Instance vs. Rule Based Learning in Controlling a Dynamic System

Danilo Fum (fum@units.it)
Department of Psychology, University of Trieste, via s.Anastasio, 12
I-34134 Trieste, Italy

Andrea Stocco (stocco@units.it)
Department of Psychology, University of Trieste, via s.Anastasio, 12
I-34134 Trieste, Italy

Abstract

The question of whether human behavior could be better explained by assuming that people are capable of extracting from their experience some general principles (*rules*) or by supposing that they store in memory concrete, individual exemplars (*instances*) of the situations they deal with was examined in 2 experiments, adopting the Sugar Factory dynamic system control task, that contrasted the predictions of the computational model by Dienes and Fahey (1995) with those deriving from the ACT-R based model developed by Dieter Wallach and coworkers (Lebiere, Wallach, & Taatgen, 1998; Taatgen & Wallach, 2002). The first experiment produced findings that could not be explained by the Dienes & Fahey's model while being consistent with the model of Wallach. The second experiment, however, obtained results that were at odds with the predictions of the latter. A new model is presented that is able to account for the results of both experiments by assuming that participants improve their performance in the Sugar Factory task by choosing, among a pool of very simple solution strategies, those that are judged increasingly more promising by the ACT-R procedural learning mechanism.

Introduction

One of the dichotomies cognitive scientists struggle with concerns the role played by abstract vs. specific knowledge in a variety of tasks ranging from classification (Nosofsky, Palmeri & McKinley, 1994) to the development of automatic actions (Logan, 1988). The question is whether the behavior in these tasks could be better explained by assuming that people are capable of extracting from their experience some general principles (*rules*), or by supposing that they store in memory concrete, individual exemplars (*instances*) of the situations they deal with. While nobody denies humans the capacity to abstract the regularities inherent in richly structured domains, the debate concentrates on whether cognition could be realized as abstract rules or, on the contrary, rule-like behavior could originate from the way specific instances are stored and retrieved from memory.

In the paper we examine the instances vs. rules issue in the context of a dynamic system control task. The following section introduces Sugar Factory (SF), the experimental paradigm used in this work. We present then two computational models, developed within the instance-based framework, that have been proposed to explain the behavior of participants in the SF task. The first model, developed by Dienes & Fahey (1995) (henceforth the D&F model) is based on the instance theory of Logan (1988). The second (henceforth the W model) has been developed by Dieter Wallach and coworkers (Lebiere, Wallach, & Taatgen, 1998; Taatgen & Wallach, 2002) within the ACT-R (Anderson & Lebiere, 1998) cognitive architecture. While based on different assumptions, the models are similarly successful in explaining experimental data, and seem therefore comparable only on other, more lenient criteria, such as parsimony of assumptions, elegance of formulation, or generality of approach. In the paper we show, however, that it is possible to draw different predictions from the models, and we present two experiments carried out to test them.

The first experiment produced findings that could not be explained by the D&F model while being consistent with the W model. The second experiment, however, obtained results that were at odds with the predictions of the W model. We therefore present a new model, also based on the ACT-R cognitive architecture, that is able to account for the results of both experiments by assuming that participants improve their performance in the SF task by choosing, among a pool of very simple solution strategies, those that are judged increasingly more promising by the ACT-R procedural learning mechanism.

The Sugar Factory Task

In the SF task (Berry & Broadbent, 1984) people are requested to keep the production P of a simulated sugar factory to a given target value by allocating an appropriate number of workers W to the job. Both P and W range between 1 and 12, and can assume only discrete (i.e., integer) values. The simulation is based on discrete events, too. At every simulation step, participants provide a number representing the size of the workforce, and the system computes the quantity of sugar that has been produced. Over a series of trials, they attempt to achieve the target by repeatedly specifying a new input and observing the resulting production. Unbeknown to participants, the system dynamics is controlled by the relation:

$$P_t = 2W_t - P_{t-1} + \varepsilon$$

i.e., the production at the simulation cycle t is computed by doubling the number of workers given in input and subtracting from it the production of the previous cycle. Because the resulting production P_t depends, in addition to the input W_t, on the previous production P_{t-1}, participants must learn to condition any new action to the result of the previous simulation cycle. Keeping the number of workers constant is not sufficient to reach steady-state production in SF; on the contrary, constant input makes the system oscillate.

