Probing the Emergent Behavior of Tabletop, an Architecture Uniting High-level Perception with Analogy-making

Douglas R. Hofstadter and Robert M. French
Center for Research on Concepts and Cognition • Indiana University
510 North Fess Street • Bloomington, Indiana 47408

Abstract

Tabletop is a computer model of analogy-making that has a nondeterministic parallel architecture. It is based on the premise that analogy-making is a by-product of high-level perception, and it operates in a restricted version of an everyday domain: that of place-settings on a table. The domain's simplicity helps clarify the tight link between perception and analogy-making. In each problem, a table configuration is given; the user, hypothetically seated at the table, points at some object. The program responds by doing "the same thing", as determined from the opposite side of the table. Being nondeterministic, Tabletop acts differently when run repeatedly on any problem. Thus to understand how diverse pressures affect the program, one must compile statistics of many runs on many problems. Tabletop was tested on several families of interrelated problems, and a performance landscape was built up, representing its "likes" and "dislikes". Through qualitative comparisons of this landscape with human preferences, one can assess the psychological realism of Tabletop's "taste".

Cognitive generality in a microdomain

Tabletop, a computer model of analogy-making, is based on the premise that analogy-making and perception are inseparably intertwined. The nondeterministic and parallel program builds representations of situations and makes analogies in a familiar but restricted domain: place-settings on a table. Though simple, the domain brings out general issues in high-level perception and analogy-making.

Imagine Henry and Eliza facing each other across a table. Henry touches an object and says, "Do this!" Eliza must respond by touching some object. The program plays the role of Eliza, with *selection* playing the role of touching. One obvious possibility would be to "touch" *literally the same* object. This option is always open, no matter what the configuration and no matter what Henry touches. Often, though, there are aspects of the situation—pressures—that make the literal-sameness option less appealing than touching some other object. The possibility of any number of pressures coexisting, and their often subtle interactions, lend the domain considerable complexity and depth.

Suppose both individuals have coffeecups before them. Most people would perceive the cups as *counterparts*. Thus, if Henry touches *his* cup, it would seem more natural for Eliza to touch *her* cup than to reach across the table to touch his. But now suppose the "counterparthood" is weakened by changing Eliza's cup to a glass (Fig. 6). Here, although the two objects remain counterparts in terms of *position*, their *categories* no longer match exactly.

However, as "cup" and "glass" are closely related categories, there remain reasons—pressures—for Eliza to see her glass and Henry's cup as counterparts. Chances are good that Eliza will rank touching her glass higher than touching his cup. Of course, if the category mismatch is further increased—give Eliza a fork, not a glass—the sense of counterparthood will be so diminished that Eliza may revert to the literal-sameness option (touching Henry's cup).

Other pressures that might influence Eliza's choice include: the arrangement, category memberships, and orientations of objects, etc. Often, perceptual groupings ("chunks"), whose plausibility depends on the physical and conceptual proximity of the items involved, play a role in determining what items Eliza is prone to see as counterparts.

Obviously, not all possible groupings can be considered by a person or by a program—after all, with just a dozen objects on the table, hundreds of potential ways of grouping them exist; moreover, if smaller chunks are allowed to be members of larger chunks (a key feature of human perception, which routinely builds up such hierarchical representations), the number is even higher. Of course, not only efficiency but cognitive plausibility militates strongly against a computer model in which brute-force strategies of any sort play any role. Thus a crucial design philosophy of the Tabletop program is that it does not routinely invoke all possible pressures in each situation; rather, it lets a limited number of context-dependent pressures *emerge* as each situation is perceptually processed.

Parallel emergent perceptual processes

The central challenge of the Tabletop project is to model the simultaneous existence and interaction of multiple pressures in a human mind perceiving (*i.e.*, building representations of) a complex situation. We stress perception rather than analogy-making, since our philosophy is that *analogies emerge automatically as a by-product of high-level perception* (see [Chalmers, French, & Hofstadter 91]). This idea is at the crux of Tabletop; we contend that the program should be judged not only on the accuracy with which it mimics human performance in its narrow domain, but also on its general principles, intended to apply to any domain, irrespective of size. (In fact, Tabletop's forerunner Copycat uses a similar architecture in a different microdomain. See [Mitchell 90] and [Hofstadter & Mitchell 91].)

