3

Action Rules Extracted by Machine Induction from Feature-Coded Self-Reports

David D. Clarke
University of Nottingham

Adam N. Letchford
University of Lancaster

A cross-validated study was performed to assess the ability of a "rule-finder" program to generate rules describing everyday behaviour. Sixty-four participants produced written reports about recent sequences of action in their everyday lives. These were input to the program as groups of numbers representing certain features, such as the time of day, the other people involved, and the reason for the action. Ten rules emerged which could predict the type or features of an unknown action in a new sample of cases with significant accuracy. This way of describing behaviour has a number of advantages. The rules can also be converted fairly simply into a "production system" for use in further modelling work.

Human action is a key concept and object of study in social psychology. Improved understanding of everyday action patterns would be useful to researchers in many different areas. A variety of alternative methods, theories, and tools for the analysis of action have grown up. These include Goffman's (1956) game and dramaturgical metaphors, Harré's (1983) ethnogenic approach, ethnomethodology, statistical sequence analysis, and so on.

One promising approach has been to treat individual actions rather as words or "morphemes" are treated in linguistics (see, for instance, Fromkin & Rodman, 1988). The object is to then write rules, like those of a grammar, to capture patterns over time (Clarke, 1983; Clarke & Crossland, 1985).

Authors' Notes: The authors are grateful to the University of Nottingham Research Fund for financial support, to the Department of Psychology for the use of computers and other facilities, and to Dr. Richard Forsyth of the University of West England for expert advice.

Author info: DAVID C. CLARK, Department of Psychology, University of Nottingham University Park, Nottingham NG7 2RD, England; ddc@psychology. nottingham.ac.uk.

Journal of Social Behavior and Personality, 1998, Vol. 13, No. 1, 33–50.

Such rules are different from the ones described by Argyle (1983) and by Harré (1983). They adopt a "realist" view of what social rules are, taking them really to exist as part of the control mechanism for action. We take the more cautious "instrumentalist" view that action-patterns may be simulated, and so partly explained at least, by suitable sets of generative rules, whether or not they are psychologically real.

For us, a rule is simply an expression of the form "If A is true then B is true." The second part "B is true" is called the target expression. When that describes a particular kind of behaviour, such as eating, we have a behavioural rule, equivalent to saying "When conditions x, y, z are true, the person will eat." This becomes particularly interesting when the conditions x, y, and z involve prior actions, events or circumstances. Then we have a rule which *might* be part of the behaviour process (provided it fits with observed behaviour, of course). Eating may occur *because* people consciously or unconsciously follow the rule "When x, y, and z are the case, then eat."

Such rules are interesting for a second reason, too. Rules of the kind "If *circumstance A* then *action B*" lead to new circumstances, which trigger new rules, new actions, and hence further circumstances. So they form chain reactions, and these can be used to simulate complex sequences of behaviour.

If the lights are off, then turn them on.
If the lights are on, then sit in a chair.
If sitting down, then pick up a book.
If holding a book, then read it.

In AI (Artificial Intelligence), such rules are called productions. Simulations which consist of lists of them are called production systems. Very sophisticated simulations can be constructed from these simple raw materials by including rules for the different circumstances which elicit different responses.

The problem, of course, is to come up with the appropriate rules to make up the simulation. This is often done by trial and error. However, a more direct method now exists for discovering behavioural rules (Clarke & Letchford, 1995). It involves the use of an AI technique known as "machine induction" (Forsyth, 1989), and more specifically the use of a computational "rule-finder." This type of program has been in use in AI for a number of years for performing machine induction for classification and prediction, especially as part of the "Knowledge Elicitation" process. However it has not been tried before for social psychological purposes, using nonspecialist participants, and everyday behaviours.