The task of controlling the sugar production is made difficult by the existence of random noise ε, uniformly distributed with values {-1, 0, +1}. The noise hinders a complete control of the system. For this reason, participants are considered to be successful in reaching the target whenever the production falls within an interval of ±1 from it.

For a more realistic interpretation, W is multiplied by 100 (hundreds of workers), and P by 1,000 (tons of sugar). Resulting values of P less than 1,000 are simply set to 1,000, and values exceeding 12,000 are set to 12,000. It is important to note that in almost all the experiments that have adopted the SF paradigm participants are given the goal to produce the target value of 9,000 tons of sugar.

A typical phenomenon that has been obtained in studies using SF is the dissociation between task performance and verbalizable knowledge. While participants progressively improve their capacity to control the system, they remain unable to describe verbally how the system works or how to reach the target value (Berry and Broadbent, 1984). There is evidence that participants do not have a working model of how the system behaves; in fact they remain unable to predict what effect a given change in the number of workers will have on sugar production (Buchner, Funke, & Berry, 1995). Moreover, they are unable to generate useful heuristics for the task (McGeorge & Burton, 1989). Initially taken as a case for the existence of a separate implicit learning system, these results are now generally explained by assuming that people rely on memorized records, or instances, of their interactions with the system. It is within this instance-based framework that the D&F and the W models have been developed.

Dienes & Fahey's model

The D&F model is loosely based on the instance theory of Logan (1988). According to the theory, each time a stimulus is processed, a separate trace is stored in memory. Distinct memory traces accumulate with experience, and permit rapid retrieval of relevant information upon encountering the appropriate stimulus. In absence of practice, responding to a stimulus requires the usage of general solution strategies. After prolonged practice, the correct response is stored in memory, and can be accessed very quickly.

While the instance theory states that every encounter with a stimulus is stored, the D&F model makes the critical assumption that only successful instances are memorized; interaction episodes not resulting in a loosely (i.e., off at most by one) correct result are not taken into account. Whenever, starting from situation $<W_{t-1}, P_{t-1}>$, an action W_t leads to a production P_t that is loosely correct, both the action and the situation are stored in memory. More particularly, two records (instances) are created: (a) the first storing the association between the current production and the action that led to it: $<P_{t-1},W_t>$, (b) the second memorizing the link between the previous workforce and the action: $<W_{t-1}, W_t>$.

On any given trial, a random selection between the instances that match the current situation is performed, and the associated action is executed. Let us suppose that W_{t-1} = 600 and P_{t-1} = 8000. Among all the instances matching the patterns $<600, W_t>$ and $<8000, W_t>$ one is randomly picked out, and the W_t associated with the selected instance is chosen as the workforce for the trial.

Dienes and Fahey (1995) noted that 86% of the first ten input values produced by their participants could be explained by assuming the following strategy: First, if P_{t-1} is above/below the target, then set W_t to a value that is different from the previous one by {0, ±100, ±200}. Second, if P_{t-1} is on the target, then set W_t to a value that is different from the previous one by {-100, 0, +100}. Finally, for the very first trial, choose a W in the range [700..900].

To replicate this behavior, Dienes and Fahey had to stuff into the model a number n of instances covering each of the three cases. At the beginning, the model tries to control the system by following the described strategy; however, as specific episodes are stored in memory, the answers are increasingly controlled by previous experience. The parameter n is critical to have the model produce a good fit. For high parameter values (n > 15), the explicit strategy was applied on virtually all trials, and the model's learning resulted smaller than the learning achieved by participants. For very low values (n < 5), the consistency of responses (for situations that were previously successful) shown by the model was considerably higher than that of the human participants.