In Tabletop, "high-level perception" means the concurrent carrying-out of the following tasks:

- initial labeling of objects in terms of basic categories;
- further labeling, on higher levels of abstraction, of already-labeled table objects;

- hierarchical building-up of *groups* (tentative perceptual chunks) on the basis of:
    - physical proximity of component items (*i.e.*, table objects or already-built groups);
    - conceptual proximity of component items;
    - structural similarity of component subgroups;
- building-up of *correspondences* (links between two items establishing them tentatively as each other's counterparts) on the basis of:
    - corresponding physical positions of the two items;
    - conceptual proximity of the two items;
    - structural similarity of the items (if they are groups);
- assignment of a time-varying *salience* to each perceived item (object, group, or correspondence);
- assignment of a time-varying *strength* to each perceived correspondence;
- competition among rival perceptual structures, giving rise to a pruning of weaker structures.

All these tasks are carried out in parallel. A key idea of the architecture is that each type of task is implemented as a sequence of small, independent *micro-actions*. Building a group, for instance, involves an escalating series of "microtests" that check the physical and conceptual distances between prospective group members. If any such test fails, the potential group is aborted; if it succeeds, the way is clear for further tests; if all requisite hurdles are cleared, the group gets built by a specific micro-action.

To carry out all these tasks requires many micro-actions, which are interleaved *at random*. For instance, a microtest checking out the attractiveness of a potential group on Eliza's side might run, followed at random by a micro-action that proposes attaching an abstract label to some object on Henry's side, followed by another microtest that checks out some other aspect of Eliza's potential group, followed by a micro-action that tests some aspect of a proposed correspondence elsewhere, etc. In sum, many different sorts of small things happen, one after another, at different places on the table. Through such interleaving of scattered local micro-actions, *large-scale perceptual structures gradually emerge in parallel in all areas of the table*. Because all these processes have mutual influences, the perceptual structures that they build tend to form conceptually coherent sets.

Tabletop's parallelism thus lies at the *task* rather than the *micro-action* level. The degree of effective parallelism is determined by the grain of the break-up of tasks into micro-actions. The finer the grain, the more evenly will emerge the different perceptual structures. A *totally unbiased* selection of micro-actions would result in all large-scale tasks getting carried out, on average, at the same speed—a completely "fair" sharing of attention over the table. However, such perceptual fairness is far from Tabletop's strategy; rather, Tabletop *accelerates* avenues of exploration that offer promise while *retarding* ones that appear uninteresting. For instance, Tabletop is not equally likely to inspect all objects on the table; at any given moment, probabilities bias its choice of what to look at. Metaphorically speaking, certain objects and areas of the table are perceptually "hot" while others are "cool", and these biases are dynamic: they change as new perceptions are made.

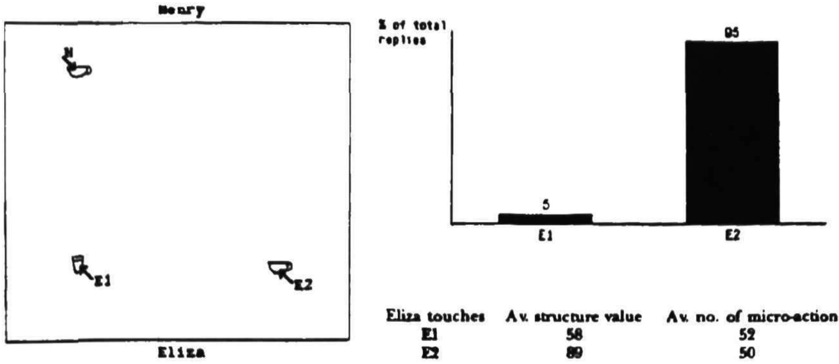
Dynamic biasing is realized by assigning each micro-action an *urgency*—effectively its probability of being chosen. Urgencies are assigned according to the perceived promise of a given micro-action. For instance, a proposed micro-action involving a salient object would get a higher urgency than one involving an object of low salience (all other things being equal). Many factors are taken into account in urgency assignment: location on the table, types of items involved, strengths and/or saliences of objects or correspondences involved, etc. Since micro-actions having high urgencies tend to be chosen swiftly, sequences of logically related high-urgency micro-actions will tend to be accelerated, while sequences of low-urgency micro-actions will tend to be slowed down. In this way, different perceptual structures emerge naturally at different speeds, depending on the program's best *a priori* estimate of their significance.

