gory learning, but not in information-integration category learning. In this study, subjects alternated a trial of categorization with a trial of Sternberg (1966) memory-scanning. In a short feedback-processing-time condition, memory scanning immediately followed categorization, whereas in a long feedback- processing-time condition, categorization was followed by a 2.5 second delay and then by memory scanning. Information-integration category learning was unaffected by manipulations of this inter-trial interval, whereas rule-based category learning was significantly impaired when subjects had only a short time to process the categorization feedback.

It is important to realize that these dissociations are not driven simply by differences in the difficulty of rule-based versus information-integration tasks. First, in several cases the experimental manipulation interfered more with the learning of the simple rule-based categories than with the more difficult information-integration strategies (Maddox et al. 2003; Waldron & Ashby 2001). Second, most of the studies explicitly controlled for difficulty differences either by decreasing the separation between the unidimensional rule-based categories, or by using a more complex two-dimensional conjunction rule in the rule-based conditions. Both manipulations increase the difficulty of rule-based categorization, yet in no case did such increases in rule-based task difficulty affect the qualitative dissociations described above.

Finally, we note that all of these dissociations were predicted in a parameter-free, a priori manner by the dual-system category-learning model COVIS (Ashby et al. 1998).

Rules work on one representation; similarity compares two representations

Todd M. Bailey

School of Psychology, Cardiff University, Cardiff CF10 3YB, United Kingdom.
baileytm1@cardiff.ac.uk **http://www.cf.ac.uk/psych/home/baileytm1**

Abstract: *Rules* and *similarity* refer to qualitatively different processes. The classification of a stimulus by rules involves abstract and usually domain-specific knowledge operating primarily on the target representation. In contrast, similarity is a relation between the target representation and another representation of the same type. It is also useful to distinguish associationist processes as a third type of cognitive process.

It is not the number of features, it is what you do with them that counts. The conceptual difference between rules and similarity has more to do with the number of object representations on which they operate than with the number of object features they process. For example, in a study of rhythm learning, Bailey et al. (1999) evaluated various models for determining which syllable in a word gets the main stress (e.g., si-mi-LA-ri-ty, not si-MI-la-ri-ty). The classical approach in linguistics involves rhythm rules that apply one after the other (e.g., Halle & Vergnaud 1987). These rules operate on a single representation, namely the one representing the phonological structure of the target word. In contrast, an exemplar model of stress assignment like the one described by Bailey et al. has comparisons between two representations at its core – the phonological representation for the target word is compared to a familiar word whose representation is recalled from memory. One could argue that the classical rules refer to only a small subset of the target word's phonological features, and that perhaps the similarity process underlying the exemplar model refers to more of these features. However, that small quantitative distinction misses the fact that the cognitive mechanisms hypothesized by these two models are qualitatively quite different. The rules require a working memory capable of representing the phonology of a single word, along with an abstract body of knowledge that effectively categorizes the target word so that it receives stress on a particular syllable. The exemplar model requires representations for two words to be juxtaposed so that a similarity relation can be computed between them (along with some additional secondary machinery to aggregate across multiple pairwise comparisons and classify based on the result).

Cognitive processes operate on representations, and these examples illustrate the distinction between unary and binary operations. Another theory of stress assignment, based on optimality theory (Prince & Smolensky 1993), is a system of soft (violable) constraints on rhythm structures (Tesar 1997). Variations of the target word with different rhythm structures are evaluated against the set of constraints, and the optimal variation, the one that is most consistent with the highest-ranking constraints, determines the stress pattern. The core of this constraint-satisfaction process is the evaluation of a single representation with respect to a set of domain-specific constraints. Those constraints, and their relative rankings, embody abstract knowledge of stress patterns. What is not involved in the constraint-satisfaction process is juxtaposition between two phonological representations. In this regard, optimality theory is similar to classical linguistic rules and qualitatively distinct from the exemplar model and its similarity comparisons. The same can be said for the "non-metrical" constraints on stress proposed in Bailey (1995).

The distinction between unary and binary operations yields sensible classifications for many models of cognitive processes, including those mentioned in the target article. Nevertheless, it may be helpful for a gross taxonomy of cognitive models to include associationist models as a third type. For example, the perceptron model of stress assignment (Gupta & Touretzky 1994) determines the location of stress using a two-layer connectionist network. The model is unary in the sense that it involves a single active phonological representation – that of the target word. However, the operation performed on this one representation is not determined by abstract domain-specific knowledge, but by a transparent mapping based on statistical properties of previous representations. Stress assignment for a word could also be determined based on the familiarity of its component chunks, along the lines of fragment models of artificial grammar learning (e.g., Perruchet & Pacteau 1990; Servan-Schreiber & Anderson 1990). Superficially, one might be tempted to think that fragment models are like exemplar models because familiar fragments, like exemplars, can be represented individually in a memory store. However, unlike exemplars, fragments are fundamentally incommensurate with the target representation in exactly the same sense that a wheel is incommensurate with a car – they stand in a part-whole relation. Decomposing the target representation into its component fragments and assessing their familiarity yields a measure of how the target relates to aggregate statistical properties of past targets of one type or another. In this sense, fragment models are similar to connectionist models and therefore belong in the same class of associationist models.

The three-way distinction between abstract unary operators (rules, constraints, etc.), binary operators (comparison to an exemplar or prototype), and associative mappings (connectionist networks, fragment models) preserves the intuitive distinction between rule-based processing and other sorts of models. It also provides a sensible way to think about certain hybrid models, like Hummel and Holyoak's (1997) theory of analogical access and mapping. In relating a proposition like "John loves Mary" to "Bill likes Susan" versus "Peter fears Beth," Hummel and Holyoak's system relies on associative mappings from individual predicate and object units (John, Mary, etc.) to a distributed semantic memory. At the same time, it maintains representations for two (or more) propositions in working memory, and relates one to the other to determine their analogical similarity. This is a good example of a hybrid model composed of an associative component and a binary similarity component. Using Pothos's proposed continuum we could describe it as a hybrid of two Similarity processes, but that is unnecessarily uninformative. The continuum between Rules and Similarity is interesting, but overlooks important qualitative differences among models of cognitive processes, including the difference between rules and similarity.

https://doi.org/10.1017/S0140525X05240013 Published online by Cambridge University Press