AFFECT AS MOTIVATION FOR COGNITIVE AND CONATIVE PROCESSES

William S. Faught
Stanford Artifioial Intelligenoe Laboratory
Stanford, California USA

Abstract

Responding to natural language requires performing two major tasks: information processing (cognition), and decision making (conation). In order to motivate and best guide an entity performing such complex tasks, motivation must originate in the self-interest of the entity. This motivation guides cognitive and conative processes at the lowest possible level by assigning values (measures of importance) to various process components and using them in making process decisions. Through such a motivational mechanism, every decision taken will be to further the entity's self-interests. We describe a model for such a mechanism and the affect process which embodies self-interests.

Introduction

Analyzing natural language input to determine its underlying semantic content is but one of the tasks necessary for an entity (human or non-human) to use natural language. Responding to natural language input requires two major tasks: 1) deriving facts about the input and the situation in which it is expressed, based on perceptual information, on the input itself, and on previous information (cognition); and 2) establishing goals for satisfying the entity's needs, desires and interests, deciding upon a plan of action to satisfy these goals, and executing these actions (conation). The motivation for the entity's actions arises not from within these two functions, but from a third function known as affect (emotion). Affect provides motivation for the entity by embodying all experience that has immediate meaning or significance to the entity. These experiences in turn motivate the cognitive and conative tasks.

We shall describe a model of the affect process and its role in motivating cognitive and conative processes. These processes are part of a model of a person whose thought processes are dominated by the paranoid mode of thought. Paranoia is diagnosed clinically from the linguistic behavior of an individual. Thus, the major task of the model is to respond to natural language input. As a result, the processes described here should be useful both in simulating cognitive processes and in natural language understanding systems.

Background

In simulating the linguistic behavior of a paranoid person in an interview situation, we found three considerations indicating the importance of modelling affect and its influence on other processes: 1) the requirements for an entity participating in an interview, 2) the centrality of affect in the theory of paranoia we are modelling, and 3) the desirability of making our theory more complete and consistent.

The requirements for an entity to participate in a psychiatric interview motivated our interest in cognitive, conative, and affective processes. When participating in an interview, a person brings many preconceived ideas about the purpose of such interviews and what happens in them. He also has expectations about how the interviewer will respond to his actions. He has needs and desires that originate from a global self-interest for survival. These motivate his interview participation by specifying goals that he wants to accomplish during the interview. He has a conative ability to structure these goals, form plans about how he intends to satisfy goals, and execute actions to carry out plans. He has a cognitive ability to observe and evaluate the actions that are taking place to determine whether his goals are being achieved or whether he must cope with some new situation. His needs and desires are tied to the success or failure of his actions as well as to the interaction of the participants in the situation. Finally, he has the ability to posit new goals and new courses of action, possibly altering the course of the interview. As he attempts to steer events, he continues to perceive the ongoing situation and measure the success of his actions. Note that the motivation for a person's participation originates in his self-interests, and expands to motivate other processes.

Affects are central to the theory of paranoia we are modelling [Colby, 1975], in that affects may radically modify the workings of other processes. In the paranoid mode, a person cognitively scans natural language input, as well as the inferences from that input, looking for evidence that judges an action, desire, or state of the self to reflect an inadequacy or defectiveness of the self. Upon finding such evidence, the person simulates acknowledgement of this inadequacy. If the inadequacy

were accepted or acknowledged as true, the "painful" affect of humiliation would result. The detection of this impending humiliation in the simulation serves as a warning not to execute an acknowledging procedure. Instead the conative ability attempts an alternative simulation in which wrongdoing is attributed to others. Since no warning signal of the affect of humiliation results, the procedure for blaming others is executed. The outcome of this alternate strategy is 1) to repudiate that the self is to blame for inadequacy, and 2) to ascribe blame to other human agents. This transfer of blame is reflected in the ongoing linguistic behavior of a paranoid patient in a psychiatric interview. Note the central role of affect in the paranoid process.