Knowledge Elicitation developed as part of Expert Systems research. Prior to the late 1970s, attempts had been made to produce a "universal problem solver." This is a computer program which could mimic, or even surpass, intelligent human reasoning, when applied to any practical problem or area of enquiry. Typically, such programs drew heavily on "means-ends analysis," the conversion of the problem into a symbolic, mathematical format (Newell & Simon, 1972). However, it soon became apparent that human problem solving was not the same in different fields. In particular, experts in various subjects used all sorts of ad hoc rules of thumb, informal intuitions and heuristics, based upon years of experience, which were not easily converted into a clear-cut, symbolic format. So, when attention turned to the simulation of expertise, such as computerised medical diagnosis, a whole new set of techniques had to be introduced into what had been a fairly simple discipline (Fox, 1984). These included heuristic software, nonstandard logics, hierarchical goal-structures and methods of eliciting knowledge from experts. This period of expansion paralleled the development of new methods and concerns in social psychology. Indeed, many of the new knowledge elicitation techniques were taken from psychology—repertory grids, card sorts, various scaling methods, and so forth (Gammack, 1987). In recent years though, machine induction has been developed, providing an entirely new approach to expert simulation. It uses a computer program to analyse the judgements of an expert, and then construct its own hypotheses as to how the judgements were made, and how they could be reproduced (Davis, 1991; Forsyth, 1984).

The underlying assumption of machine induction is that the reasoning of an expert can be recovered by systematic and thorough analysis from a sample of decisions and the variables on which they were based. Hence it requires a number of cases, each measured on a number of variables, together with the expert's judgement or decision in each case. This may be the classification of an object (like a geologist's opinion that a rock is "metamorphic"), or the choice of a plan of action (like a doctor's decision to start a patient on chemotherapy). The examples are then analysed by the rule-finder program, to uncover the rules which the expert seems to be following. The rules may then be checked for their plausibility, novelty and usefulness, and also validated on fresh data. In some respects the process resembles a linear discriminant function analysis, although here the results may involve: (a) several rules in combination, (b) hierarchical groups of rules forming a decision tree (Quinlan, 1979), or (c) "nonlinear" rules involving elaborate combinations of variables (Forsyth & Rada, 1986). In this respect most rule-finders are more sophisticated than simple discriminant function analyses. However, the results are no harder to understand or interpret. In fact they are often much easier to use.

This brings us to the present situation, where social psychology needs better techniques for uncovering and formalising the rules behind everyday patterns of action. Machine learning programs are being used to uncover the decision principles of experts in fields like geology and medicine, but not the decisions of everyday life. Social psychologists have not yet become generally aware of these tools. Or, if they have known of them, have probably not considered them relevant to their field. It is clear, however, that a powerful analogy can be drawn between the knowledge engineer, formalising the unconscious knowledge of the expert, and the social psychologist, studying how ordinary people make everyday judgements and decisions.

Rule-finders could be used in social psychology in a number of ways. They could distinguish likely from unlikely actions in a given circumstance; plausible from implausible sequences of behaviour; true from false reports of a given event; or likely from unlikely next actions or utterances given a brief case history. Our particular interest is in the processes which generate action patterns over time. So we decided to look at the "circumstance-action" rules which are implicit in people's descriptions of particular episodes of behaviour. This was an exploratory study, so there were no prior hypotheses to be tested.

METHOD

Participants

Sixty-four participants were chosen to fit a 3-way stratification. There were equal numbers of males and females, in four age groups (16–25, 26–39, 40–55, 56–75), from four geographical quadrants of England. All participants were of British nationality, and all were either lower middle class or upper working class. The majority were undergraduate or postgraduate students, and the rest were local people, living in the vicinity of The University of Nottingham. Everyone who was approached agreed to participate. All were naive with respect to psychology in general and the details of this study in particular.

Materials

Questionnaires. Each participant was given a short questionnaire asking them to report sequences of three actions from their everyday life. These are called "triplets." They also gave what they thought was their reason for the central action in each group of three. This was the "key" event, the one we tried to derive rules for. The events immediately before and after the key event were recorded too, so that the predictive rules could refer to those where appropriate. For instance, travelling might be

TABLE 1 Typical Event Record for the Category "Relaxation"

What you did:	Watch television
Others present/involved:	Husband and son
Time of day:	8 p.m.
Reason:	I was tired
What you did before:	Washed up
Others:	Son
What you did after:	Went to see friend
Others:	The friend

something which typically occurred just after work. Or personal hygiene might be something which typically occurred just before mealtimes.

