

OUTLINES OF A COMPUTER MODEL OF MOTIVATION

G. R. Kiss
M. R. C. Speech and Communication Unit
University of Edinburgh

Abstract

Plans for programming a computer model of human motivation, for use in clinical applications, are described. A system based on parallel cooperating goal-processes is assumed, with the 'will' interpreted as the resource allocator between them. Wishes and wants are distinguished on the basis of resources allocated for action. An interpretation of mental effort in terms of computational work is suggested. The generation of actions through planning, decision-making and forming intentions are discussed in relation to 'regnant' goals. The dual nature of motivation manifested in active and reactive aspects of behaviour are emphasized.

Introduction

The information-processing model of human motivation to be described in this paper has been constructed as a part of a project directed at the investigation of computer techniques in clinical psychology and psychiatry. The project is in an early planning stage. The ultimate aim is to construct computer-based systems capable of intelligent and helpful interaction with people seeking psychiatric help. The general orientation of the project is based on the assumption that such systems will require an adequate model of human psychology at the level usually found in personality theories within psychology. The motivation model put forward here is intended as an element towards the development of such a comprehensive model on the basis of computational (information-processing) concepts.

Our point of view of personality is that of a set of basic resources organised into a heterarchical system. These resources are formed by information structures which include both processes and data structures (objects). The heterarchical organisation implies that these resources are used in complex interaction with each other in order to synthesize high-level psychological functions (like memory, perception, thinking). There can be no absolute separation of these resources from each other and from the higher-level functions. The partitioning of personality into such resources will always depend on the particular point of view adopted for some specific purpose. This is the general problem of representation in modelling complex systems. It must be recognised that there is no unique solution to this problem, but that solutions may be better or worse depending on the purpose for which they are used.

We shall adopt a division of personality into resources, which is reasonably close both to common sense and to traditional views in the psychology of personality. The proposed division has five resources:

1. Central control (consciousness and attention)
2. Knowledge and reasoning (concepts, skills, beliefs, inference, etc)
3. Motivation (goals, values, wishes; wants, will, intentions)
4. Affect (emotions, attitudes, dynamic

evaluation, interrupts)

5. Somatic interfaces (sensory input, motor output, psychosomatic processes)

Our long-term program is to discuss each of these resources in detail; to show that they are themselves complex systems composed of parts; and to indicate how they interact with each other. Implementation as computer programs will be attempted experimentally over the next few years, but the development of a comprehensive framework has higher priority. This paper deals with motivation. The reason for tackling this aspect first is that this area is currently rather neglected in both psychology and artificial intelligence, while it is clearly one of the most central problems in personality theory.

The General Framework

Motivation needs to be treated within a general theory of human action. There is a considerable literature on this problem both in psychology (Mischel, 1969; Cofer & Appley, 1965; Ryan, 1970; Atkinson, 1964; Bindra & Stewart, 1971; James, 1890; to mention but a few), and in philosophy (Anscombe, 1957; Kenny, 1963; Meiland, 1970; Peters, 1958; White, 1968). Our approach is eclectic, and owes much to such sources. At the same time it has some new features, resulting from our attempt to formulate a model which could be implemented in the form of computer programs.

An agent (a person) executes an action in order to bring something about. Something is brought about by making it happen, by causing events. An action is the exercise of the power to cause events (cf. White, 1968, pp. 2-8). Events are changes in a situation. Situations are states of the world (McCarthy, 1968).

A person has an internal representation (a model) of the situation at any given time. This model includes the representations of the environment, of other persons in the situation, and of the person himself. Various mental processes, internal to the person, are also parts of the situation and may be represented in the model. The same holds for the mental processes and objects (information structures) of other persons in the situation.

The fundamental problem of motivation is to characterize the relationship between a person's purposes and actions. The purpose towards which an action is directed we shall call a goal. In general, a goal is a representation of a situation. The person acts in order to bring this situation about, i.e., cause the situation to change through events so that as a result of the action it will agree with the goal.

Not all representations of situations are goals. When a representation is a goal, this indicates that the situation is regarded as desirable by the person, i.e. it has value for him. In the final analysis it is this desirability (value, utility) of goals which motivates behaviour. We shall not in this paper try to analyse how values are attached to goals or why this

happens. As to the question of how, we see scope for both learning and genetic mechanisms. As to the question of why, we adopt a position labelled 'rational hedonism' by G. A. Miller (1962), and assume that values originate from pleasure, but not necessarily sensory pleasure. Pursuing this point further would take us into the area of affect, so we leave further discussion for a future paper. Let us note however, that here we reached one particular interface between the motivational and affective resources.

