Locally-to-Globally Consistent Processing in Similarity

Robert L. Goldstone
Indiana University
rgoldsto@ucs.indiana.edu

Abstract

SIAM, a model of structural similarity, is presented. SIAM, along with models of analogical reasoning, predicts that the relative similarity of different scenes will vary as a function of processing time. SIAM's prediction is empirically tested by having subjects make speeded judgements about whether two scenes have the same objects. The similarity of two scenes with different objects is measured by the percentage of trials on which the scenes are called the same. Consistent with SIAM's prediction, similarity becomes increasingly influenced by the global consistency of feature matches with time. Early on, feature matches are most influential if they belong to similar objects. Later on, feature matches are most influential if they place objects in alignment in a manner that is consistent with other strong object alignments. The similarity of two scenes, rather than being a single fixed quantity, varies systematically with the time spent on the comparison.

Introduction

The similarity of two things is not simply a relation between the two things; it is a relation between the two things and the comparison-maker. Similarity assessments must be constructed by a process that compares the items in question. Sometimes the process is straightforward. The similarity of cigars and cigarettes is easily determined. Determining the more abstract similarity between cigarettes and time bombs (Ortony, Vondruska, Foss, & Jones, 1985) seems to take a longer time. The fact that similarity develops along a time course suggests that similarity does not immediately impinge upon our perceptual system. Instead, perceptual and cognitive processes actively <u>build</u> a conception of similarity.

The time course of similarity assessments provides a useful tool for investigating the comparison process. If we dispatch with the assumption that similarity is "out there" in the objective world, then the question of "How does similarity develop?" becomes crucial. One method for understanding how comparisons are made is temporal analysis.

Dynamic Models of Similarity

General models of similarity have not often addressed temporal aspects of processing (Carroll & Wish, 1974; Tversky, 1977). These models do not consider similarity to be a dynamically evolving quantity. Instead, their equations for similarity give single "endpoint" estimates. However, specific process models have been developed for some specialized tasks. For example, similarity has often been measured by the time elapsed, or the errors made, when subjects determine if two displays are different. The assumption made is that the longer it takes to respond that the displays are different, or the more times that different displays are erroneously thought to be the same, the more similar the displays are. Specific processing mechanisms have been hypothesized to account for how this speeded same/different task is executed (for a review, see Farrell, 1985). The speeded same/different task will be used to measure similarity in the experiment to be reported. The speeded same/different task cannot replace subjective ratings as a method for investigating similarity, but it does provide a converging measure that is relatively immune to experimenter demands and subjects' high-level reasoning strategies.

Recently, a general model of similarity has been developed called SIAM that also hypothesizes a dynamic time course for comparisons (Goldstone 1991; Goldstone & Medin, in press). According to SIAM (Similarity as Interactive Activation and Mapping), when structured scenes are compared, the parts of one scene are aligned, or placed in correspondence, with the parts of the other scene. Emerging correspondences influence each other as processing continues. With sufficient time, the strongest correspondences will be those that are consistent with other correspondences. Similarity is determined by a process of interactive activation between feature and object correspondences. The degree to which features from two scenes are placed in correspondence depends on how strongly their objects are placed in correspondence. Reciprocally, how strongly two objects are placed in correspondence depends on the correspondence strength of their features.

The details of SIAM are discussed elsewhere (Goldstone, 1991). Essentially, SIAM's network architecture is composed of nodes that excite and inhibit each other. Nodes represent hypotheses that two entities correspond to one another in two scenes. For the present purposes, two types of nodes are important: feature-to-feature nodes, and object-to-object nodes. Each feature-to-feature node represents an hypothesis that two features correspond to each other. One feature-to-feature node is assigned to every possible pair of alignable features. As the activation of a feature-to-feature node increases, the two features referenced by the node will be placed in stronger correspondence. Object-to-object nodes represent hypotheses that two objects correspond to each other.

Network activity starts by features being placed in correspondence according to their physical similarity. After this occurs, SIAM begins to place objects into correspondence that are consistent with the feature correspondences. As objects begin to be put in correspondence, activation is fed back down to the feature (mis)matches that are consistent with the object alignments. In this way, object matches influence activation of feature matches and feature matches influence the activation of object matches concurrently.

Activation spreads in SIAM by two principles: 1) nodes that are consistent with one another send excitatory activation to each other and 2) nodes that are inconsistent inhibit one another. Nodes are inconsistent if they produce many-to-one mappings, and are consistent otherwise. Processing in SIAM starts with a description of the scenes to be compared. Scenes are described in terms of objects that contain feature slots that are filled with particular feature values. On each "slice" of time (cycle), activation spreads between nodes. Nodes that are highly active are weighted heavily in the similarity assessment.