Sixty-four different questionnaires were used, each asking for four types of key event randomly selected from the eight types below:

- Personal hygiene (bathing, washing hair, brushing teeth, etc.)
- Cleaning/tidying/organising
- Mealtime
- Relaxation (TV, radio, music, light reading, etc.)
- Talking
- Shopping
- Travelling (bus, car, train, cycle, on foot)
- Work (manual, office, heavy housework, etc.)

The categories had been developed in a pilot study, where they were found to provide sufficient coverage of the everyday events in people's lives. A similar grouping was used by Clarke, Allen, and Salinas (1986). For each category, the participant was required to recall one event from the past week, and then to supply the following information:

- the time of day the event occurred
- the reason the participant believed he or she acted in that way
- whether any other people were present or involved
- what the participant had done just previously
- what the participant did subsequently
- whether other people were involved in the preceding and subsequent events

Table 1 shows a typical event record. This is, of course, only one approach that could be taken to the coding of self-reported actions. Numerous behaviour classification schemes have been devised for different research purposes, using categories, features or dimensions (see for example Clarke, 1983, for a review of the approaches and taxonomies).

TABLE 2 Example of Two Predictive Rules Produced by BEAGLE for a Given Target Expression

Target	*[Category = i]*	*Phi*
Rule 1	[var1 × var2] > [var4 + 6.28]	.92
Rule 2	[var8 < var3] & [var7 > .5]	.78

Note: When the rules are true, the target is more likely to be true. The efficiency of each rule is expressed as a Phi-coefficient.

The rule-finder software. The program chosen was BEAGLE—"Bionic Evolutionary Algorithm Generating Logical Expressions" (Forsyth, 1987). Like many other rule-finders, this requires a database of examples and a "target expression." It then produces further expressions which are strongly associated with the truth or falsity of the target expression. An example of a pair of rules, in the raw form produced by the program, is shown in Table 2.

This means the target expression is likely to be true when the other two (predictive) rules are true. That is, a participant is likely to report a behaviour of type "i" (working, say), when variables 1, 2, 3, 4, 7, and 8 have certain combinations of values. Firstly, variable 1 multiplied by variable 2 should be greater than variable 4 plus 6.28. This state of affairs is associated with "type i" behaviour strongly enough to give a Phi coefficient of .92. Secondly, variable 8 should be less than variable 3, and variable 7 should be greater than .5. This state of affairs is associated with "type i" behaviour strongly enough to give a Phi coefficient of .78.

Rules are illustrated in raw form in Tables 2 and 5 to give the reader a feel for the kind of output the program produces, and the amount of "hand translation" that has gone into the final "clear" version of the rules in Table 7. There is no need to work through Tables 2 and 5 trying to decode each rule in turn. For instance, Rule 2 in Table 5

(time ≤ 19.5) or (reason = 1), → PHYS > 1

is the origin of Rule 1 in Table 7.

IF time before 7:30 p.m. OR reason = external (or both), THEN physically active.

The correspondence is clear, even if the former rule seems obscure at first sight.

Unlike typical rule-finders, BEAGLE can form any combination of variables involving logical operators (AND, OR, NOT), comparators (such as GREATER-THAN, LESS-THAN, EQUAL-TO), arithmetic

TABLE 3 Contingency Table Showing the Probability of a Target Expression Being True Given Particular Combinations of Truth-Values for the Rules Identified

Rule 1	*Rule 2*	*p(Category = i)*
F	F	3%
F	T	17%
T	F	35%
T	T	91%

operators (+, -, ×, ÷) and numerical values. This is a powerful format which is not restricted to simple linear functions (as so many statistical procedures are). Also, the program can handle mixtures of interval, ordinal and nominal variables, which is necessary for the present study.

In addition to individual rules, BEAGLE produces tables showing how pairs of rules interact (see Table 3).

This is how BEAGLE works. It belongs to a class of computer programs called genetic algorithms, because they are based on evolutionary principles. Initially, it produces a set of randomly constructed rules. The most efficient of these are selected, and then chopped and spliced together. Occasionally rules are mutated to form new versions which enter the selection process. Then the cycle repeats. This iterative procedure mimics neo-Darwinian evolution. Symbolic expressions (corresponding to chromosomes) are halved and recombined (as in meiosis and mating), mutated occasionally, and then selected according to their "fitness" to provide the basis of future generations. The process halts after a fixed number of cycles, and the surviving rules are harvested, and evaluated statistically.