Since the relative saliences of table items are critical in determining what makes a region of the table "hot", one might ask, "What makes an item salient?" Many factors are involved, including: location relative to the touched object; conceptual proximity to the touched object; being a group (as opposed to a mere object); size of group; physical position in a group; prior perception of other objects in the same category; having a counterpart or not; etc. Obviously, some of these attributes vary over time, so that saliences also change, which implies that various areas of the table become stronger or weaker probabilistic foci of the program's attention. Not only saliences of objects but also strengths of correspondences play a role in determining an area's (probabilistic) perceptual attractiveness, and a similar list of factors is taken into account in the computation of each correspondence's (dynamically varying) strength.

Any group can be disbanded and any correspondence taken down. Often the reason for dismantling a perceptual structure is the discovery of another structure of comparable or greater strength. Thus Tabletop's perceptual process is a rough-and-tumble contest among conflicting interpretations (often just fragmentary), the outcome of which, in the end, is hopefully a strong set of mutually-reinforcing perceptual structures.

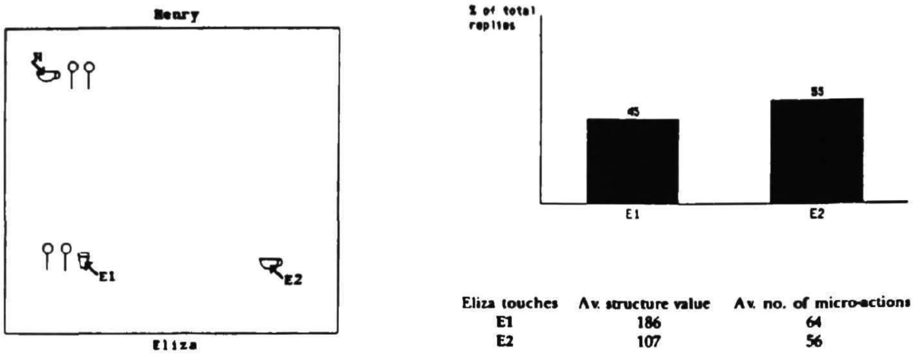
Structure value is a dynamically varying number that represents the total strength of all currently existing perceptual structures. This number can be considered a measure of how well the program has so far done in "making sense" of the scene before it; at the end of a run, structure value can serve as a "quality measure" of the answer produced by the program. An important pressure on Tabletop is to maximize structure value; counterbalancing this, however, is a competing pressure—time pressure—that pushes for the program to finish within a reasonable amount of time.

As perceptual structures emerge around the table, *mappings* also emerge. Indeed, a mapping is just one type of perceptual structure: a family of one or more mutually compatible (often mutually reinforcing) correspondences. A mapping, needless to say, is an analogy. The basic premise of Tabletop, then, is that *analogy-making is a high-level by-product of perception*. In other words, analogies



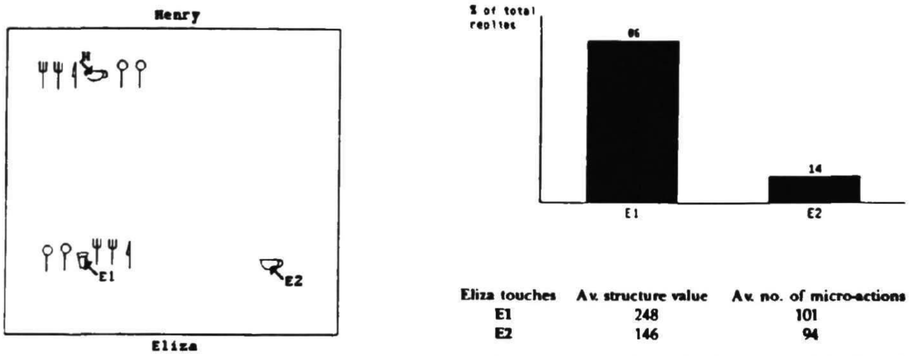
Eliza touches	Av. structure value	Av. no. of micro-actions
E1	58	52
E2	89	50

