational, and affective processes in both Person and Robot using CD+. Even though the primary purpose here is to elucidate the computational and representational processes in IAS (robots), there are two purposes in elucidating similar processes in the Person involved as well. First, Person can itself be another robot engaging in natural language communication with the first robot. Second, a robot or IAS can model the "mental" processes of another person or robot as well, which is the block indicated in the bottom left corner of Figure 2. Therefore, in the following, we elucidate the processes taking place in the person as well as the robot.

4.1 Motivation and Sentence Concoction/Generation

Before an utterance is made, the utterer must begin with an idea in mind. This idea could be just a thought to be shared, or a want to be conveyed. Suppose a person (Person) thinks of asking a robot (Robot) to bring her a tool, Tool(X), from the table. This would be her "WANT," which if the robot could succeed in listening to her and satisfying it, she would be *Pleased*. At the very top part of Figure 3, this is represented after the same fashion as the representation of Figure 1(b). Firstly, here are SAs associated with Person and Robot, and SAs such as these will be omitted in subsequent figures to avoid clutter. The PTRANS concept is Physical TRANSfer, used and explained in [22] and [12]. There is a "from" location (Loc(Table)) and a "to" location (Loc(Person)) and a Direction (D) of transfer. So, in this case, Person conceptualizes (CONCP) that if Robot were to PTRANS Tool(X) from Table to her, she would be *Pleased*. "I" on the right most side of the PTRANS representation represent the "Instrument" that Robot might use for this purpose, such as using its legs to propel itself along the ground. The PTRANS process involves a series of steps. Suppose Robot is currently next to Person and Table is some distance away, Robot would first turn its body and face Table, viewing from far to see that Tool(X) is on Table, then mentally, through a PS process, plot a path to Table, and after executing, say, a pickup action, bring Tool(X) to Person. This sequence of events is first worked out in Robot's REAC module (i.e., the usual non-language related problem solving and planning process). This is the HOW in the CD+ Elaboration (CD+E) pointed to from PTRANS. The desired Pleased state in the WANT conceptualization is a motivational force that propels Person to proceed to look for solutions to realize the *Pleased* state. *Pleased* is a basic and ground level emotion, as discussed in [12].

Now, as discussed in [12], when a WANT is conceptualized, there may or may not be a solution to satisfy the object of the WANT. Therefore, the exact "HOW" is "IRRELEVANT" in the concept of WANT. Hence, there could be a situation in which "I want to get rich but I can't." If a solution exists, then the concept of CAN comes into play. So, if "Robot CAN bring Tool(X) from Table to Person," then it means a solution exists. This representation of CAN is given in [12] and will be shown in Figure 4. The entire CONCP encased in a box is called a MOTIVATION CONCEPT (M-CONC).

Next, having the WANT of a certain event (namely the Robot bringing Tool(X) to the Person), Person then WANTs to communicate the concept that she WANTs this certain event to happen to Robot. There is a general rule that says if an IAS (a human is a natural IAS while a robot is an artificial one) wants something, a motivation to achieve the state of "*Pleased*" will drive the IAS to do one of three things: 1. Carry out a planning or problem solving process to achieve the state of Pleased by herself or itself; 2. Request the help of someone else to do so; and 3. Command a servile agent to do so. This is encoded as the first causal link near the top of the figure. Now, note that at the very top of the figure, the WANT conceptualization (the topmost double arrow) is encased in a gray box with a label "1." This entire conceptualization "1", that the Person wants Robot to do something, is now the object of an MTRANS (Mental TRANSfer) process intended to "mentally" transfer the WANT conceptualization "1" from Person to Robot, that will make Person *Pleased*. (I.e., Person WANTs her WANT, currently in her mind, to be MTRANS to Robot's mind. In computational terms, "mind" is simply the internal memory and processing mechanisms of the human or robot involved.)

In order to realize this communication, Person concocts a sentence by Mentally BUILDing (MBUILDing) the sentence from considering the conceptualization involved (labeled "1"), together with the grammar of the language involved, the intended tone of the sentence, the emotional state of Person, etc. (the intended tone is dependent on the existing context of communication). The sentence constructed is "Robot, please bring me Tool(X) from the table," as shown in the figure as comceptualization "2". (The fact that Person WANTs Robot to do something is not explicitly stated in the sentence, but it is implied. Person could also have stated more explicitly, "Robot, I want you to bring me Tool(X) from the table.) The MBUILD concept is discussed in [22] and [12]. The MBUILDing of the sentence takes place in BF(LCRC(PERSON)). The precise process of converting an internal meaning representation in CD+ form to a grammatical surface sentence for communication is relegated to future work.