For this purpose they are put into pairs which use different features of the data, and tried out on a new set of cases. This "test set" is a holdback sample, which has been kept apart from the "training set" in which the rules were first discovered. It provides an extra stringent test of the adequacy of the rules by reducing the danger of "overfitting." That is the description of anomalies in a particular set of cases which would not generalise to others. It is a form of "cross-validation," similar to that used in factor analytic research, for example. For each pair of rules, a "discrimination table" is drawn up. This shows the number of times the target expression was true, cross-tabulated with the number of times both predictive rules were true, only one was true, or both were false, in the test set of cases. The table is then evaluated by the program, using the chi-square statistic.

This process has proved to be effective at searching the enormous "space" of possible rules. It is not only quick and straightforward to carry out, but it also tends to converge rapidly on near optimal sets of rules with most kinds of data. It captures linear and nonlinear relationships in and between the qualitative and quantitative, sequential and static data, in a way that is impossible with most other statistical techniques. It is described in detail in Forsyth (1989). Genetic algorithms in general are discussed in a number of places (cf. Booker, Goldberg, & Holland (1989), Davis (1991), Holland (1975, 1992), and Mitchell (1996)).

Design

The main dependent variable was the category of the key event, which could take one of eight values. The eight independent variables were:

- the time of day at which the event occurred
- the categories of the preceding and following events
- four additional "features" for all the events in each triplet, which the investigators coded in later (see below), and
- the reason given for the main action.

The information was expressed in a numerical code. In addition, the age and sex of the participant were recorded for each triplet of events.

The data were also analysed in separate runs using other features of the key event as the dependent variable, such as the amount of physical activity involved. All of these analyses are reported below where they led to valid rules.

We expected to find rules predicting the category of the key event in each triplet and some of its other features from the 14 other variables, with significant accuracy. This would show there are systematic patterns of action in these descriptions of daily activity, and would go some way towards capturing what they are.

Procedure

Each participant was allocated a different version of the questionnaire at random, to be completed at home. Questionnaires were distributed personally by the researchers. Participants were given no financial incentives to complete them. Overall, this yielded 256 action triplets. For each triplet, the first and last events were coded using the eight categories described above. For example, an event such as "I had a cup of tea" would be assigned to category four, "relaxing." As a reliability check, the classification was repeated by an independent judge on 48 of the events, with agreement on 34 of them. The probability of this agreement occurring by chance is 10^{-14} on a simple ratio test ($Z = 12.22$).

The researchers then coded each action on four additional features:

TABLE 4 An Example Record from the Action Database

Sex	*Age*	*Time*	*Category*	*Pre-Category*	*Post-Category*	*"Key" Features*	*Pre-Features*	*Post-Features*	*Reason*
1	24	20.5	1	4	1	1 2 1 3	2 2 1 2	1 3 1 3	3

- the amount of physical activity required (on a 3-point scale)
- the amount of psychological activity or "mental effort" required (also on a 3-point scale)
- whether the action would typically be initiated by the participant (two categories), and
- whether it was an individual action, joint action, or one where other people would typically be present but not involved (three categories).

These four features were also checked for reliability. All showed significant agreement. On physical activity Spearman's r_s was .66 ($p < .001$). On psychological activity Spearman's r_s was .45 ($p < .005$). On initiation the Phi-coefficient was .46 ($p < .002$). On participation there were 39 out of 48 agreements ($Z = 7.04$, $p < 10^{-8}$, ratio test).

The reasons given by the participants for each key event were coded as one of four possibilities: (a) due to an external stimulus; (b) psychological state or trait; (c) goal-directed but spontaneous; or (d) goal-directed and "scripted."

This classification was checked for reliability by the same method as the features, and a ratio test on 43 agreements out of 64 yielded a Z value of 7.79 ($p < 10^{-7}$).

The database now contained 256 records, each consisting of 19 numbers showing the presence or absence of the various features. An example is given in Table 4. This was then divided randomly into a "training set," for the rule-induction itself, and a "test set" for cross-validating the rules. The chosen ratio was 55%–45%, or 144 records in the training set, and 112 in the test set. The training set was input to BEAGLE, along with suitable target expressions for predicting the category or features of the main event, such as [category = 3], [psychological activity > 2], or [participation = individual].