Figure 1. In Figs. 1 through 5, representing the *Surround* family, H touches his cup. Though the literal-sameness answer (*i.e.*, touching H's cup) is always possible, the main rivalry is between E's glass and cup. The variants explore combinations of pressures by surrounding E's glass and H's cup with various sets of objects. Fig. 1, the "base case", has no surrounding objects; here, just two pressures contribute to the decision: category membership and physical position. The former favors the cup (category *identity* is better than category *proximity*). What about the latter? People are more likely to seek a corner object's counterpart in the diagonally opposite corner than in the mirror-image corner, so such a bias was built into Tabletop. Therefore, position pressure also favors the cup. Overall, then, the pressure in favor of E's cup is very strong; indeed, Tabletop chooses her glass only 5% of the time.



Eliza touches	Av. structure value	Av. no. of micro-actions
E1	186	64
E2	107	56

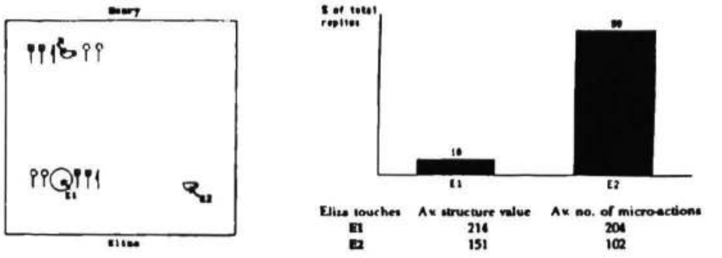
Figure 2. When humans look at this setup, they effortlessly perceive two groups — one containing H's cup, the other containing E's glass. (Of course other groupings are possible, but virtually never come to mind.) Tabletop is similarly inclined; it sees a group consisting of H's cup and two spoons (the spoon pair is likely to be seen as a *subgroup*), and a group consisting of E's glass and two spoons (also likely to be seen as a subgroup). Not only do these groups map onto each other as wholes, but their subgroups (if seen) map strongly onto each other, thus pushing the structure value up and increasing the pressure for mapping H's cup onto E's glass. Indeed, Tabletop now touches the glass 45% of the time. As might be expected, the average structure value when it does so is significantly higher than when it touches her cup. This is a case where highest frequency and best structure disagree.



Eliza touches	Av. structure value	Av. no. of micro-actions
E1	248	101
E2	146	94

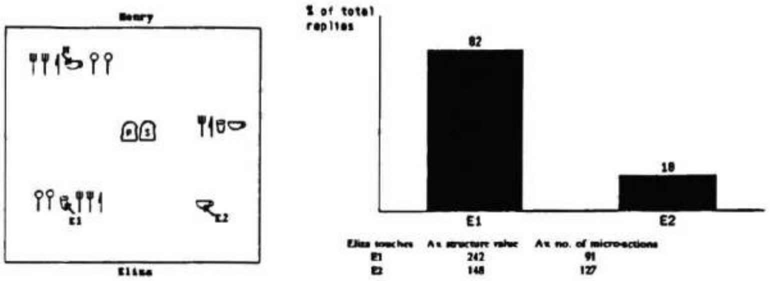
Figure 3. The groups around H's cup and E's glass are very similar: they have the same number of objects and their subgroups are identical. Humans see the mappings as very strong and see E's cup as a loner, thus feel much pressure to touch her glass. Not only does the program do this far more often than touch her cup, but the structure value for the former averages far higher than for the latter.