The goals for action may come from a variety of sources. Psychologists and philosophers tend to make distinctions on the basis of time (past or future reasons), internal or external factors (Peters, 1958), physiological or psychological factors, etc. We shall not be concerned with these distinctions, important though they are, since in our view they do not form part of the motivational mechanisms themselves. Our criterion for any factor to become motivational is that the person should set up a corresponding goal with an assignable value. Events leading to setting up of goals may be found in physiological processes, social circumstances, physical circumstances, psychological processes, and elsewhere.

Wishes and Wants

We shall describe the result of setting up a goal, i.e. the result of attaching a value to a description of a situation, by saying that the person has a wish for that goal. A wish is simply a recognition of the fact that a certain situation is desirable. Having a wish for a goal does not imply action. It is thus possible to wish for goals which are impossible at the time, or ever. When the goal of a wish is reached, the wish is satisfied (fulfilled). By definition, a wish cannot be fulfilled by the result of an action of the person having that wish (although it may be fulfilled as a consequence of an action), since the person does not intend to do anything about his wish. (The distinction between the results and consequences of an action is due to von Wright, 1963.)

The difference between a wish and a want is that the person is intending to do something in order to reach the goal of the want. Note that 'intending to do something' does not mean the same as intending to take a particular action which will bring about the goal. Having a want of a goal simply means that the person is allocating some resources towards taking some action. This process of allocating resources is the function of the will, to which we now turn.

The Will and Mental Effort

The function of the will is to allocate resources to the processes of reaching goals. Will power is the investment of resources. The resources of a person are finite and ongoing goal processes have to compete for them. These resources include both the physical and the psychological, although here we shall be concerned only with the latter.

At the most fundamental level the resources are space (for memory) and time (for processing) in an information processing system. Together these resources can be used to define the amount of computational work required for some process (Savage, 1972).

Many treatments of motivation in the past have made use of an energy analogy to convey the notion that

one important aspect of motivation is the 'energising' of behaviour. Two clear examples are provided by Freudian and ethologist (e.g. Tinbergen, 1951) theories. These theories discuss motivation in terms of the availability, storage, and utilisation of 'psychic energy', sometimes incorporating 'hydraulic' ideas like reservoirs, overflow, releasing valves, etc. It has always been recognised, both by these writers and their critics, that the energy concept is only a loose analogy. The relation of 'psychic energy' to physical energy has never been clearly indicated by Freud, and in his later writings it seems to be treated as an 'inferential abstraction without specific physical referent' (Cofer & Appley, 1964, p. 597). Rapoport (1960), however, has unearthed a rather interesting quote from Freud's 'Three Essays on Sexuality', in which the concept is defined as follows 'the quantum of psychic energy is a measure of the demand made upon the mind for work'. This definition is closely related to the ideas we are putting forward here.

We propose that mental work is to be measured in terms of the computational work performed, i.e. in terms of the number of equivalent logical operations which need to be executed in order to perform a computation (see Savage, [1972] for a detailed quantitative treatment of computational work). The 'psychic energy' concept is to be replaced by the concept of computational work.

Let us now return to the will and mental effort. We have interpreted the will as the allocation of resources to competing processes. Since we are confining ourselves to psychological (mental) processes, the resource is computational work. We now propose that mental effort is the manifestation in consciousness (central control) of the allocation and expenditure of mental work. We need to distinguish at least two aspects. Subjective feeling of effort is involved in completing a process if the amount of computational work involved is large. This is the 'slow dead heave of the will' idea of William James (1890, Vol. II, p. 534). Feeling of effort is also involved in matters relating to the interruption of ongoing activities and the initiation of new activities. This kind of effort is often labelled 'concentration' if it is a matter of protecting a process from interruption, or an act of will if it is a matter of interrupting a pleasurable process in order to do something else. An example of the latter is William James's (1890, Vol. II, p. 524) famous description of getting up on a cold morning.

We also need to distinguish between resource allocation for purely symbolic internal processing without external action, and allocation involving the formulation and execution of intentions and external actions. Special status needs to be given to external actions, because this is what distinguishes daydreaming and fantasy about one's wishes, from doing something about them.

In summary, the will involves resource allocation, partly in the sense of the amount of computational work a process can use, and partly in the sense of the relative priorities of processes over each other as to when they can use up computational work.

It can be seen that in some respects our interpretation of the will parallels the situation in a modern time-sharing computer system where several users compete for the computational resources (CPU, memory, peripherals, etc), and a supervisor allocates quotas so that everyone gets a fair deal.

Let us close this section by pointing out that the will is an interface between the motivational and central control mechanisms and is thus another example of the many points of interaction between the components of the personality system described in an earlier section.