The program was set to derive three rules per target expression initially. Each rule was the result of 500 "generations," or cycles of the evolutionary procedure. Forty-two rules were discovered. Twenty five of them were shown to be redundant by another component of BEAGLE. The 17 remaining rules are described below.

TABLE 5 The 17 Nonredundant Rules in Raw Form with their Phi-Coefficients

1	(time > 10.5) & (reason = 4)	→ PSYCH > 2	.30
2	(time ≤ 19.5) or (reason = 1)	→ PHYS > 1	.35
3	(postphys < 3) or (postrait > 1)	→ INIT = 2	.41
4a	(postclass < 3)	→ CLASS = 1	.43
4b	(postinit ≥ preclass)	→ CLASS = 1	.38
5	(precat = 8) & (reason > 2)	→ CLASS = 2	.31
6a	(postclass > postinit)	→ CLASS = 3	.50
6b	(prephys > 1) & (preclass = 3)	→ CLASS = 3	.46
7a	(time < 8)	→ CAT = 1	.39
7b	(preinit < 2) & (postcat < 3) & (precat = 4)	→ CAT = 1	.55
8	(time = 14.5)	→ CAT = 2	.24
9	(precat = 1) & (time < 10)	→ CAT = 3	.50
10	(time > 14) & (psyc < 3) & (reason = 2)	→ CAT = 4	.60
11	(time ≥ 10.75) & (preclass > 1) or (postphys < 2)	→ CAT = 5	.27
12	(prephys > 2) & (precat > 4) & (postphys < 2)	→ CAT = 6	.39
13	(reason > 2) & (preclass = 3) & (precat < 3)	→ CAT = 7	.45
14	(time ≥ 12) & (reason = 4) & (postcat = 5, 7, or 8)	→ CAT = 8	.50

RESULTS

Table 5 shows the 17 nonredundant rules in the original form in which they were produced. The majority of rules are already interpretable. Those which proved to be significant will be re-presented later in ordinary English. In that form they are remarkably accessible compared with factor loadings or regression weights, say. Rules dealing with the same dependent variable are grouped together. Rule 8 is clearly defective. It says that the main event is likely to be cleaning/tidying/organising (category two) when the time is 2:30 p.m. As time was a continuous variable, this rule was treated as ill-formed. It also had a very low Phi-coefficient, so it was discarded. (Note that genetic algorithms of this type, like evolution itself, are very powerful but probabilistic and imperfect. They produce different outcomes on different occasions, and occasionally come up with something bizarre.)

The remaining rules were cross-validated as follows. When a target expression yielded only one rule, a table was set up showing how well it performed on the unseen data in the test set. Rule 1, for example, gave the following table (which shows the number of true positives, false positives, and so on).

A chi-square test then showed whether the performance of the rule was significantly better than expected by chance. In this case chi-square

TABLE 6 Decision Table for Rule 1

	Actual Category	
Predicted Category	*Target Expression: True*	*Target Expression: False*
Target Expression: True	5	8
Target Expression: False	25	62

was .51 (df = 1, $p > .1$, one-tailed). So this rule failed on cross-validation, and was rejected.

When two rules were found for the same target expression, BEAGLE produced a more complicated contingency table, like the example in Table 3. This showed how likely the target expression was to be true for every combination of truth values of the predictive rules. A point-biserial correlation was then calculated for that rule pair using all the cases in the test set. For each case, the correlation paired the probability predicted by BEAGLE for those circumstances (0–1) with the actual value reported by the participant (true or false). This is the most precise way to evaluate the rules, when it is possible. However, it requires a minimum of two predictive rules for the target expression in question.