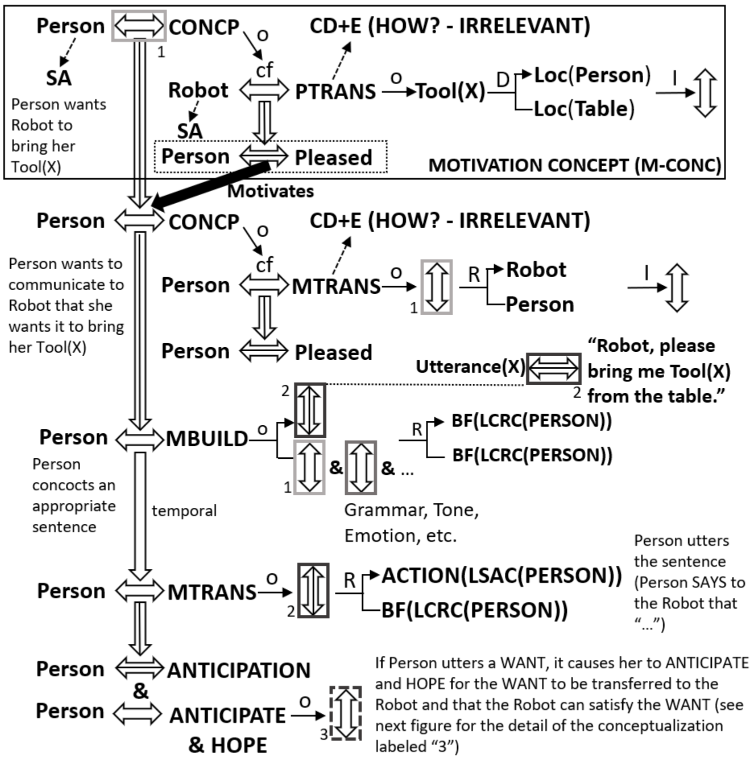


Fig. 3. The conceptual, motivational, and affective processes involved in the utterance "Robot, please bring me Tool(X) from the table." See text for explanation. The black "Motivates" arrow is not part of the representation, but an indication of the source of the causal link involved.

After the sentence is concocted in BF(LCRC(PERSON)), it it MTRANSed to ACTION(LSAC(PERSON)) to be emitted as an utterance. The link between MBUILD and MTRANS is a "temporal" one, not a causal one, as the second step simply follows the first step as part of the process (temporal links are indicated as a thickened arrow without a line running down its middle).

When a sentence, whether one that is a command or request, or just a factual statement, is uttered toward a recipient, a state of ANTICIPATION is entered (which is part of the implication of command or request) and the utterer then ANTICIPATEs something, as shown in the bottom of Figure 3. This is the first affective state that emerges in the present communication process and will be discussed in detail in the next section. This ANTICIPATION is accompanied by HOPE as the prospect is positive [25].

4.2 Affective States and Illocutionary Forces

In the non-linguistic sphere, when a certain action is emitted by an IAS, it is expected to cause certain effects. Similarly, in the linguistic sphere, an emitted utterance is expected to cause some effects. This has been investigated by speech act theorists [26]. If an utterance is meant to communicate certain information to the recipient, there may be no immediate overt actions or responses expected in the recipient, but the information conveyed may cause future actions or responses, or in the least, it may cause certain changes in the beliefs of the recipient. If the utterance is in the form of a command or request, immediate actions and responses are expected. The utterance is said to have an "illocutionary force." [26]

At the bottom of Figure 3 we show that Person enters am affective state of ANTICIPATION and she ANTICIPATEs something. In Figure 4(a) we show the functional representation of ANTICIPATE, an action that accompanies the affective state ANTICIPATION. If an Agent ANTICIPATEs a certain conceptualization, she conceptualizes that the conceptualization involved will happen in the future. It is shown in Figure 4(a) that the object of Agent's CONCP is labeled with an "f", which means it resides in the prospective part of the EXPC (Figure 2), PROSP(EXPC). This formulation of an affective state and its associated causal consequences is in consonant with the cognitive appraisal theory of emotion [25]. The same approach will be adopted with the other affective states in subsequent discussions. Specifically, in the situation depicted in Figure 3 in which Person asks Robot to bring her Tool(X) from Table, she ANTICIPATEs both the facts that "Robot WANTs to PTRANS Tool(X) from Table to Person so that Person is *Pleased*" and "Robot CAN PTRANS Tool(X) from Table to Person so that Person is *Pleased*," as shown in Figure 4(b).