Wallach's model

The W model exploits the computational mechanisms provided by the ACT-R (Anderson & Lebiere, 1998) cognitive architecture. ACT-R distinguishes between declarative and procedural knowledge. Declarative knowledge is represented through *chunks*, frame-like structures composed of labeled slots with associated filler values. Chunks are used in the W model to encode the interaction episodes experienced by participants. While the D&F model stores only successfully instances, i.e. instances in which an action (loosely) led to target, the W model stores every interaction episode, irrespective of the result. Moreover, while the D&F model allows the same situation to be coded as multiple identical instances, the W model, following the general policy of ACT-R, does not duplicate identical chunks.

Procedural knowledge is represented in ACT-R through production rules, or *productions*. The condition of a rule specifies some chunks that must be present and active in declarative memory for the production to apply; the action specifies some actions be taken. Productions are used to retrieve and transform declarative knowledge.

The participants' performance is explained by assuming a match between the present situation and the encoding of past instances. On each trial, a memory search is initiated based on the current situation and the target production value in order to retrieve an instance that, in a similar situation, led to success. In case a match is found, the number of workers retrieved from the instance is used as the new input value.

ACT-R allows a partial match between the pattern that drives memory search and the retrieved chunk. Instances that only partially match the retrieval pattern are penalized by lowering their activation proportionally to the degree of mismatch. In case of partial match, the penalty is computed according to the formula:

$$penalty = MP\sum_{s}(1 - sim(required_{s} - actual_{s}))$$

where: *MP* is a mismatch penalty parameter, and *s* represents each slot in the matched chunk.

To calculate the similarity between two numbers *a* and *b* representing the sugar production in the retrieval pattern and in the retrieved instance chunk, respectively, the following function (Lebiere, 1998) is used:

$$sim(a,b) = \frac{min(a,b)}{max(a,b,1)}$$

The model exploits further procedural knowledge for implementing the solution strategy described in the preceding section. Similarly to the D&F model, also the behavior of the W model increasingly changes from an algorithmic, strategy-based approach to memory-based processing as more and more instances of previous interactions are available to the system.

Discriminating between the models

The two models are similarly successful in their fit with the data reported by Dienes & Fahey (1995). In synthesizing the results of the comparison, Lebiere et al (1998) wrote "[The W model] slightly overpredicts the performance found in the first phase, while the D&F model slightly underpredicts the performance of the subjects in the second phase. Since both models seem to explain the data equally well, we cannot favour one over the other" (p.186). Apparently, an identification problem has arisen due to the fact that the empirical data are not sufficient to discriminate between the proposed models. In such a case other criteria (e.g., parsimony, generality, or even elegance) are generally proposed for the evaluation. From this point of view, the W model, because it requires fewer ad hoc assumptions and it relies on a general unified theory of cognition, seems to be preferred.

In fact, there is a way to discriminate between the predictions of the models. As it has been previously mentioned, almost all the experiments carried out within the SF paradigm adopted the same value as the production target: 9,000 tons of sugar. It is interesting to wonder what would happen if such a value is varied by assuming, for instance, that the participants should reach and maintain a production of 3,000 tons.

According to the D&F model, no difference between the two conditions should be obtained. The relation that controls the system dynamics makes in fact every production level comprised between 2,000 and 11,000 tons equally probable. Any particular target value is used by the model as a filter to establish whether an instance should be taken or not into account. Only successful instances, i.e. instances that led to a production identical (plus/minus noise) to the target are memorized, but the target value by itself does not appears in the stored instances nor does it play any functional role in determining which instance is used to control the system behavior.

Things are, however, different in the case of the W model due to the ACT-R-based partial matching mechanism it relies upon. Among the chunks that match the retrieval pattern, ACT-R chooses the most active one. When the match is only partial, however, the activation value of the chunk is lowered proportionally to the degree of mismatch computed according to Lebiere's similarity function.

The function returns different results for the same absolute difference between the retrieved and the expected values, depending on the values that are taken into account. For instance, a mismatch of 1,000 tons in the value of the current production P_t is evaluated differently for a target value of 3,000 vs. 9,000 tons. In the first case, a retrieved value of 2,000 tons leads to *sim*(2,3) = 0.67 while in the second case a retrieved value of 8,000 tons leads to *sim*(8,9) = 0.89. Because the activation mismatch penalty is higher for the less similar values (i.e. for the 3,000 tons in comparison with the 9,000 tons condition), the consequence is that more mismatched chunks would be retrieved in the latter case. As a result, the 3,000 tons condition, being characterized by better precision, takes advantage from the retrieval of more successful instances, and is therefore facilitated in comparison of the 9,000 tons condition.