With single rules and with rule pairs, there is more to the evaluation than just saying how likely a rule is to be true, as a percentage, say. A predictive rule which is usually true is a bad rule, with too many false positives. What is needed is a predictive rule which is true IF AND ONLY IF the target expression is true. It is the strength of association between the predictive rule and the target expression that matters. Even the evaluation of single rules has to take account of the number of times:

(a) the target expression is true when the predictive expression is true
(b) the target expression is true when the predictive expression is false
(c) the target expression is false when the predictive expression is true
(d) the target expression is false when the predictive expression is false

With rule pairs there are eight situations any measure has to summarise. These are the times:

(a) the target expression is true when both predictive expressions are true
(b) the target expression is true when the first predictive expression is true and the second is false
(c) the target expression is true when the first predictive expression is false and the second is true, and so on

TABLE 7 The Significant Rules

1. IF time before 7:30 p.m. OR reason = external (or both)	THEN physically active*
2. IF next-action not joint 3. IF next-action initiation high AND previous-action not joint	THEN individual**
4. IF previous-action work AND reason = goal-directed	THEN others present*
5. IF next-action not individual and next-action initiation low 6. IF previous-action physically active AND previous-action joint	THEN joint**
7. IF time before 8 am 8. IF previous-action not initiated AND previous action relaxation AND next-action personal hygiene or cleaning/tidying/organising	THEN personal hygiene***
9. IF time after 2 p.m. AND previous-action not psychologically active AND reason = state/trait	THEN relaxation**
10. IF time after midday AND reason = script AND next-action talking, travelling, or work	THEN work*

$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .005$.

After all the rules had been tested appropriately, ten were found to make predictions about the test set of cases that were significantly better than chance. These are shown in Table 7 as English expressions.

DISCUSSION

On the whole BEAGLE did well at producing reliable rules for each target expression. Although it failed to yield good rules for several of the event categories (and also for two of the features), this only tended to occur for the more open-ended ones. Meals can occur at various times and in various contexts depending on the occupation of the participant. Similarly "travelling" ranged from a short cycle ride to a holiday abroad. The second category in particular (cleaning/tidying/organising) is what Sinclair and Coulthard (1975) call a "ragbag" category. Clearly then the program was not at fault in failing to find rules for these categories.

Even the nonsignificant rules gave positive chi-square values, showing they had picked out more positive instances in the test set than expected by chance (except in one case). Several rules only just reached significance at the five percent level, but this reflects the sample size, rather than a lack of sensitivity in the technique. In future studies it would be helpful to focus on a narrowly defined and fairly uniform sample of people and activities. This would reduce variability and give rules which were less bland as a result. Nevertheless, the broad categories used here are adequate as a general introduction and illustration. For more specific applications of this approach in the field of road accident research, for example, see Clarke, Forsyth, and Wright (in press).

The rules that emerged were quite transparent on the whole. For example, the last rule shows that the key event "work" was often found after midday, as part of a standardised routine (such as a script), and was followed by either talking, travelling or more work. It is a characteristic of BEAGLE rules that they are readily interpretable, even by a lay person, as long as the variables are clearly defined. This is specially the case when they are paraphrased in natural language terms, as in Table 7. This is in contrast to methods like factor analysis or discriminant function analysis, where only the gist of the findings is comprehensible to a lay person. In addition, the program provides the contingency tables described above, which give the probability of new cases belonging to each of the possible types. These too are easy to understand and use. When assessed against unseen data, they provide a cross-validation of the findings, and a strong criterion for the acceptance or rejection of particular rules.

Our aim is to demonstrate the potential uses of a method which is new to this field rather than to provide substantive empirical results, but certain features of the rules are still worth commenting on.

The first part of Rule 1 in Table 7 is straightforward, in that one would expect people to be less physically active in the evening. It is less clear though why the participants attributed physical activity to external factors. It could just mean that the people in this sample tend to remain inactive unless prompted by an external stimulus.

Rules 2 to 6 are fairly straightforward, showing that individual (or joint) activities tend both to follow and to precede other similar activities. This could mean simply that people tend to be alone, or with others, for extended periods of time rather than just for isolated events.

Rule 7 is even simpler. It shows that participants tended to cite events from the start of the day as examples of personal hygiene. This is probably because most people have some kind of routine in the morning, making it the most memorable and consistent source of examples, not because morning is the only time when the activity occurs.

A similar effect is also apparent in Rule 8. Although it seems to provide a rather different kind of information at first sight, it says the preceding action has to involve relaxation and lack of initiation—two features which are consistent with sleep.

Rule 9 shows that the participants believe relaxation is governed by a psychological state or trait. They would often put "because I was tired," as their explanation, although some preferences were mentioned too, such as "I like to have a break."

Lastly, Rule 10 shows that work was often accounted for by some kind of standardised goal.