First, let us consider the representation for "Robot WANTs to PTRANS Tool(X) from Table to Person so that Person is *Pleased*." This is a transfer of what Person WANTs to what Robot WANTs. Now, for Person to reasonably assume that Robot would WANT to *Please* her, it must be assumed that Robot has either a SERVILE or an ALTRUISTIC attitude. In situations in which Robot or other recipient(s) of the utterance is REBELLIOUS or UNCOOPERATIVE, then this situation does not obtain.

In Figure 5 we depict that Robot is indeed SERVILE or ALTRUISTIC and hence in Figure 4(b) Person ANTICIPATES that "Robot WANTs to PTRANS Tool(X) from Table to Person so that Person is *Pleased*." Hence, Robot being *Pleased* is caused by Person being *Pleased*. CD+ can be used to represent the connections between attitudes such as being SERVILE, ALTRUISTIC, COOPERATIVE, REBELLIOUS, or UNCOOPERATIVE and whether the entity/IAS involves WANTs to do certain things. The details of these are left to future work.

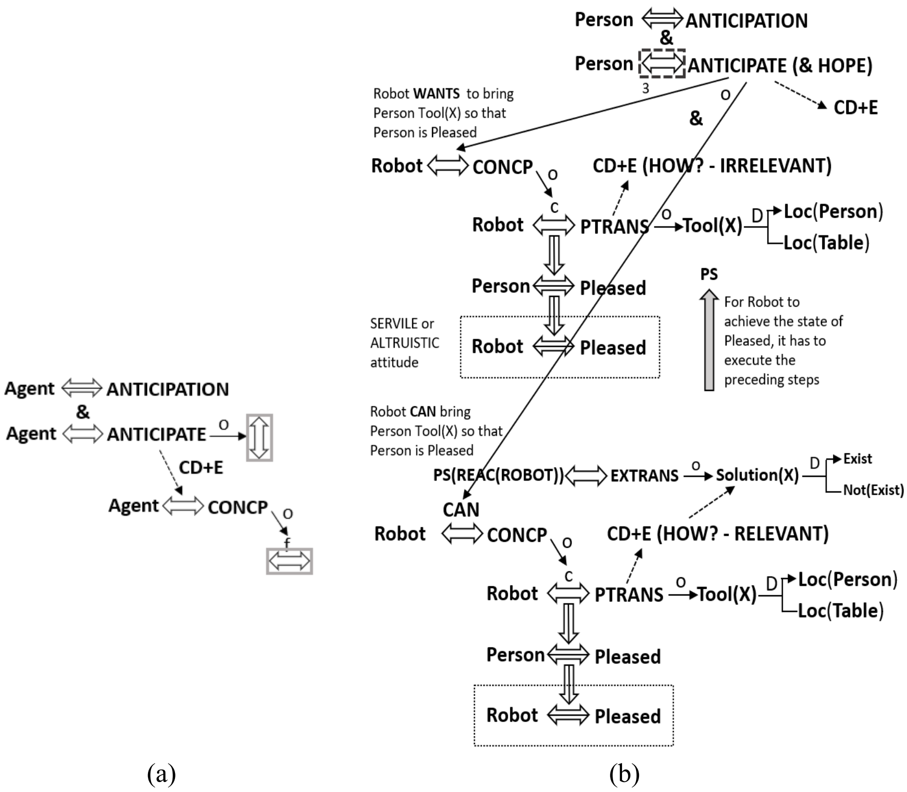


Fig. 4. (a) The general representation of ANTICIPATION. There is a state of ANTICIPATION, and an object referred to by the ANTICIPATE action. (b) The specific case of ANTICIPATION at the bottom of Figure 3.

The locus of the illocutionary force is this. In any IAS, ultimately it will do whatever *Please*s it. The arrow labeled PS shows the flow of Robot's actions: in order to please itself, it has to please Person, and in order to please Person, it has to PTRANS Tool(X) from Table to Person, if it CAN.