Finally, in striving for completeness and consistency in our model and its underlying theory, we recognized the importance of modelling all of the phenomena in one formalism, from the very top-level global strategies to the simplest cognitive operation For a theory to be complete, it must explain all behavior associated with the phenomena being modelled. A person dominated by the paranoid mode is nevertheless subject to periods of normalcy, and this normalcy must also be adequately modelled. Also, a person has a general background strategy behind his interview participation strategy. A person in an interview situation not only has specific desires and goals of what he wants from the interview, but also has an general strategy for survival which is motivating all action in the system and must also be modelled

For a theory to be consistent, it must explain similar phenomena in similar formalisms. Thus the normal and paranoid modes of processing should be modelled in the same formalism. This principle also applies to the two functions of cognition and conation — present in some form in almost any task involving human thought processing. For consistency, we would like motivations for one type of task ["find all x (moon rocks, blocks, bacteria) for which P(x) is heavy, is red, is toxic"] represented and structured in the same manner as other motivations ["whenever an input is typed on the teletype, process it" or "if x is a new rule of inference, store it for later use"]. A single strategy should motivate all cognitive and conative processes in a similar manner.

Affeot: The Model

Our model of affect (emotion) is derived from differential emotion theory [Izard, 1971]. According to that theory, affect is one of the five subsystems of personality (the others: homeostatic, drive, cognitive, motor). The function of affect is to provide the principal motivation for human behavior, by invoking and directing other processes The affect system consists of : 1) several discrete affects, each having its own distinct motivational and phenomenological properties, (specifically, the distinct sets of phenomena that activate them and their distinct motivational responses); 2) a strategy for choosing which affects to favor in order to maintain and enhance the state of the self; and 3) a process that activates individual affects, applies the choice strategy, and invokes other processes as a response. The eight affects we distinguish [Tomkins, 1962] are three positive: enjoyment-joy, interest-excitement, and surprise-startle; and five negative: fear-terror, anger-rage, distress-anguish, shame-humiliation, and contempt-disgust. The terms positive and negative are not meant to convey an ethical judgement, but instead express a global strategy of self-maintenance and self-enhancement. The affects refer to pleasure and pain, either physical, mental, or emotional pleasure or pain (e.g., fear of physical pain, interest in emotional pleasure). The strategy is to maximize positive affects and minimize negative ones. This self-interest strategy determines how affect will respond in directing other processes to satisfy affect requirements.

Affect has complex chemical, neurophysiological, neuromuscular, and phenomenological aspects. Affects are activated by perceptual activities such as perceiving an object or situation of importance in a person-environment interaction, or by cognitive activities such as memory or imagination of previous and future events. (We postulate that both activities function as activators in the same manner.) Less common affect activations result from spontaneous motor, endocrine, or neuromuscular activity. Activation entails establishing patterns of electrochemical activity in the nervous system (through the limbic system) for a particular affect. These patterns initiate neuromuscular activity, primarily in facial activity and facial patterning, and secondarily in body response.

When' this activity is sensed at the phenomenological level, it becomes a conscious experience of a discrete affect which is meaningful and significant by itself. The power of affect is in this experience; the experience is intrinsically rewarding for positive affects and punishing for negative ones. The experience is conscious (in the sense of being

independent of cognition, not as opposed to subconscious): here we find the distinction between drive and affect. Drives constitute needs of the system, but drives only motivate through affect. Drives send information about the entity's physical needs to the affect process. Affect may amplify, attenuate, or ignore the information, thereby modulating the impact of the drive on consciousness. (Without the conscious experience of the need, the need has no motivating influence.) Thus affect can motivate without drive, but drive cannot motivate without affect.

Being the only motivational process, affect must match the complexity of the cognition and conation processes it motivates. The complexity of the affect process arises from a) the variety of cue-producing facial and body responses for each of the eight discrete primary affects, and b) the complex response of affect in interacting with the processes affect motivates. The cues from facial and body reponses are used by cognition to determine the cause of the affect (the experience itself does not indicate the activator, although past experience may be used as described in the next section) and to determine a best possible plan by conation. The richness of cue-differentiation allows cognition to distinguish among the phenomenologtcally distinct significant experiences of affect. We shall see the complexity of the interaction of affect response with other processes in the next section.