In summary, while the D&F model predicts no difference in trying to control the SF system with the target of maintaining a production of 3,000 vs. 9,000 tons, the W model predicts that the first condition would be easier than the second one. A series of 2,000 runs with the W model confirmed the analysis reported above.

Experiment 1

To test the predictions deriving from the D&F vs. the W model a first experiment has been carried out.

Method

Participants

The participants were 88 undergraduates (33 males and 55 females) aged 19 to 33 (median = 23), enrolled in a General Psychology course. The data of one participant were missed due to a computer failure.

Design and Procedure

The main independent variable manipulated in the experiment was the target value of the sugar production that participants were trying to reach, i.e. 3,000 vs. 9,000 tons. Participants were tested individually in single sessions. Each session was divided in two blocks of 40 trials each. The main dependent variable recorded in the experiment was the number of hits obtained in each phase, i.e. the number of times participants were able to reach the target value according to the loose scoring criterion described above. The participants were not informed about the scoring criterion. The experiment adopted a mixed design with Target value as between subjects and session Phase as within subjects factors.

The participants interacted with the system through an interface that displayed the current values of the production and of the workforce. The number of workers was set initially to 600, and the sugar production to 6,000 tons. The participant set the new workforce by pressing one of the F1-F12 function keys and, after a small interval, the system displayed the resulting value of the production.

Results

The main results of the experiment and the predictions of the W model are reported in Table 1.

Table 1: Experiment 1 results and W model predictions

	First Phase		**Second Phase**	
	Particip.	W model	Particip.	W model
3000-3000	10.57	10.09	13.18	12.06
9000-9000	7.26	8.79	9.26	11.10

The ANOVA computed on the participants' data showed as significant the main effects of the Target ($F(1,85)=17.69$, $MSE=569.71$, $p<.0001$) and of the Phase ($F(1,85)=14.86$, $MSE=232.19$, $p=.0002$), but not their interaction. As predicted by the W model, the participants in the 3,000 condition gave a better performance in comparison with the 9,000 tons condition, while participants of both groups improved their capacity to control the SF system from the first to the second phase. While the W model overestimated somewhat the participants' performance in the 9,000 tons condition, it was however capable of predicting the general qualitative structure of the results, a feat that apparently goes beyond the capabilities of the D&F model.

Experiment 2

Both the D&F model and the W model assume that the performance in the SF task is controlled (with the exception of the very few interaction episodes) by the retrieval of previous instances from memory. An interesting prediction that can be derived from both models is that a change in the target value between the first and the second phase of the experiment should be detrimental for performance. In fact, a change in the target value between the two phases destroys the possibility to take advantage, in trying to control the system in the second phase, from previous experience because all the instances that have been stored refer to interaction episodes that are almost useless in the new condition. As a consequence, no positive transfer between the two phases should be obtained. To test this prediction, a second experiment has been carried out.

Method

Participants

The participants were 88 undergraduates (39 males and 49 females) aged 19 to 32 (median = 25), chosen from the same population utilized in the first experiment.

Design and Procedure

The design and the procedure of the experiment were identical to those adopted in the previous one. Only the production values used in the different Target conditions were varied: One group switched from a target value of 3,000 tons in the first phase to a value of 9,000 in the second. The other group experienced the same target levels but in the reversed order.

Results

Table 2 shows the results and the W model predictions (obtained after 2,000 runs) for the experiment.

Table 2: Experiment 2 results and W model predictions

	First Phase		Second Phase	
	Particip.	W model	Particip.	W model
3000-9000	9.39	9.77	8.70	9.11
9000-3000	6,86	8.67	12.84	8.19

From the ANOVA, a significant ($F(1,86)=25.59$, $MSE=487.78$, $p<.0001$) interaction Target x Phase was obtained together with a significant ($F(1,86)=16.18$, $MSE=308.46$, $p=.0001$) main effect of the Phase. While the participants in the 3,000-9,000 condition gave an essentially similar performance in the two phases, the participants in the 9,000-3,000 condition doubled their performance in the second phase. This result comes completely unexpected and is at odds with the predictions of the W model that completely misses the interaction, and is therefore unable to explain the pattern of behavior found in the experiment.