As mentioned above, there is a difference between WANT and CAN [12]. The primary difference is that WANTing something to happen (e.g., PTRANSing something from one place to another) does not imply that a solution exists for the thing to happen, but CAN implies that the solution exists.

Therefore, there could be a situation that "I want to go from here to there but I can't". Hence, the representation for "Robot CAN PTRANS Tool(X) from Table to Person so that Person is Pleased" shown in Figure 4(b) is that PS(REAC(ROBOT)) returns a solution (Solution(X)) for Robot to PTRANS Tool(X) from Table to Person. (EXTRANS stands for EXistential TRANSformation in which something goes from non-existence to existence or vice versa – which is used to represent the existence of a Solution(X) – see [12]).

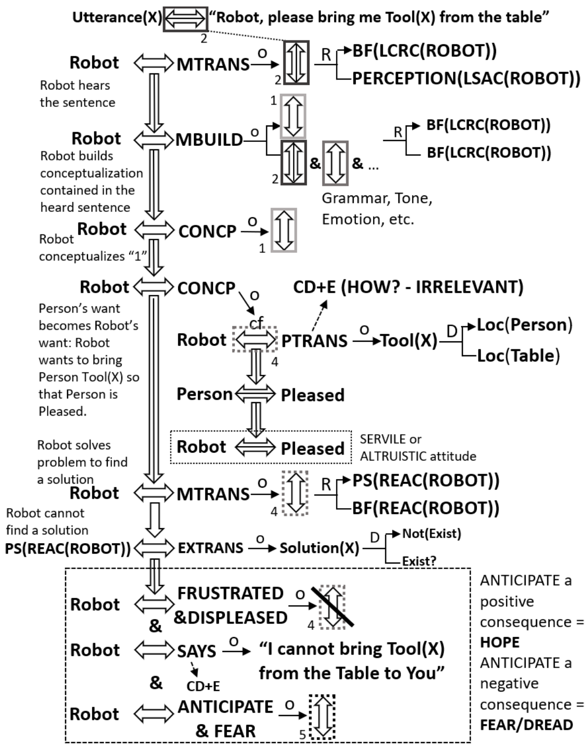


Fig. 5. Robot's internal processing in response to Person's utterance in Figure 4. See text for explanation. Unlike in Figure 3, ANTICIPATE here is accompanied by FEAR as it is anticipating a negative prospect [25].

4.3 Sentence Understanding, Actions, and Affective States

Now that the utterance has been made and presumably received by Robot, the first step of the process on Robot's side is to MTRANS the received utterance into BF(LCRC(ROBOT)) for further processing as shown in Figure 5. This causes Robot to MBUILD the conceptualization corresponding to the received utterance, based on the grammar of the language, the tone present in the sentence, the current emotional state of Robot and the perceived emotional state of Person, etc. This MBUILDed conceptualization is labeled "1", which is the same as conceptualization "1" in Figure 3. Assuming Robot has either a SERVILE, COOPERATIVE, or ALTRUISTIC attitude, this causes it to create the next conceptualization capturing the fact that Robot will be *Pleased* if Person is *Pleased* due to Robot carrying out a certain task. This then motivates Robot to seek a solution to conceptualization "4", which is "Robot PTRANS Tool(X) from Table to Person." This conceptualization is MTRANS from BF(REAC(ROBOT)) to PS(REAC(ROBOT)).

Suppose PS(REAC(ROBOT)) cannot find a solution subsequently. This situation is represented in CD+ using EXTRANS showing that a solution does not exist (see [12] for the concept of CANNOT). It causes Robot to enter states of FRUSTRATED, DISPLEASED and FEAR (unlike for the case of ANTICIPATION in Figs. 3 and 4(a), these are not shown in Figure 5 to avoid clutter) and it is also FRUSTRATED, DISPLEASED, and FEARful about the un-attainment of conceptualization "4" as shown in Figure 5 ("un-attainment' is represented as a slash across the conceptualization and is directly related to the concept of CANNOT). FRUSTRATION and the other emotions can also arise if PS(REAC(ROBOT)) can find a solution but Robot cannot execute it due to other situations that are not anticipated in the PS process.