For our task of simulating human behavior, it is necessary and useful to model affect as it occurs in human personalities. To construct any entity with a task to perform, it may be useful to construct an affect-like process representing the entity's self-interests. The requirements of such an entity typically entail finding and executing a "best" sequence of actions to perform a task. For the entity to strive to do its best, it must perceive its relative success of performance. The relative success of performance must then take on importance and be significant to the system. The experience of this significance is an immediately rewarding or punishing experience, as in affect. Such experiences become synonymous with the entity's self-interest. Using an affect-like process, motivation for all action would stem from this self-interest, and insure that all action would be taken for the purpose of enhancing the original task. Finally, if the entity performs a task that is complex, we would expect the motivation of the entity towards its task (and more specifically its self-interest) to be equally complex in guiding the entire system.

Affeot as Motivation

In discussing affect as motivation for the entire system, we shall concentrate on the motivation affect provides for cognitive and conative processes. We define cognition as the process by which information is stored in memory, extended (typically by inductive or deductive processes), transformed from sensory input into an internal representation, and made available to other processes As seen from other processes, the function of cognition is to provide information relevant to some task. We define conation as the process by which goals are established, decisions are made as to what actions will satisfy these goals, and the derived actions implemented by performing the appropriate task. Conation is not necessarily a conscious process. Subconscious decisions about subconscious goals can be made because the goal and decision is common (driving a car), or because of extreme affect conditions to be described later. To make the distinction clearer; we theorize that conation is the decision-making and -implementing function of the system for any decision that is not in the hardware of the system. Cognition is the information-translating function of the system, translating sensory input into cognitive structures, manipulating structures internally, and translating cognitive structures to motor output. Conation may establish a goal and call cognition for a plan to satisfy the goal. Cognition may call conation to establish subgoals as part of a plan. Note that part of what usually is termed cognition, e.g. establishing a goal of proving a conjecture and then initiating an attempt of the proof, we term conation. Affect motivates these processes through 3 types of interaction:1) Affect initiates the processes, and interrupts them on condition of extreme affect levels. 2) When interrupting the processes, affect associates significance values (measures of importance) with the new goals or actions, or the new information that caused the interrupt, for later use. 3) Whenever choices are made about the importance of a goal or action (in conation), or importance of information (in cognition), the criteria of importance used to make the choice are from current affects or from significance values originally determined by affect.

Affect's most direct form of motivation stems from its ability to initiate and interrupt other systems. The motivation for the model originates in the rewarding or punishing experiences of affect. When these experiences occur, the affect strategy determines the most important affects to satisfy. If there is not already a conative activity in existence that is satisfying more crucial affects, then affect initiates the conative process to determine and fulfill a goal of enhancing current positive affects or reducing negative affects. Conation

(optionally) initiates cognition as part of its processing to determine goals or actions, or to perceive the results of actions. As an example, in its simplest sequence of processing, the affect process detects the presence of extreme negative affects and initiates a conative process for reducing those affects. Conation immediately chooses an action and executes it, without any call to cognition (e.g., crying after being hurt). A more typical sequence is for affect to note that affect levels are normal and then initiate a conative process with a goal of finding a plan for enhancing positive affects. Conation then uses cognition to determine the plan and monitor the execution of the plan.

Once conation and cognition have been initiated, affect may interrupt either process due to extreme affect levels. Such an interrupt may initiate an entirely new process or redirect an existing process by adding new affect information to a current process. Adding affect information entails attaching importance to a step in a current process based on the intrinsically rewarding or punishing experience of current affects. Affect information provides the criteria for redirecting processes, depending upon the significance of the affect response. For example, if cognition constructs a new belief that invokes a sharp change in affect, affect may interrupt to cause conation to assign to cognition a new task of examining the inferences that led to the belief. Similarly, if conation executes an action that invokes sharp affect change, affect may interrupt conation to derive a new goal more suitable to the new affect conditions.

The homeostatic and drive systems are only indirectly influenced by affect. The homeostatic system is separate from the other systems and carries on its functions according to its own needs. The requirement of those needs is to attain or return to preferred states of the system in spite of repeated oscillations out of the states. Affect's only influence upon this system is from facial and body changes of affect responses. Drive system needs satisfy tissue demands for food, water, air, etc. Drive needs are cyclic in nature and let their needs be known to affect through affect-activation. Because drive needs are only one class of affect activator, other activators may cause affect to ignore drives, attenuate drives, or amplify them, depending upon the other concurrent activators. Affect's response to the drive system is limited to directing conation and cognition to satisfy drive needs.