Despite the diversity of participants by age and gender, significant rules emerged relating actions to circumstances and other events. This implies these rules were consistent across participants. However *no* rules emerged which used age or gender information as predictors. This is in keeping with the largely "situationist" orientation of action analytic research in general. The organisation of the action stream itself often seems to be crucial in shaping choices. It tends to dominate any individual characteristics as a predictor of specific events.

Rules predicting the key event use features of the subsequent event as often as those of the preceding event. This is interesting because it is uncommon to find statistical models which use features of future events as "predictors" for the present, although a great deal of research has been done over the years on the psychology of expectations. Normally it is sound practice to exclude such "future" variables from a list of predictors. However, since they may represent goals and expectations, it is reasonable to include them in this kind of study of behavioural rules. After all, a present action may be chosen because of what it leads to, as much as for what it follows. This argument should not be over used though. The same patterns will occur in the data whether the earlier event was carried out in order to produce the later one, or simply caused it.

A limitation of the study as it stands is the low reliability of some of the coding categories, which may have affected the results in two ways. Inconsistencies in the work of the original coder may have reduced the number of rules which cross-validate successfully. Consistent idiosyncrasies may have produced patterns which do cross-validate, but rely on a particular interpretation of the coding scheme. Further studies of this kind should use more reliable coding procedures. However, some rules, such as Rule 4a in Table 7, are formed solely from variables with high reliability scores. A study based on self-report data should perhaps not be taken as final and conclusive in any case. The rules that emerge should really only be treated as detailed and novel hypotheses for examination in other kinds of empirical work.

For serious applications, beyond the scale of this initial exploration of rule-finding with everyday behaviour, it is usually possible to devote more resources to the creation of appropriate coding schemes. Clarke, Forsyth, and Wright (1995) conducted a large-scale study of the sequences of events leading up to road accidents, using police accident files and rule-finder programs. For this they created a special purpose machine-readable coding language called TRAAL (The Transport Related Action Analysis Language) with its own coding manual, to avoid such problems. They also used the rule-finder itself to assess coder reliability. When the program was able to identify which coder was which from the cases they had worked on, the coding was held to be unreliable. It was then revised until the coders were indistinguishable. This unusual approach to coder reliability has the added advantage that it produces descriptions of the distinctive style of each coder, not just a figure showing the extent of their disagreement. This makes it easier and quicker to resolve any discrepancies (Clarke et al., in press).

Overall, it seems the rule-finder approach offers a potentially useful method for uncovering regularities in the stream of actions. It can supplement the more orthodox statistical techniques already in use, and in some respects may provide a better alternative.

Firstly, the rule-finder is a comparatively powerful technique for data analysis. It performs as well as more orthodox statistical procedures when relations in the data are comparatively linear. It outperforms many of them when faced with nonlinear patterns, disjunctive combinations, and so on. (Disjunctive rules are those involving one condition *or* another, like Rules 1, 8, and 10 in Table 7. These are particularly troublesome for other techniques when the two conditions can be met at once, but the prediction only applies when one condition or the other is true, but not both.)

During extensive pilot work, we used BEAGLE in various ways. These included predicting attributions of responsibility in road-traffic accidents (from situational variables); ethical judgements (from demographic variables); and the degree of intimacy a person has with an acquaintance (from content-analyses of written descriptions) (Clarke & Letchford, 1995). Two PhD theses on the incidence and causes of drug abuse in Kuwait (prior to the Gulf War) used BEAGLE, too. The program successfully predicted the probability of abuse from personality inventory scores (Al-Azmi, 1991), and from sociodemographic information (Al-Najjar, 1992). In all these cases, the program was compared with conventional statistical methods (t tests, ANOVAs, linear regression, and discriminant functions), and found to be equally sensitive overall. Furthermore, it was generally easier to use, and performed better when the regularities in the data were nonlinear.

The second advantage of using a rule-finder is that the rules can easily be converted to such a clear format that they can be presented directly to the participant(s) from whose actions they are derived. This could be an aid to lay people in their self-understanding, perhaps in a clinical context.

Thirdly, the rule-sets can provide a straightforward basis for creating computational "production systems." These are simulations based on rules in the form "If *circumstance A* then *action B*." Fuzzy logic can then be used to generate predictions, as the present system also provides probability values for its rules.