The state of FRUSTRATION always follows the situation when an Agent WANTs something but it cannot be obtained, as shown in Figure 6. The state of *Displeased* also accompanies this based on a rule that states that if the object of a WANT is not achievable or satisfiable, the IAS involved will be Displeased. FEAR comes from the fact that Robot has a model of Person's negative response to the un-attainment of her WANT, and it is reflected in its ANTICIPATION of conceptualization "5", which is shown in Figure 7 as "Person is DISAPPOINTED and DISPLEASED that conceptualization "4" is not attainable." FEAR is the anticipation of a negative consequence that may happen to the agent itself [25]. It is not shown here that a DISAPPOINTED and DISPLEASED Person toward Robot may take negative actions toward it in some way.

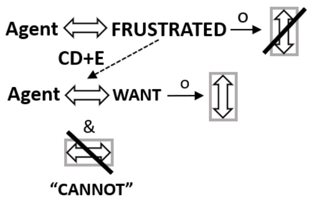


Fig. 6. The meaning and representation of FRUSTRATED

Other than the consequences above, another response to not being able to find a solution includes Robot communicating this to Person by uttering "I cannot bring Tool(X) from the table to you" as shown in Figure 5. The symbol "SAY" has a CD+E that is the same processes in Figure 3 when Person concocts and utters a sentence, but we omit the detailed CD+E here. If instead Robot is able to find a solution, it would go ahead to carry out the task and say "Here is Tool(X)" when handing it to Person.

What motives Robot to explain its failure is the communication rule that states: "if others are displeased with your failure to do something on request or command, do communicate about it, including explaining the reason involved, because this will placate the other person, which in turn should reduce your own frustration, displeasure, and fear." This entire rule could be stated in CD+ for the system to interpret and execute.

4.4 Continuing Communication

Following Robot reporting that it is not able to bring Tool(X) to Person, Person enters the state of being DISAPPOINTED and DISPLEASED and the object of the DISAPPOINTment and DISPLEASure is the un-attainment of conceptualization "4", as shown in Figure 7. Instead of just keeping quiet, which is a possible response on the part of Person if she is no longer concerned about the un-attainment of "4" or she is taking some time to ponder her response, a typical immediate response on the part of Person is to try and understand the cause of the un-attainment of "4." To this end, Person asks "Why can't you bring Tool(X) to me?" as shown in Figure 7.

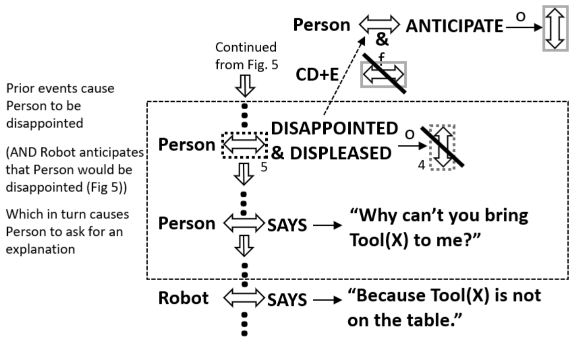


Fig. 7. Possible communication continued from Fig 5.

As the UGALRS and CD+ representational framework articulated in [12] is a fully explainable framework, when in the problem-solving process, PS(REAC(ROBOT)) fails to return a solution, because the steps of processing everywhere in UGALRS using CD+ representations are explicit, the cause(s) of the PS failure is easily identified. Hence, the Robot would respond with "Because Tool(X) is not on the table." What causes Robot to respond is the illocutionary force present in Person asking the Why question (i.e., it is in Robot's CONC, where general knowledge is stored – Figure 2 – that Person would be *Pleased* if her Why question is answered to, and would be *Displeased* if this is not so. This knowledge is also represented in CD+ form in CONC. These are the representations of the illocutionary forces involved.) Robot may feel further RELIEVED from being FRUSTRATED, DISPLEASED, and FEARful after providing this explanation, because providing an explanation may cause Person to be more Pleased.

5 Discussion and Conclusion

Even a small set of relatively simple utterances by an intelligent system involves many complex and intricate processes encompassing conceptual, motivational, and affective aspects. In this regard, we used UGALRS and CD+ to elucidate some of these processes. For the sake of clarity and simplicity, some processes have been omitted in these diagrams, but the texts have discussed some of them, particularly the communication rules underlying the generation of sentences. Future research could delve further into explicitly representing these rules. It is worth noting that CD+ is comprehensive enough to represent these rules and the situations under which they are triggered – i.e., the reasoning processes themselves are also representable using CD+, as has been amply illustrated in [12].