When interrupting another process, affect assigns significance values (measures of importance) to the various process components that caused the interrupt. By process components we mean beliefs and inferences for cognition, and goals, plans and actions for conation. Significance values are the vehicle for attaching significance or importance to process components. These values originate in self-interest, in the affect experiences that have immediate meaning for the entity. Significance values are the measure of importance of a particular process component to process decisions — importance in terms of the self-interest of the entity. For example, when cognition infers a new belief that activates intense affect, the concept, event, situation, rule of inference, or deductive path that led to the new belief will have a significance value associated with it. Later cognitive processing of the same concept, etc. will detect the importance of the concept from the previously assigned significance value, which may invoke a memory of the associated affect, or may reinvoke the affect itself. Similarly when conation structures a goal, determines a plan, or executes an action that invokes (or releases) intense affect, the goal or action may have a new significance value associated with it. When later the same goal is established or the same action executed, its importance and significance to the entity will be determined by the previously set value. Significance value mechanisms are not part of cognition, but are an inherent part of the decision structures within the processes themselves.

The most complex and intricate aspect of affect's motivation is the use of significance values in ongoing cognitive and conative processing. After significance values have been associated with previously-referenced process components, the processes use these previous values, and significance values from current affect interrupts, to guide their processing, in order to best reflect the self-interests of the system.

In cognition, such guidance takes the form of determining relevant information (relevant to the entity's self-interests) about a situation to be perceived or a problem to be solved. The two sources of decision criteria for process guidance are; a) the global statement of the cognitive task to be performed, and b) information discovered by local processing as the task is performed. These criteria are used in guiding which inference path to try next, which concept to elaborate, or what strategy to use for a particular task. Specifically, one task of cognition is to perceive a particular situation to obtain whatever facts are most relevant. Significance values from the global statement of the task are consulted initially to determine the facts to focus attention upon, and again whenever the statement of the task is reviewed to refocus attention. Additionally, the association of significance values with perceived objects or concepts provides amplification or attenuation of interest in the current local perception task as a source of locally-obtained information. Another cognitive task is to solve a given problem. Affect provides the significance of the solution as the criteria

for decisions about what inference rules to use or which beliefs to expand upon. As beliefs or rules of inference are used, their importance is determined by their previously set significance values, generating local information for process guidance.

In conation, guidance from affect help6 determine goals and actions, and control the execution of actions. The top-level goal with which affect initiates conation is a product of the currently activated affects and affect's global strategy. This goal is given its significance value (importance of satisfying the goal) by the current affects. When subgoals of this goal are established, they take their significance values from previously assigned values for this particular subgoal, and the current value for the supergoal. The criteria for choosing the appropriate action to satisfy a goal are the value given to the goal, and the values of the possible actions. There exists a set of actions associated with the current goal each having significance values previously assigned. These values are from previous executions of these actions that attempted to satisfy the goal. In determining the appropriate action, cognition may be called to find a new set of actions or examine an already-determined 6et of actions to find the best plan. Cognition returns its candidate plans, and values associated with them. The entire set of values, from cognition, from previous actions, and from current goals, are used to determine the appropriate action. (Note that the significance value is the criteria. Thus a decision may be made in spite of strong contrary "rational" evidence. Or the value from the goal and one previously used action may be so strong as to bypass cognitive planning altogether and instead execute the extreme-valued action.) When the execution of an action is monitored, the results of the actions may invoke new significance values through perception of their effects. These may change the commitment to finishing the action in terms of resources allocated because of a new evaluation of the potential for satisfying the original goal.

The interaction between affect interrupts, assigning significance values upon interrupt, and later use of the values gives the entity the ability to determine importance in making process decisions. This importance derives originally from the fundamental self-interest of the entity, and therefore gives the entity the motivational intricacy it needs for complex tasks

Application: The Paranoid Mode

Our task of modelling a paranoid person in a psychiatric interview situation provides a rich field of examples for the theory. We shall concentrate first on the interview situation and then the paranoid mode.