Lastly, rule-finders provide a useful alternative to the way in which scientific psychology normally progresses, with researchers forming various hypotheses and then testing them experimentally and statistically. Instead, these programs can help fill theoretical gaps by proposing hypotheses to be tested. They act as heuristic devices for suggesting novel features and patterns for further study by other means.

REFERENCES

Al-Azmi, M. (1991). *The effect of personality, drugs, and alcohol abuse on criminal behaviour in Kuwait*. Unpublished doctoral dissertation. University of Nottingham, England.

Al-Najjar, M. (1992). *Drug-abuse in Kuwait: The influence of personality, peers and sociodemographic factors*. Unpublished doctoral dissertation. University of Nottingham, England.

Argyle, M. (1983). *The psychology of interpersonal behaviour* (4th ed.). Harmondsworth, England: Penguin.

Booker, L.B., Goldberg, D.E., & Holland, J.H. (1989). Classifier systems and genetic algorithms. *Artificial Intelligence, 40*, 235–2821.

Clarke, D.D. (1983). Language and action: A structural model of behaviour. *International Series in Experimental Social Psychology, Volume 7*. Oxford: Pergamon.

Clarke, D.D., Allen, C.M.B., & Salinas, M. (1986). Conjoint time-budgeting: Investigating behavioural accommodation in marriage. *Journal of Social and Personal Relations, 3*, 53–69.

Clarke, D.D., & Crossland, J. (1985). *Action systems: An introduction to the analysis of complex behaviour*. London: Methuen.

Clarke, D.D., Forsyth, R.S., & Wright, R.L. (1995). The analysis of pre-accident sequences. *Contractor Reports, 305*, 1–63. Crowthorne, Berkshire, England: Department of Transport/Transport Research Laboratory.

Clarke, D.D., Forsyth, R.S., & Wright, R.L. (in press). Behavioural factors in accidents at road junctions: The use of a genetic algorithm to extract descriptive rules from police case files. *Accident Analysis and Prevention*.

Clarke, D.D, & Letchford, A.N. (1995). Rules from behaviour: The use of a computational rule-finder as a social psychological tool. *British Psychological Society, Social Psychology Section Newsletter, 33*, 4–13.

Davis, L. (1991). *Handbook of genetic algorithms*. New York: Van Nostrand Reinhold.

Forsyth, R.S. (Ed.). (1984). *Expert systems: Principles and case studies*. London: Chapman and Hall.

Forsyth, R.S. (1987). *PC/BEAGLE user guide*. Nottingham, England: Pathway Research.

Forsyth, R.S. (Ed.). (1989). *Machine learning: Principles and techniques*. London: Chapman and Hall.

Forsyth, R.S., & Rada, R. (1986). *Machine learning: Applications in expert systems and information retrieval*. Chichester, England: Ellis Horwood.

Fox, J. (Ed.). (1984). *Expert systems: State of the art report*. Oxford, England: Pergamon.

Fromkin, V., & Rodman, R. (1988). *An introduction to language*. New York: Holt, Rinehart, and Winston.

Gammack, J.G. (1987). Different techniques and different aspects of declarative knowledge. In A.L. Krodd (Ed.), *Knowledge Acquisition for Expert Systems: A Practical Handbook*. New York: Plenum.

Goffman, E. (1956). *The presentation of self in everyday life*. Harmondsworth, England: Penguin.

Harré, R. (1983). *Personal being*. Oxford, England: Blackwell.

Holland, J.H. (1975). *Adaptation in natural and artificial systems*. Ann Arbor, MI: University of Michigan Press.

Holland, J.H. (1992, July). Genetic algorithms. *Scientific American, 267*(1), 44–50.

Mitchell, M. (1996). *An introduction to genetic algorithms.* London: Bradford/MIT Press.

Newell, A., & Simon, H.A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice-Hall.

Quinlan, J.R. (1979). Discovering rules from large collections of examples: A case study. In D. Mitchie (Ed.), *Expert systems in the micro electronic age.* Edinburgh, England: Edinburgh University Press.

Sinclair, J.M., & Coulthard, R.M. (1975). *Towards an analysis of discourse.* London: Oxford University Press.