A New Model

If we take the results of both experiments jointly into account, we are puzzled by a striking finding: independently of the situation experienced in the first session phase, those participants that in the second phase shared the same target value gave essentially similar performances. The mean number of second phase hits in the 3,000 condition was, in fact, 13.18 for the first experiment, and 12.84 for the second. In the 9,000 condition the hits were 9.26 and 8.70 for the first and second experiment, respectively. This finding suggests the idea that some situational factor could be more relevant than the memory of past instances in determining the participants behavior in the SF task. Grounded upon this idea, we developed a new ACT-R model (named the P, or procedural, model) that tries to explain the data obtained in our experiments.

The P model is based on the assumption that participants can exploit a set of very simple strategies in choosing the workforce value for the SF task. By combing the SF literature, and by looking at the participants' interaction traces, we can identify some of these strategies. The model we developed comprises the following ones:

Choose-random. Choose randomly a value between 1 and 12.

Repeat-choice: Repeat the value of *W* chosen in the previous interaction episode.

Stay-on-hit: Whenever the previous *W* choice resulted in a success, repeat it. This strategy can be considered as a more selective variant of the previous one.

Pivot-around-target: Choose for *W* the value of the target, plus/minus noise.

Jump-up/down: If the resulting *P* is lower than the target, increase the value of *W*; if it is higher, diminishes it. There exist several possible variants of this strategy. The one employed in the model was the *Jump-on-Middle,* i.e., choose as the new *W* a quantity that lies midway between the previous value and the upper/lower limit of the distribution (i.e., 1 when decreasing, and 12 when increasing).

All these strategies have been implemented through ACT-R productions that are let compete for execution. ACT-R selects the productions on the basis of their expected utility. The expected utility depends, among other things, from the success probability of the production. Whenever the subsymbolic ACT-R production parameter learning mechanism is activated, successful applications of a production increase its expected utility, and therefore augment the chance that the production will be chosen for execution in a next occasion.

Table 3 reports the results obtained by running the P model for 2,500 runs in each condition of both experiments, and compares them with the actual participants' data.

Table 3: Experimental results and P model predictions

	First Phase		Second Phase	
	Particip.	P model	Particip.	P model
3000-3000	10.57	9.35	13.18	12.30
9000-9000	7.26	7.02	9.26	8.83
3000-9000	9.39	9.32	8.70	9.04
9000-3000	6.86	6.96	12.84	12.12

The P model makes predictions similar to the W model in the first experiment, and is able to capture in the second experiment the interaction originated from the anomalous result obtained in the 9000-3000 condition.

Even more important than a good fit is the capability of the model to provide an insight for the data, and to be able to suggest an explanation for them. In our case, the most puzzling results obtained in the experiments are: (a) the fact that the behavior in the second phase seems to be a function of only the target value sought for in that phase, independently of what the participants have done in the first phase, and (b) the interaction arising from the exceptionally good performance shown by the participants in the 9,000-3,000 condition of the second experiment.

These results suggest that situational factors play an essential role in determining the quality of the performance in the experimental conditions. In other words, some conditions are inherently easier than others, due to the dynamics of the SF system and to the way it deals with out-of-range values.

