

MODELLING INTENTIONAL BEHAVIOR GENERATION

William S. Faught
The Rand Corporation
Santa Monica, California 90406

Two problems of modelling intentions are identified: representing intentional and non-intentional (mechanical) actions in a single model consistently, and implementing a control structure that accounts for how intentional actions are initiated. These two problems stem from a common source: the interaction of intentional and nonintentional actions. We review a simulation model of cognitive and motivational processes based on a production system architecture that includes both types of actions. We then describe a meta system addition to the simulation model composed of nonintentional actions.

1. MECHANICAL VS. INTENTIONAL ACTION

One example of a model of intentional (and nonintentional) actions is the PARRY3 system [2]. PARRY3 is a simulation model of a paranoid patient in an interview situation. The system (computer program) communicates with a psychiatrist in English via a teletype [3]. The simulation model uses a conventional production system (PS) architecture. The PS has conditions on the left hand sides (LHS) and actions on the right hand sides (RHS) and runs in the normal forward manner: testing conditions and performing actions.

The major elements in the model are action sequences, external conditions, affects, beliefs, inferences, and intentions. An action sequence is a set of linked (with tags) PS rules that represent a sequential set of actions that the system is to perform. Representing an action sequence in sequential rules (as opposed to a single action) allows higher priority actions to interrupt the sequence. An example of an action sequence:

How do you like it there?	Question
Why do you want to know?	Clarification question
Maybe it upsets you.	Clarification
It's ok. But I am upset.	Answer

External conditions (on left hand sides of rules) represent information from the system's interface to the world. Examples are: the interviewer's linguistic input, facts about and measurements of the input, amount of time since the last input.

The model has eight affects representing eight primary emotions (e.g., anger, fear, shame) that identify the significant experiences for the model-patient (simulation), which in turn provide motivation for the model-patient's actions. Affects on a rule's LHS represent affect states that cause behavior. Affect levels on a rule's RHS represent incremental affect modifications performed before the succeeding PS interpretation cycle.

The model also has a number of beliefs (e.g., the interviewer wants to help) and inferences to manipulate them. Beliefs are represented in a standard format: as propositions with modifiable truth values representing the amount of evidence that they are true. Beliefs on a LHS represent conditions; beliefs on a RHS represent actions to add evidence that the beliefs are true. Inferences are PS rules with beliefs on their RHS. An example inference rule is:

```
(INPUT & (INTOPIC = MENTALSTATE)
& (SHAME = HIGH)) => (DOCABNORMAL)
& (FEAR 40)
```

That is, if there is an input, and its topic is the model-patient's mental state, and shame is high, then conclude the interviewer is abnormal and increment the fear affect.

Finally, the model has a number of intentions representing situations that are (believed to be) advantageous to the model-patient (e.g., introducing a new topic). Intentions are similar to beliefs - they have modifiable values and can occur on LHS and RHS of rules.

Values of intentions represent the intention's activation level or "strength of commitment" to achieving the goal. Intentions on a LHS instigate actions. Intentions on a RHS represent actions to (de-)activate the intention (modify its strength value). Additionally, each intention has several special PS rules associated with it that the intention activates and that test whether the intention's desired situation (goal) has come about. If the system has achieved the goal, the rules reduce the intention's strength (so that it is no longer active). An example rule activating an intention is:

(DDHELP & (FEAR LOW)) => PHELP

That is, if the interviewer desires to help, and if fear is low, then set the intention for the model-patient to get help. Subsequent rules use the intention PHELP as a condition to determine what new topic to introduce and when to introduce it (e.g., in a conversational lull).

Generating mechanical actions in this simulation, then, occurs when PS rules with external conditions, affects, and beliefs instigate actions in a stimulus-response manner. Generating intentional actions is consistent with this formulation: intentions (as well as affects, beliefs, and conditions) are LHS conditions which instigate actions. The system transforms an intention from an information structure to a causative agent when the system "activates" the intention by increasing its strength as part of a cognitive decision process. Once the intention is activated, however, the system's cognitive (belief) processes no longer have direct control over the action. The action is subject to the normal conflict resolution algorithm of the PS rule set, and other actions (most likely based on high affect content) may override the intention's action.

2. THE PROBLEM OF CONATION

The second difficulty of modelling intentional action is the problem of conation [1]. The question is: Who performs intentional actions? A mechanical action, such as avoiding a fearful stimulus, can be modelled with a stimulus-response process. The problem is explaining the causal mechanism behind a person's intentional actions. The usual conative paradigm identifies the cause of intentional action B as another process A. However, we have gained nothing if we claim that the person caused process A. If instead we claim that process A is the person's want or desire to do

process B, then it, too, is an intentional action, and we are in danger of an infinite regress. (We will ignore a third strawman explanation, "the Ghost in the Machine" [4], in which the mind is a "ghost" that resides in the body.)