The strong mappings push so hard for touching E's glass that one might wonder what would *ever* induce Tabletop to pick E's cup. Two factors are involved. One is, Tabletop sometimes simply fails to build those mappings. On such runs, the pressures do not so greatly favor her glass. More rarely, Tabletop may build the mappings but simply choose (stochastically) to ignore them and touch E's cup. Though this may seem irrational, people often act similarly. In a survey, subjects were asked to draw all relevant correspondences in this setup. Some, after drawing a line linking the two spoon-groups, another linking the fork-groups, and a third linking the knives, *ignored* all these lines and chose E's cup. In this light, the "anomalous" 14% of runs in which Tabletop touches the lone cup seem justified.



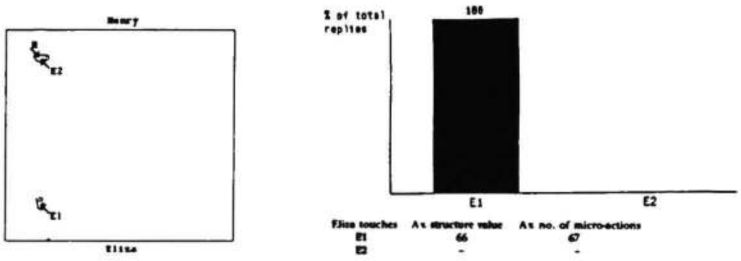
Eliza touches	Av. structure value	Av. no. of micro-actions
E1	214	204
E2	151	102

Figure 4. E's glass has been replaced by a plate, conceptually remote from the touched object. This should shift the pressures back to favoring the isolated cup. Indeed, Tabletop now touches E's cup 90% of the time, and her plate just 10%. Still, the structure value associated with the plate-answer remains over 40% higher than that for E's cup.



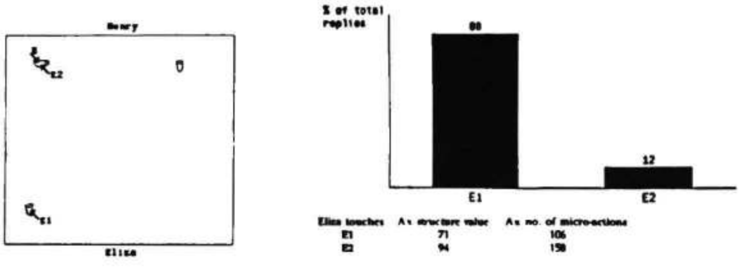
Eliza touches	Av. structure value	Av. no. of micro-actions
E1	242	91
E2	148	127

Figure 5. Distractors have been added to Fig. 3, more than doubling the number of potential object-correspondences. But if Tabletop's focusing mechanisms operate well, this should have little effect on the amount of processing and the distribution of answers. Indeed, there is little contrast between these results and Fig. 3. The average run length is almost exactly the same as in Fig. 3, which had no distractions at all. Thus Tabletop essentially ignores objects in unlikely locations on the table, focusing its attention primarily on *a priori* preferred regions. (However, when no objects are in *a priori* preferred regions, Tabletop does examine *a priori* unlikely regions.)



Eliza touches	Av. structure value	Av. no. of micro-actions
E1	66	67
E2	-	-

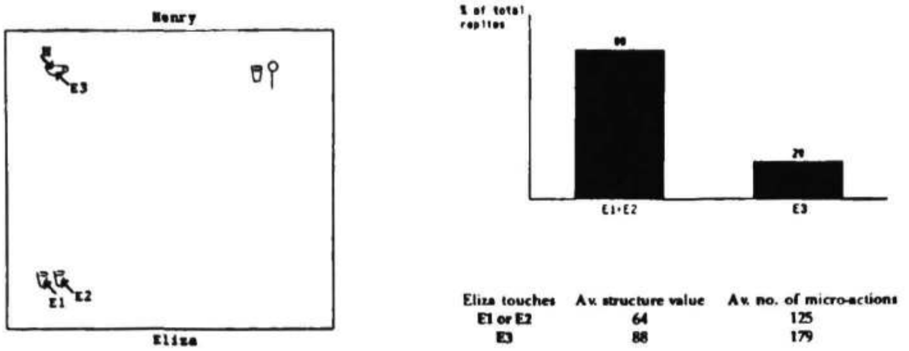
Figure 6. Figs. 6 through 10 represent the *Blockage* family. The cup and glass facing each other are not identical, but almost so: the Slipnet nodes "cup" and "glass" are very close. Also the glass is in a favorable position with respect to the cup. There is thus much pressure to choose the glass, and Tabletop always does so here. In variants, the pressures for touching H's cup are increased by creating correspondences that "usurp" E's glass. In this, the base case, there is no attempt at blockage.