A typical motivation for a person to participate in an interview is for the reduction of an ongoing condition of distress. This distress could be due to an unsolved problem or unresolved conflict in which: a) the conflict tends to evoke one (or more) of the five negative affects periodically in a person's life, and b) the goals or actions that the person applies are not adequate or appropriate to reduce the negative affect. Periodic discomfort activates the affect distress-anguish. One plan for reducing the affect distress is to eliminate its cause by getting help from someone, thus motivating participation in an interview.

Once in an interview, a person generally has a strategy for interaction to help achieve his goal. We envision the goat to be to get help, with subgoals of setting up the interview, going to the interview, and participating in the interview. A typical strategy for interview participation is: 1) establish rapport with the interviewer by answering questions and talking with him, 2) explain the self's problem, especially the discomfort from certain situations, 3) seek confirmation and support for himself, and 4) get specific instructions for the solution to the problem. As the person embarks upon this strategy, he monitors its progress. For example, before telling about his problem he must perceive whether the interviewer is able to understand him and help him, that is, whether rapport can be established.

When establishing rapport with the interviewer, typical goals are to answer the interviewer's questions, to volunteer demographic information, and to follow the interviewer's direction. To answer the interviewer's question, the person uses a plan to listen to the question, and then try to answer it. This plan calls a motor action to physically listen to the interviewer for his question, store it in a cognitive structure, and then determine what information the interviewer is seeking. If the input from the interviewer is a command, the plan involves determining what action is to be taken that will satisfy the interviewer's request within the context of the goal of trying to establish rapport with him. As the person acts to satisfy this goal, he monitors his actions to see if they are accomplishing his intentions. In this case, he examines the input from the interviewer to see if the interviewer is understanding him and accepting him by performing in the expected role. Note that the motivation for all of these actions, including accepting a

natural language input and interpreting it, stems from the original goal of relieving distress.

Once a person is establishing rapport successfully with the interviewer (at least according to the conclusions of his cognitive processes), he can initiate the next action in his plan for interview participation, that of explaining his problem. The plan calls for relating events and situations to the interviewer from cognition, and listening for and interpreting the interviewer's responses to evaluate whether rapport is being maintained. Throughout the remainder of the interview the interaction continues in a similar manner -- the original plan motivated by affects is executed, resulting in subgoals by conation of listening to and interpreting the interviewer's input, cognitively determining the intervewer's abilities and intentions, explaining the problem to the interviewer and answering his questions, and monitoring the conation process at various levels to insure that goals are being satisfied.

As the participation plan is being executed, other affects modify the plan. Three examples are anger, interest, and fear. When establishing rapport, if the interviewer does not act according to his expected role but instead starts an abnormal sequence of actions (e.g., telling about his own problems), the person's anger may be activated. Such anger may provoke an action of verbally attacking the interviewer, a conative decision that could be made automatically without cognition, *or* refusing to participate further, a conative decision made on the basis of a cognitive judgement that the interviewer doesn't want to help. The strength of the affect response stems from a significance value attached to similar previous experiences in which a person expected a favorable situation to occur and was prevented from experiencing it.

Interest comes into play in an interview as an auxiliary affect to be enhanced. A person participating in an interview receives attention from the interviewer. The importance placed on that attention depends on the person's attitude towards the interviewer, the current need for attention, and the strength of other affects being satisfied. If the interviewer shows interest in some aspect of the person's life, the interest affect may interrupt the current conative process to put more importance on relating well to the interviewer so as to enhance the affect interest. The conative process would resume its former task with a newly acquired tendency to interact positively with the interviewer. This tendency would remain strong until satisfied or until another affect (e.g., fear) interrupts and becomes more important.

Fear in an interview refers to fear of emotional harm, in order to get help, a person must reveal much of himself to the interviewer. As a result, it is easy for the interviewer to lower the person's self-esteem, causing emotional pain. The fear affect is sensitive to emotional pain, particularly if the fear affect has previously placed great significance on the ability to perceive when such pain may occur. If the interviewer shows an intention of proving the person wrong, fear may interrupt and place a high value on telling only facts that defend the person. Later, if it is established that the interviewer only wants to prove the person wrong to help, and if fear has subsided, a high value may be placed *on* telling the rest of the truth.