Psychologists have identified up to 161 types of motivations in humans [27]. For robots, the number of motivations could be simpler and smaller in number [19]–[21]. However, for an IAS or robot to understand humans and hence be able to interact with them effectively, it could benefit from having a model of the humans' motivations, as shown in Figure 2 and discussed in this paper. To contain the discussion, in this paper we have only dealt with a small number of motivations; future work certainly calls for expansion in this direction, and as already demonstrated in [12], it is possible to do this within a UGALRS plus CD+ framework.

It should be noted that the topic of emotions would entail a broader selection of plausible emotions that could similarly be useful for characterizing and communicating about internal states in robots and humans than what has been discussed in this paper [10], [28], [29], [30]. For instance, while Ortony's cognitive appraisal theory of emotion [25], which is amenable to computational treatment, has been partially capitalized here, there is a fuller set of emotions that has not been addressed in this paper. In future, a more comprehensive selection of emotions and treatment of affective processes could enable intelligent systems to handle a wider range of communicative scenarios. Despite the fact that this paper only covers a subset of these vast conceptual, motivational, and affective spaces, its main contribution is to articulate a general framework of knowledge representation and processing to link these together and elucidate the respective functions they serve in the complex process underlying natural language communication.

Other important future work includes: 1. The transformations between the surface sentences illustrated in many places in this paper and their corresponding deep level "meaning" representations (Section IV(A)). This has also not been fully developed in Schank's original CD work [22]–[24]. The transformation must take into consideration grammar, tone, emotion, etc. 2. The roles, representations and causal consequences of various attitudes such as SERVILE, UNCOOPERATIVE, etc. (Section IV(B)). 3. Extension of the current framework and paradigm to cover a wider range of communication. 4. A computational implementation of the representation and processes involved. 5. The learning of the various representations illustrated in this paper.

Even though a computational implementation of CD+ is not reported in this paper, its viability as a computationally tractable representational scheme is reflected in the original work of Schank and colleagues [22]–[24], which has demonstrated that a computational implementation of CD could handle natural language question-answering and communication processes that benefit from the deep meaning representations of CD. It follows then that CD+, as an enhancement of CD, will share similar prospects for computational tractability.

As a final discussion point, the aspect of learning will be an important consideration. Given the complexity of the representations discussed in this paper, a system that lacks the ability to learn could not be scaled up nor become a practical system. While Ho [12] has discussed how learning could be done in the framework of CD+, we must first understand the kind of representations that are needed for intelligent processes, in this case, language communication processes, before we understand what it is that is to be learned. This paper hopes that by contributing to the elucidation of the intricate and complex conceptual, motivational, and affective processes involved in natural language communication between social agents, it could, when appropriately extended, hopefully bring about a fuller characterization of language communication in general.

References

[1] A. Clark, C. Fox, and S. Lappin, Eds., *The Handbook of Computational Linguistics and Natural Language Processing*. Hoboken, New Jersey: Wiley-Blackwell, 2012.

[2] R. Mitkov, Ed., *The Oxford Handbook of Computational Linguistics*. Oxford: Oxford University Press, 2005.

[3] Z. Wang, Z. Hu, S.-B. Ho, E. Cambria, and A.-H. Tan, “MiMuSa-Mimicking Human Language Understanding for Fine-grained Multi-class Sentiment Analysis,” *Neural Comput Appl*, 35, pp. 15907–15921 (2023.

[4] T. B. Brown and et. al., “Language Models are Few-Shots Learners,” 2020. doi: doi.org/10.48550/arXiv.2005.14165.

[5] E. Cambria, A. Hussain, C. Havasi, and C. Eckl, “Sentic computing: exploitation of common sense for the development of emotion-sensitive systems,” in *Lecture Notes in Computer Science*, 5967, Springer, 2010, pp. 148–156.

[6] E. Cambria, Q. Liu, S. Decherchi, F. Xing, and K. Kwok, “SenticNet 7: A Commonsense-based Neurosymbolic AI Framework for Explainable Sentiment Analysis,” in *Proceedings of LREC*, 2022, pp. 3829–3839.