Eliza touches	Av. structure value	Av. no. of micro-actions
E1	71	106
E2	94	158

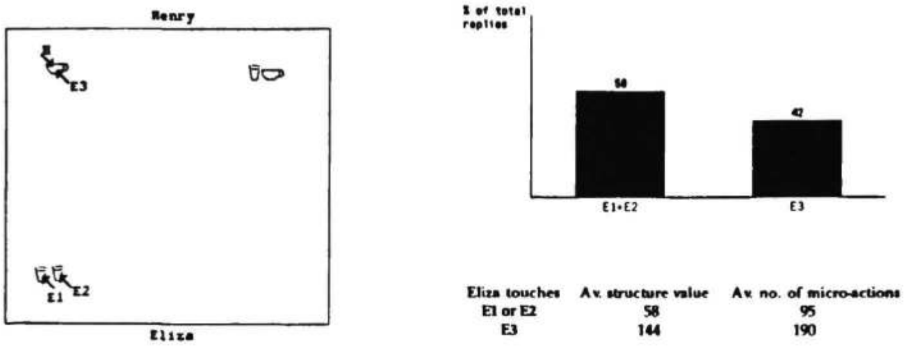
Figure 7. A glass has been added; being on H's side, it is most unlikely to be touched. Unlike the additions in the *Surround* family, this addition creates no new group. One might thus expect that this addition, like the distractions in Fig. 5, would have little effect on Tabletop's answers or resources expended. But another effect — the distant glasses' *identicality* — gives rise to a pressure to build a *correspondence* between them. When this is done, the glasses are seen as part of a single, albeit weak, structure, which exerts a *blockage* effect. (Correspondences and groups are both perceptual chunks and are similar in many ways; the former, however, tend to be weaker since their constituents, usually being far apart on the table, are not tightly bound together.)

Many subjects (40% [French 1992]) saw the glasses as counterparts. When this happens, since E's glass cannot be the counterpart *both* of H's glass *and* of his cup, just one answer remains: the literal-sameness answer, H's cup. Tabletop occasionally (12% of the time) sees the glasses as counterparts and touches H's cup. Note that the structure value of this answer is 30% higher than for E's glass, though the latter is chosen far more often. In addition, runs on which Tabletop chooses H's cup average roughly 50% longer than for E's glass. Once again, this is not surprising: answers involving deeper perception should take longer to find than those with less.



Eliza touches	Av. structure value	Av. no. of micro-actions
E1 or E2	64	125
E3	88	179

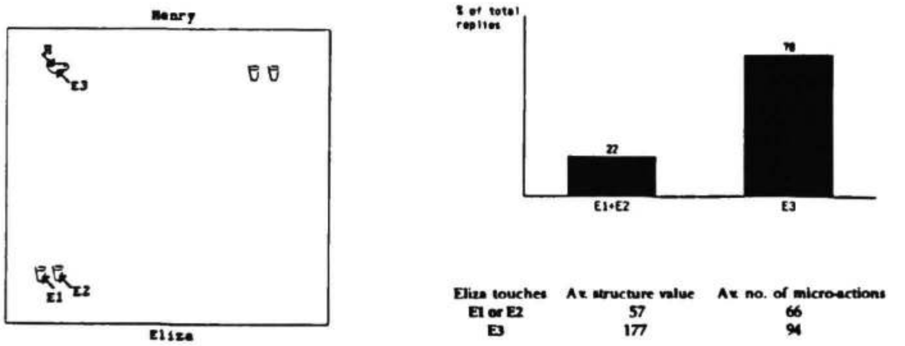
Figure 8. Two objects have been added, strongly suggesting groups. Tabletop almost always builds the group on E's side, since the two glasses are not just neighbors but identical objects. The group on H's side has less appeal, since "spoon" and "glass" are distant Slipnet nodes. Still, on many runs, both groups get built. When, in addition, a diagonal correspondence between them is built, despite its weakness, it "usurps" both glasses on E's side, forcing Tabletop to go for H's cup.