Thus, the interview task requires numerous decisions in the cognitive and conative processes due to different motivations. These motivations are applied to the decisions by the value mechanism based on affect.

The theory of the paranoid mode of thought provides an example of the direct control of affect over other processes. Within the context of a psychiatric interview and its ongoing goals and actions, a person dominated by the paranoid mode uses his cognitive ability to look for potential humiliation in the input (as one of his goals in an interview). The past experience of the person makes humiliation so painful that not only insults will activate shame, but al6o the slightest evidence of an inadequacy of the self implied by the other person will activate shame. We envision an ever-widening sphere of concepts within the cognitive process acquiring significance values leading to shame over some time span in the life of the paranoid person, until, in severe cases, almost everything in the person's life activates the shame affect and the person breaks down.

Once shame is activated, it becomes extremely disruptive to cognitive and conative processes because of the high significance placed upon the goal of removing shame's activator Typically the activator is a statement of belief or leading question by the interviewer implying an inadequacy of the self. The paranoid person seeks an alternate explanation, such as to posit an alternate belief about the interviewer's incompetence or evil intentions. The incongruence of this new belief with the current processes, and the strong facial and body response of shame lead to disruption of the current processes and initiation of new processes. These new processes seek to extract the person from the uncomfortable situation, either by quitting altogether, partially withdrawing by refusing to answer questions, or attacking the interviewer. The end result of affect's direct control is for a new plan of survival to be implemented based on a belief of danger and controlled by the highly excited affect state.

Implementation

A prototype of the model has been implemented in LISP [Taught, Colby, and Parkison, 1974, also Colby, Parkison and Faught, 1974] and is currently being tested according to a multidimensional analysis. The current implementation is available for interviewing at the Stanford Artificial Intelligence Laboratory through the ARPA net. The prototype contains simulations of the affective, cognitive, and conative processes, but the interaction among the processes is still too primitive. For instance, the value mechanism of motivation by affect is implicit in the data structures used by the processes, and is static over the duration of an interview.

Our intention is to implement conation, cognition, and affect in a production system. A mechanism to handle significance values will be implemented by dynamically ordering the production rules for rule selection. Each set of rules with the same beginning token is ordered when the matcher selects among them. Each rule has a significance value, a vector of the affects and their strengths when previously used. When a set of rules is called, the calling token has its own set of significance values. This global information is used to order the rules. Local information is provided when a rule matches. In this case, the local value modifies the catling value after the rule has been applied, modifying the value of the token that called the rule.

The major thrust of the implementation will be to model the conation process at the lowest possible level, making sure that any action that occurs in the system which is not at the hardware level Is a specific conative process. In this way we will insure that the affect process directs all goal-directed processing towards the entity's self-interests.

Summary

In order to motivate and best guide an entity performing complex tasks, motivation must originate in the self-interest of the entity. Motivation must guide cognitive and conative processes at the lowest possible level by assigning significance values (measures of importance) to various process components and using them in making process decisions. Through such a value mechanism, every decision taken will be to further the entity's self-interests.

Acknowledgements

The author would like to thank Dr. Kenneth Colby and Roger Parkison for their helpful comments and suggestions

This research is supported in part by Grant PHS MH 06645-13 from the National Institute of Mental Health, and in part by the Advanced Research Projects Agency of the Department of Defense under Contract DAHC15-73-C-0435.

The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the National Institute of Mental Health, the Advanced Research Projects Agency or the U. S. Government.

References

Colby, K.M. Artificial Paranoia, Pergamon Press, to be published *January* 1975.

Colby, K.M., Parkison, R.C. and Faught, B. Pattern-Matching Rules for the Recognition of Natural Language Dialog Expressions, Stanford Artificial Intelligence Laboratory Memo AIM-234, June 1974. Also published in the American Journal of Computational Linguistics, Microfiche 5, September 1974.

Faught, B., Colby, K.M., and Parkison, R.C. The Interaction of Inferences, Affects, and Intentions in a Model of Parnoia, Stanford Artificial Intelligence Laboratory Memo AIM-253, December 1974.

Izard, Carroll E The Face of Emotion, Appleton-Century-Crofts, New York, 1971.

Tomktns, Silvan S. Affect, Imagery, and Consciousness, Springer Publishing Company, New York, 1962.