[7] R. Mao, Q. Liu, K. He, W. Li, and E. Cambria, “The Biases of Pre-Trained Language Models: An Empirical Study on Prompt-Based Sentiment Analysis and Emotion Detection,” *IEEE Trans Affect Comput*, 2023.

[8] K. He, R. Mao, T. Gong, C. Li, and E. Cambria, “Meta-based Self-training and Re-weighting for Aspect-based Sentiment Analysis,” *IEEE Trans Affect Comput*, 2023.

[9] A. Kumar, T. Trueman, and E. Cambria, “Gender-Based Multi-Aspect Sentiment Detection using Multilabel Learning,” *Information Science 606*, pp. 453–468, 2022.

[10] Z. Wang, S.-B. Ho, and E. Cambria, “A review of emotion sensing: Categorization models and algorithms,” *Multimed Tools Appl*, vol. 79, pp. 35553–35582, 2020.

[11] Z. Wang, S.-B. Ho, and E. Cambria, “Multi-level fine-scaled sentiment analysis with ambivalence handling,” *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 28, no. 4, pp. 683–697, 2020.

[12] S.-B. Ho, “A general framework for the representation of function and affordance: A cognitive, causal, and grounded approach, and a step toward AGI,” 2022. doi: 10.48550/arXiv.2206.05273.

[13] N. Howard and E. Cambria, “Intention Awareness: Improving upon Situation Awareness in Human-Centric Environments,” *Human-centric Computing and Information Sciences*, vol. 3, no. 9, 2013.
[14] L. Talmy, *Toward a Cognitive Semantics Volume I and II*. Cambridge, Massachusetts: The MIT Press, 2000.
[15] V. Evans and M. Green, *Cognitive Linguistics: An Introduction*. Mahwah, New Jersey: Lawrence Erlbaum Associates, 2006.
[16] R. W. Langacker, *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford University Press, 2008.
[17] D. Geeraerts, *Theories of Lexical Semantics*. Oxford: Oxford University Press, 2010.
[18] J. van Eijck and C. Unger, *Computational Semantics with Functional Programming*. Cambridge: Cambridge University Press, 2010.
[19] B.-K. Quek, “A Survivability Framework for Autonomous Systems,” Ph.D. Thesis, National University of Singapore, 2008.
[20] K. Quek, J. Ibañez-Guzmán, and K.-W. Lim, “Attaining operational survivability in an autonomous unmanned ground surveillance vehicle,” in *32nd Annual Conference on IEEE Industrial Electronics*, 2006, pp. 3969–3974. doi: 10.1109/IECON.2006.348001.
[21] K. Quek, J. Ibañez-Guzmán, and K.-W. Lim, “A survivability framework for the development of autonomous unmanned systems,” in *9th International Conference on Control, Automation, Robotics and Vision*, 2006, pp. 1–6. doi: 10.1109/ICARCV.2006.345336.
[22] R. C. Schank, “Identification of conceptualizations underlying natural language,” in *Computer Models of Thought and Language*, R. C. Schank and K. M. Colby, Eds., San Francisco: WH Freemann&Company, 1973, pp. 187–247.
[23] R. C. Schank, *Conceptual Information Processing*. Amsterdam: North-Holland Publishing Company, 1975.
[24] R. C. Schank and R. P. Abelson, *Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structure*. Mahwah, New Jersey: Lawrence Erlbaum Associates, 1977.
[25] A. Ortony, G. L. Clore, and A. Collins, *The Cognitive Structure of Emotions*. Cambridge: Cambridge University Press, 1990.
[26] J. R. Searle, *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press, 1970.
[27] J. Talevich, S. Read, D. Walsh, R. Iyer, and G. Chopra, “Toward a comprehensive taxonomy of human motives,” *PLoS One*, vol. 12, no. 2, 2017.
[28] R. Plutchik, *Emotions and Life: Perspectives from Psychology, Biology, and Evolution*. American Psychological Association, 2002.
[29] Y. Susanto, A. Livingstone, B. Ng, and E. Cambria, “The Hourglass Model Revisited,” *IEEE Intell Syst*, vol. 35, no. 5, pp. 96–102, 2020.
[30] M. Amin, E. Cambria, and B. Schullerl, “Will Affective Computing Emerge from Foundation Models and General Artificial Intelligence? A First Evaluation of ChatGPT,” *IEEE Intell Syst*, vol. 38, no. 2, pp. 15-23, 2023.