Eliza touches	Av. structure value	Av. no. of micro-actions
E1 or E2	58	95
E3	144	190

Figure 9. The group on H's side now has more appeal, as "cup" and "glass" are far closer in the Slipnet than "spoon" and "glass". Often the two objects are seen as physically *and* conceptually close. This makes for a stronger group, which in turn makes for a stronger diagonal correspondence, leading Tabletop to choose H's cup more than twice as often as in Fig. 8.

Sometimes E's group is built but not mapped to anything as a unit; in such runs, the touched cup tends to be mapped onto one of E's glasses. There is pressure to map her other glass onto H's glass (diagonally opposite identical objects make strong counterparts). But there is also counterpressure: to map E's two glasses, which have been grouped and are thus a conceptual unit, onto unrelated objects would be to disrespect their unity. Yet Tabletop does this occasionally, in which case the structure value suffers markedly. When Tabletop goes for H's cup, the structure is much better than when it chooses one of E's glasses. Also note that Tabletop takes significantly longer to build the structure that gives rise to the better answer.



Eliza touches	Av. structure value	Av. no. of micro-actions
E1 or E2	57	66
E3	177	94

Figure 10. A turning point in the *Blockage* family: Tabletop chooses H's cup over half the time. The reason is simple. Both glass–glass groups are very strong, as is the correspondence between them—strong enough, it turns out, to make Tabletop very reluctant to break it by mapping H's cup onto either of E's glasses. Tabletop thus picks H's cup 78% of the time. (Human subjects chose H's cup 66% of the time [French & Hofstadter 1991].) As one might expect, the average structure value for this answer is better than when Tabletop chooses one of E's glasses. Also as usual, it tends to take Tabletop longer (by about 40%) to get the answer having better structure.

represent the highest (most abstract) level of perception. It would thus not be exaggerated to describe Tabletop as a model of high-level vision. Of course, the raw input to Tabletop must be thought of as being the output of a prior module that carries out perceptual processing at a lower (and more modality-specific) level. Tabletop is not a model of all of vision, but of vision's high end—the end that interfaces with concepts at various levels of abstraction.

Many aspects of Tabletop's architecture can barely be hinted at herein. ([French 92] gives a much fuller presentation.) In particular, *conceptual proximity* is implemented in the *Slipnet*, a network in which each node represents a concept (strictly, the *core* of a concept), and links to other nodes establish a metric defining conceptual distances. Each node (conceptual core) has a dynamically varying "halo" (the *full* concept)—a diffuse region centered on it and probabilistically including nearby nodes. For instance, the degree to which the node "glass" is, at any moment, included in the halo of "cup" represents the current likelihood of those two concepts to be "equated" (the likelihood that their non-identity will be "forgiven") in the act of considering whether a particular cup and glass deserve being deemed counterparts (at least tentatively).

Tabletop's overall "personality"

[French & Hofstadter 91] presented Tabletop's architecture along with a few runs on three problems. Though of interest, this afforded only a limited perspective on the program's behavior. Because of its stochastic nature, Tabletop follows different pathways on different runs, and thus often comes up with different answers on different runs. Therefore, to get a feel for the program's overall behavior, one must run it not only on many different problems, but *many times on each given problem.* Only thus can one gain a clear perspective on how different combinations of pressures "pull" the program. Since the heart of the model is its ability to handle multiple interacting pressures, this is a key test.