As an alternate approach, consider a (hypothetical) system performing mechanical actions. At some point the system's actions fail and the system stops. We construct an additional process of mechanical actions to locate the error and resume or restart the original mechanical actions. This process would use beliefs and remaining information from the original actions to hypothesize what the failed action was and specify the appropriate recovery actions.

The features of this paradigm are: (1) All actions can be considered mechanical at some level; (2) Mechanical actions can manipulate intentions. In particular, mechanical actions can activate intentions causing intentional actions; (3) The system senses these activated intentions in a manner similar to other (affect and external) conditions. Their activation causes intentional actions.

In this paradigm, an intentional action is the result of a mechanical action that is one level of abstraction higher than the intentional action. The processes involved in the higher level abstraction comprise a meta system with the lower level being an object system.

3. A META SYSTEM

We shall describe our efforts at implementing such a system. (The system is more fully specified in Faught [1978]. It has been incorporated into the PARRY3 system.) The simulation model of cognitive and motivational processes (described above) serves as an "object" system; i.e., the meta system models the object system's actions. The meta system consists of (a) beliefs, conditions, and inferences to model the situation, and (b) actions (rules) to determine facts and manipulate intentions. Meta system constructs are represented in the same format as object level beliefs, inferences, and actions, and are acted on by the same PS interpreter.

The meta level action sequences and inferences are invoked only when certain events disturb the system. The system invokes meta actions whenever a high negative affect condition occurs, when an action failure occurs (detected when no PS rules match for one cycle), or when the interviewer asks the model-patient directly

for an opinion or for information on its internal state (e.g., "What do you want to do right now?"). The meta actions are contained in an action sequence having special conditions on its LHS, e.g., no PS rules matching during the previous PS interpreter cycle. The action sequence carries out the following individual actions: (1) Attempts to determine three beliefs comprising the current situation--what actions the system has been performing, what actions the system is currently performing, and what actions the interviewer is performing. (2) Determines whether the system finds the current situation desirable, i.e., whether the system likes the current situation according to affect measures of the situation (e.g., whether negative affect levels are high). (3) Sets an intention to continue the current situation if it is desirable. (4) If the situation is undesirable, the system determines which states are possible to obtain, chooses the most desirable one, and sets an intention to perform an action leading to the selected state.

As a result of performing this action sequence, the system typically activates an intention to start an action. This intention is part of the normal object-level system; if no higher priority action is being performed, the intention will initiate an action. Note that the meta level routines do not have direct control over the object system's processing; their conative control consists of activating beliefs and intentions, relying on PS rules to detect the beliefs and intentions and initiate the corresponding action.

The meta level processes are not a replica of the object system. The meta processes (beliefs, inferences, actions) contain certain object-level abstractions - e.g., beliefs about what action the system was performing. However, both levels are driven by the same affects and external conditions, and both modify the same affects. The PS interpreter matches rules from both levels during any one cycle. What distinguishes the meta level is its set of abstractions about object-level processing.

Several criteria come to mind when evaluating the meta system's effectiveness: (a) Does the meta system perform a rich set of cognitive operations to model the system's actions? (b) Are the system's actions sometimes independent of the immediate external environment? (c) Can the meta system monitor and restart object-level actions effectively, i.e., with a minimum of object-level actions. While the meta system can restart some intentional actions, the system as a whole fails to exhibit these

qualities. The reason seems to be a lack of complexity, not of details of the external world but of details of its own actions' components. For example, the system has only a limited number of beliefs that represent the actions it is performing at any one time; most of the beliefs are mutually exclusive. There is none of the subtlety required for restarting failures of complex actions. Restarting procedures would include interpreting what an action without an explicit goal was attempting to accomplish, the typical methods of restarting that action, and the consequences of the action's failure. The system needs an extensive set of (situation, potential action, typical outcome) triples and similar triples for meta level processing. The lack of subtle, interlocking meta and object-level data appears to be a major obstacle to eliciting effective meta level processing.

REFERENCES

- [1] Boden, M.A. "The Structure of Intentions." Journal for the Theory of Social Behavior. 3:1 (1973) 23-46.
- [2] Faight, W.S. Motivation and Intensionality in a Computer Simulation Model of Paranoia. Basel: Birkhaeuser Verlag, 1978.
- [3] Faight, W.S., Colby, K.M. and Parkison, R.C. "Inferences, Affects, and Intentions in a Model of Paranoia." Cognitive Psychology, 9 (1977) 153-187.
- [4] Ryle, G. The Concept of Mind. London: Hutchinson, 1949.