The Control of Multiple Goal Processes

In the previous section we have already introduced a point of view in which each goal is associated with a process directed at the achievement of the goal. We are using the concept of a process in the technical sense in which it has come into use recently in computer science in connection with the design of time-sharing systems (see, for example, Colin, 1971). The several processes belonging to the coexisting set of goals of a person are executed as parallel, cooperating processes. This means that the processes share both control and access environments in the sense of Bobrow & Wegbreit (1972).

The degree to which the goal processes communicate with each other and pass control among themselves is assumed to be much greater than that occurring in time-sharing computer systems, so that in this respect our analogy becomes inaccurate.

The parallel cooperating execution of goal processes implies that the achievement of several goals is being carried forward simultaneously over some time interval, since, although control resides in one particular process at any moment, the mutual interruption, suspension, initiation, and resumption of the processes ensures that some effort is invested in a range of goals at the same time.

This multiple-process view also makes it possible to incorporate another important aspect of human motivation, namely, that any particular action is usually chosen in such a way that it contributes to several goals simultaneously. The need for multiple goal evaluation is recognised and discussed by Fikes, Hart & Nilsson (1972). We propose to make use of the standard AI technique of generating possible actions in the planning stage (see below) on the basis of one goal process (which we shall call the *regnant* goal process, following Murray, 1938), and evaluating it against the current states of the other existing goal processes.

Let us now turn to another aspect of controlling goal-directed activity. Why does activity stop? This depends on the nature of the goal. Some goals can be reached completely, eliminating any associated wish or want. Others can only be reached partially and some evaluation of the state of a goal tree is needed in order to decide whether some associated wish or want is to be wound up or not. This is Simon's (1967) idea of 'satisficing'. Some goals require the indefinite continuation or recurrence of some activity or situation. A goal process may also stop because it ran out of the allocated resources, or because it was interrupted by some other process.

Planning and Intentions

We shall now turn to the processes which may intervene between the setting up of goals, wishes, and wants on the one hand, and the actual execution of actions on the other. These processes are planning, choice (decision-making), and forming intentions.

Once a goal has become regnant (in the sense of assuming control), a plan for action needs to be formulated to bring the goal about. Planning is assumed to be a complex activity, recursively using all the resources of the system. Clearly, it is heavily dependent on the knowledge and reasoning resources in retrieving possible actions and evaluating them in terms of their results and consequences relative to all relevant current goals in the system. These kinds of knowledge are referred to by Sloman (1972a) as the 'resource store, resource catalogue, and a store of environmental information (beliefs)'.

Planning is a process which uses both knowledge of a general kind (knowledge of the world in general), and also knowledge of a current situation. The influence of situational knowledge is underestimated in some current AI programs, since they mostly operate in environments containing only one agent (e.g. a robot). The depth of planning which can be usefully undertaken depends heavily on the predictability of the environment, as can be seen in game-playing programs which have to take into account events produced by other agents. For this reason planning cannot be separated out as an independent, self-contained process, but has to be fitted into the heterarchical organisation as an activity taking place in parallel and in interaction with the other ongoing process.

We shall not go into the details of the planning process itself, since planning and problem-solving are perhaps the best developed areas in AI and we expect that the goal-reduction and state-space approaches would be readily applicable for our purposes.

The process of planning usually uncovers a set of alternative possibilities. These alternatives need to be evaluated with respect to a large range of criteria including, among other things, relevant current goals, preferences, expectancy of success (Atkinson, 1964), likes, principles, etc. (see Sloman, 1972b, for a more detailed discussion [although Sloman regards these criteria as motivating factors, with which we disagree]). The process of taking these criteria into account in an evaluation leads to a choice (decision-making) between the alternatives. The evaluation of alternatives with respect to relevant current goals makes it possible for wishes to have an influence on eventual behaviour even though wishes do not directly generate actions, since they have no resources for this. Presumably this indirect influence is the mechanism behind 'Freudian slips'.

The outcome of decision-making is the adoption of the best plan for execution. The plan itself is a prescription of an action, activity, or of the performance of a complex procedure (actions, activities, and performances are distinguished according to the nature of their results and the manner in which the result is obtained, [see Evans, 1967, for these, and even finer distinctions]). The state of affairs resulting from the adoption of a plan we shall refer to as having an *intention*. A person intends to do something (intends to carry out an action) when he has a plan which he has decided to execute. The amount of detail which may be present in the plan at the time of this decision may vary widely. If the plan is sketchy, the whole of the system may need to be called recursively to make its execution possible.

Active and Reactive Aspects of Motivation

The interaction between a person and the environment has both active and reactive aspects. The active aspect is involved in the execution of actions based on regnant goals. The equally important reactive aspect is concerned with actions based on events in the situation. This distinction relates only to the way in which actions are initiated and not to the way in which their execution is controlled.