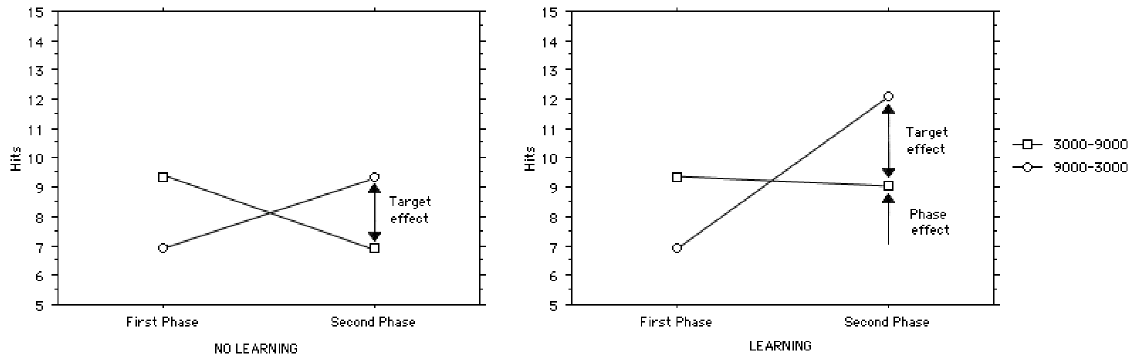

Figure 1: The model predictions without (left) and with (right) the learning mechanism activated.

The success probability of a production is in fact different in different conditions: for instance, the *Repeat-Choice* rule has an averaged (between phases) success probability of 15.18% in the 3,000 tons condition vs. 11.32% in the 9,000 one. Analogously, the *Jump-on-Middle* (up and down) rule has success probabilities of 2.44% vs. 1.22 in the 3,000 and 9,000 tons conditions, respectively. This simple fact is sufficient to explain the Target main effect in the first experiment.

Not only the success probability for the same production is different in different conditions, but also the success probability of *different* production in the *same* condition are unequal. In other words, some productions, in the SF domain, are inherently better than others: so, for example, the productions *Repeat-Choice* or *Stay-on-Hit* are always better than *Choose-Random* or *Jump-on-Middle.*

At the beginning, the model considers all productions as having the same expected utility. When the ACT-R production learning mechanism is activated, however, it increasingly tends to prefer the most successful productions, and this results in the improvement that is generally obtained from the first to the second experimental phase.

The puzzling results obtained in the second experiment derive from the combined influence of situational (Target) and learning (Phase) factors. Figure 1 illustrates the idea. Without learning, only situational factors are involved. In this case the model predicts a symmetrical situation for the two conditions. When learning is activated, it interacts with situational factors originating the results obtained in the second experiment.

It is worth to emphasize that we do not claim that participants are unable to store any memory of their interactions with the system or that instances are useless to explain human cognition. The model simply states that instance-based learning does not play a significant functional role in controlling a simple, but not trivial, dynamic system like SF, at least for the time scale covered by the experiments.

References

Anderson, J.R., & Lebiere, C. (1998). *The atomic components of thought.* Hillsdale, NJ: Erlbaum.

Berry, D.C., & Broadbent, D.E. (1984) On the relationship between task performance and associated verbalizable knowledge. *The Quarterly Hournal of Experimental Psychology, 36A,* 209-231.

Buchner, A., Funke, J., & Berry, D.C. (1995). Negative correlations between control performance and verbalizable knowledge: Indicators for implici learning in process control tasks? *The Quarterly Journal of Experimental Psychology, 48A,* 166-187.

Dienes, Z., & Fahey, R. (1995). Role of specific instances in controlling a dynamic system. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 848-862.

Lebiere, C., (1998). The dynamics of cognition: An ACT-R model of cognitive arithmetic. (Technical Report CMU-CS-98-186) Pittsburgh, PA: School of Computer Science, Computer Science Department, Carnegie Mellon University.

Lebiere, C., Wallach, D., & Taatgen N. A. (1998). Implicit and explicit learning in ACT-R. In F. Ritter & R. Young (Eds.), *Proceedings of the Second European Conference on Cognitive Modelling*. Nottingham, UK: Nottingham University Press, 183-193.

Logan, G. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492-528.

McGeorge, P., & Burton, A.M.. (1989). The effect of concurrent verbalization on performance in a dynamic systems task. *British Journal of Psychology, 80,* 455-465.

Nosofsky, R.M., Palmeri, T.J., & McKinley, S.C. (1994). Rule-plus-exception model of classification learning. *Psychological Review, 101,* 53-79.

Taatgen, N.A., & Wallach, D. (2002). Whether skill acquisition is rule or instance based is determined by the structure of the task. *Cognitive Science Quarterly,* ***2,*** 163-204.