We have probed Tabletop's "personality" by running it many times on a great variety of configurations. Inevitably, once any problem was devised, several close variants would spring to mind in which the altered pressures would alter the appeal, to humans, of various answers. By testing Tabletop on such tightly interrelated *families* of problems, we learned how it responds to diverse combinations of pressures.

In the figures we sample two families, each represented by five problems. In each problem, the table is shown on the left, with the object Henry touched indicated by an arrow with an "H". Possible responses by Eliza are indicated by arrows labeled "E1", "E2", etc. On the right, a bar graph is shown; each bar represents the frequency of one answer. All problems were run 50 times. On each run, a monitor recorded the *answer*, the final *structure value*, and the *run-length* (total number of micro-actions). Below the graph is a table giving, for each answer, the average final structure value and the average run-length for all runs yielding that answer.

Of particular interest are cases where the highest-frequency answer is not the answer having the highest final structure value. Such cases, rather than reflecting a defect of the architecture, reflect an inevitable fact about high-level perception: deep perceptions are often hard to discover; it is easy to be distracted by routes having more surface appeal. Thus Tabletop often prefers "shallow" answers, provided they have at least a modicum of plausibility, over "deep" ones (where depth is measured by structure value). It is, however, a virtue of Tabletop's parallel stochastic architecture that, by allowing simultaneous exploration at different rates along rival routes showing different degrees of promise, it is not always seduced by surface glitter, and can on occasion come up with deeper visions.

By exploring several families of "Do this!" problems, each family having many members, we built up a "performance landscape" of the program—a surface in the abstract multidimensional space of all Tabletop problems, where each dimension roughly corresponds to a given pressure. The "ridges" in this landscape represent critical combinations of pressures where the program switches from one preference to another (e.g., Fig. 10 in the Blockage family). Likewise, "peaks" and "valleys" correspond to clear and stable "likes" and "dislikes" on the program's part. By making qualitative comparisons of the locations of Tabletop's ridges, valleys, and peaks with our own personal preferences, as well as with statistics summarizing the preferences of experimental subjects, we were able to assess the psychological realism of Tabletop's "taste". (Experimental results can be found in [French 92].)

Analogy-making as high-level perception

From our point of view, the Tabletop program did a creditable job, on a qualitative level, of simulating the taste of a typical human playing the role of Eliza. (Readers can look at the bar graphs and decide for themselves whether they agree.) Despite this success, we reiterate our contention that the program is not to be judged primarily on this basis, but rather on its overall architecture, in which analogy-making falls out as a natural by-product of high-level perception, a cognitive activity that is realized by parallel processes guided by dynamically evolving pressures that emerge in response to the situation being faced.

References

[Chalmers, D., French, R., & Hofstadter, D. 91] "Perception, High-level Representation, and Analogy-making: A Critique of Artificial Intelligence Methodology". Technical Report, Center for Research on Concepts & Cognition, Indiana U., Bloomington, 1991. To appear in *Journal of Experimental and Theoretical Artificial Intelligence*, 1992.

[French, R. 92] "Tabletop: An Emergent, Stochastic Model of Analogy-Making". Ph.D. diss.,. Computer Science & Engineering Dept., U. of Michigan, Ann Arbor, 1992.

[French, R. & Hofstadter, D. 91] "Tabletop: An Emergent, Stochastic Model of Analogy-Making". In *Proceedings, 13th Annual Cognitive Science Society Conference*, pp. 708–713. Hillsdale, NJ: LEA, 1991.

[Hofstadter, D. & Mitchell, M. 91] "The Copycat Project: A Model of Fluid Concepts and Analogy-Making". Tech. Rpt., Ctr. for Resch. on Concepts & Cog'n, Indiana U., Blgtn, 1991. To appear in Barnden, J. & Holyoak, K. (eds.) *Advances in Conn'ist & Neural Computation Theory. Vol. 2: Analogical Connections*. Hillsdale, NJ: LEA, 1992.

[Mitchell, M. 90] "Copycat: A Computer Model of High-level Perception and Conceptual Slippage in Analogy-making". Ph.D. diss., Comp. & Communic'n Sci's Dept., U. Mich., Ann Arbor, 1990.