The execution of a plan based on a regnant goal takes place in an environment which is only partially under the agent's control. Events may occur in this environment which have not been foreseen in the plan. Such unexpected events call for the initiation of reactions to ensure that the execution of current plans can continue.

Reactions to unexpected events may call for means-ends analysis, planning, and the formulation of intentions in much the same way as active behaviour. However, the distinctive characteristic of reactive behaviour is that it implies a constant monitoring of events through various levels of attention in the perceptual processes, in order to determine their relevance to the currently existing goals. We assume that this is done by setting up perceptual schemas (possibly of quite a crude nature) to watch out for. When such a pattern is detected by low-level sensory and perceptual processes, attention is drawn to the event by causing an interruption of the ongoing activity in central control in order to enable a more detailed evaluation of the event against all relevant goals. Depending on the priorities of these goals, planning and action may be set into motion by central control. The evaluation is aided by the response of the affective system (to be discussed in detail elsewhere) which contributes a quick, simplified evaluation and partly determines whether an interrupt should occur (Simon, 1967).

Since events are evaluated against all relevant goals, we now have a mechanism through which wishes can exert an influence on behaviour, even though the will has not invested resources in them for external action. The reaction to an event which is relevant to a goal can be influenced by the existence of the associated wish. One form of this influence may well be the changing of the wish into a want, if the event indicates that previously existing obstacles to action towards the goal have been removed. This mechanism is, effectively, the 'releasing mechanism' of the ethologists, and the 'incentive stimulation' of Bindra (1969), and others.

We believe that it is extremely important to incorporate both aspects, active and reactive, of the motivational phenomena in our model. Undue emphasis on the 'inner urge' aspect, characteristic of many psychodynamically based theories of motivation, or on the 'environmental pull', both unduly restrict the scope of a theory.

REFERENCES

- ANSCOMBE, G. E. M. *Intention*. Oxford: Blackwell, 1957.
- ATKINSON, J. W. *An Introduction to Motivation*. Princeton: Van Nostrand, 1964.
- BINDRA, D. & STEWART, Jane, (eds). *Motivation*. Harmondsworth: Penguin, 1971.
- BOBROW, D. G. & WEGBREIT, B. A model and stack implementation of multiple environments. BBN Report No. 2334, 1972.
- COFER, C. N. & APPELEY, M. H. *Motivation: Theory and Research*. New York: Wiley, 1965.
- COLIN, A. J. T. *Introduction to Operating Systems*. London: Macdonald, 1971.
- EVANS, O. C. States, activities and performances. *Australasian J. Philosophy*, 45, 293-308, 1967.
- FIKES, R. E. HART, P. E. & NILSSON, N. J. Some new directions in robot problem solving. In: B. Meltzer & D. Michie (eds), *Machine Intelligence 7*. Edinburgh University Press, 1972.
- JAMES, W. *The Principles of Psychology*. (Page refs to Dover Edition, 1950). New York: Holt, 1890.
- KENNY, A. *Action, Motivation and Will*. London: Routledge and Kegan Paul, 1963.
- MCCARTHY, J. Programs with common sense. In: M. Minsky (ed), *Semantic Information Processing*. Cambridge, Mass.: MIT Press, 1968.
- MEILAND, J. W. *The Nature of Intention*. London: Methuen, 1970.
- MILLER, G. A. *Psychology, the Science of Mental Life*. Harmondsworth: Penguin, 1966, original edition 1962.
- MISCHEL, T. (ED), *Human Action*. New York: Academic Press, 1969.
- MURRAY, H. A. *Explorations in Personality*. Oxford: Oxford University Press, 1938.
- PETERS, R. S. *The Concept of Motivation*. London: Routledge and Kegan Paul, 1958.
- RAPOPORT, D. On the psychoanalytic theory of motivation. In: M. R. Jones (ed), *Nebraska Symposium on Motivation*. Lincoln: Univ. of Nebraska Press, 1960.
- RYAN, T. A. *Intentional Behavior*. New York: Ronald, 1970.
- SAVAGE, J. E. Computational work and time on finite machines. *J. ACM*, 19, 660-674, 1972.
- SIMON, H. A. Motivational and emotional controls of cognition. *Psychological Review*, 74, 29-39, 1967.
- SLOMAN, A. Some steps towards the design of a mind: a bird's eye view. Department of Computational Logic, Univ. of Edinburgh, Report No. 59, 1972a.
- SLOMAN, A. How is it possible to use available knowledge and abilities purposefully? Unpublished MS., Univ. of Sussex, 1972b.
- TINBERGEN, N. *The Study of Instinct*. Oxford: Oxford Univ. Press, 1951.
- WHITE, A. R. (ED), *The Philosophy of Action*. Oxford: Oxford Univ. Press, 1968.
- VON WRIGHT, G. H. *Norm and Action*. London, 1963.