SIAM shares architectural commonalities with McClelland and Rumelhart's (1981) interactive activation model of word perception and Marr and Poggio's model of depth perception (1979), and is highly related to the SME (Falkenhainer, Gentner, and Forbus, 1990) and ACME (Holyoak and Thagard, 1989) models of analogical reasoning. In ACME, SME, and Marr and Poggio's model, there are pressures against developing many-to-one mappings, and pressures in favor of developing mutually consistent mappings. The models of McClelland and Rumelhart, Holyoak and Thagard, and Marr and Poggio are all examples of what Marr (1982) calls "cooperative algorithms." Cooperative algorithms create globally consistent mappings by local interactions between units. SME also moves from locally determined mappings to globally consistent mappings with more processing. As we will see, SIAM incorporates a similar local-to-global processing principle.

A Behavioral Prediction of SIAM

In SIAM, object correspondences depend on feature and object correspondences[1]. SIAM

---

[1] In the full version of SIAM, object correspondences also depend on role

initially begins to place objects in correspondence on the basis of their featural overlap; the more featural commonalities two objects have, the more strongly they will be placed in correspondence. However, the strength of an object correspondence is also influenced by its consistency with other object correspondences. If two objects from one scene correspond to a single object in another scene, then the two correspondences are inconsistent and will decrease each others' strength. SIAM, like ACME and SME, predicts that object correspondences will become increasingly influenced by other object correspondences with time, as activation spreads between nodes .

One prediction of this temporal processing is that feature matches that are inconsistent with the set of globally consistent correspondences should tend to influence similarity less with time. Globally consistent feature matches should become more influential with time. A set of mappings between objects is globally consistent if it a) yields only one-to-one mappings, and b) maximizes the number of matching features that belong to corresponding entities. Even though object A from scene 1 may be most similar to object B from scene 2, these objects may not be a part of globally consistent set of mappings. In particular, if other objects from scene 1 are also fairly similar to B, and other objects from scene 2 are fairly similar to A, and if we only allow one-to-one correspondences, then placing A in correspondence with B may not maximize the number of feature matches between aligned objects.

In SIAM, object correspondences will first be based on feature matches, the only information available. Objects that are featurally similar will begin to be placed in correspondence. With time, object correspondences will be inhibited by inconsistent object correspondences, and excited by consistent object correspondences. By these interactions, object correspondences that are consistent with many other object correspondences become stronger. In turn, the feature matches that belong to these globally consistent correspondences will receive more weight. In this manner, the global consistency of feature matches comes to influence similarity more with increased processing time.

correspondences that serve to align objects that play similar roles in their scenes.

Experimental Support for a Local-to-global[2] Processing Shift

To test the influence of processing time on globally consistent and inconsistent feature matches, subjects are shown pairs of scenes; sample scenes are shown in Figure 1. Each scene contains two butterflies, and each butterfly contains four features. Subjects must decide whether the two scenes contain the same butterflies within a specified deadline. A symbolic representation is shown below each of the butterflies in Figure 1. For example, the target scene is composed of butterflies "AAAA" and "BBBB," where the letters refer to different values along the four dimensions {body shading, head type, tail type, wing shading}. The butterfly "XABA" has feature matches on the second and fourth dimensions with "AAAA", and a feature match on the third dimension with "BBBB."

First, consider trials in which the target scene is compared to the base scene. Both of the butterflies in the target scene have more matching features in common with the base scene's left butterfly than right butterfly. The base butterfly "BABA" has two matches in common with both of the target scene's butterflies. Thus, if we only consider the locally preferred mappings, we would map both target butterflies onto the top butterfly of the base. However, if the global consistency of object mappings is maintained, then this many-to-one mapping is not permitted. The best globally consistent mapping is to map the left butterflies onto each other, and the right butterflies onto

2 The term "local-to-global," as used here, is only distantly related to previous researchers' (e. g. Navon, 1977) claim for a processing shift from global (holistic) to detailed/analytic similarities. The current claim concerns the increasing importance of globally consistent features matches on similarity.

each other. In short, "BBBB" corresponds to "BABA" if we only consider local feature matches, but "BBBB" corresponds to "XXXB" if we consider the influence that object correspondences can have one another.

The target scene is also compared to two derivatives of the base scene. Each derivative differs from the base scene by only a single feature. For Figure 1A, one of the local feature matches is removed, leaving all of the globally consistent matches intact. For Figure 1B, one of the globally consistent matches is removed, and all of the local matches are preserved. The empirical questions of primary interest is "Is the target scene more similar to the scene in Figure 1A or 1B, and does the relative similarity of the scenes depend upon the processing time allowed?"

Thirty-three undergraduates were presented with 608 displays each. On half of the trials, two copies of the target scene were displayed. On these trials, the subjects' correct response was "same." On the other half of the trials, displays consisted of the target scene and one of the other three scenes shown in Figure 1. Butterfly position, dimension order, dimension values, scene location, and display type were all randomized. Displays were presented on Macintosh SE30 computers.

The subjects' task was to press a key with one hand if the butterflies of one scene were the same as the butterflies of the other scene, and to press a key with the other hand if the two scene's

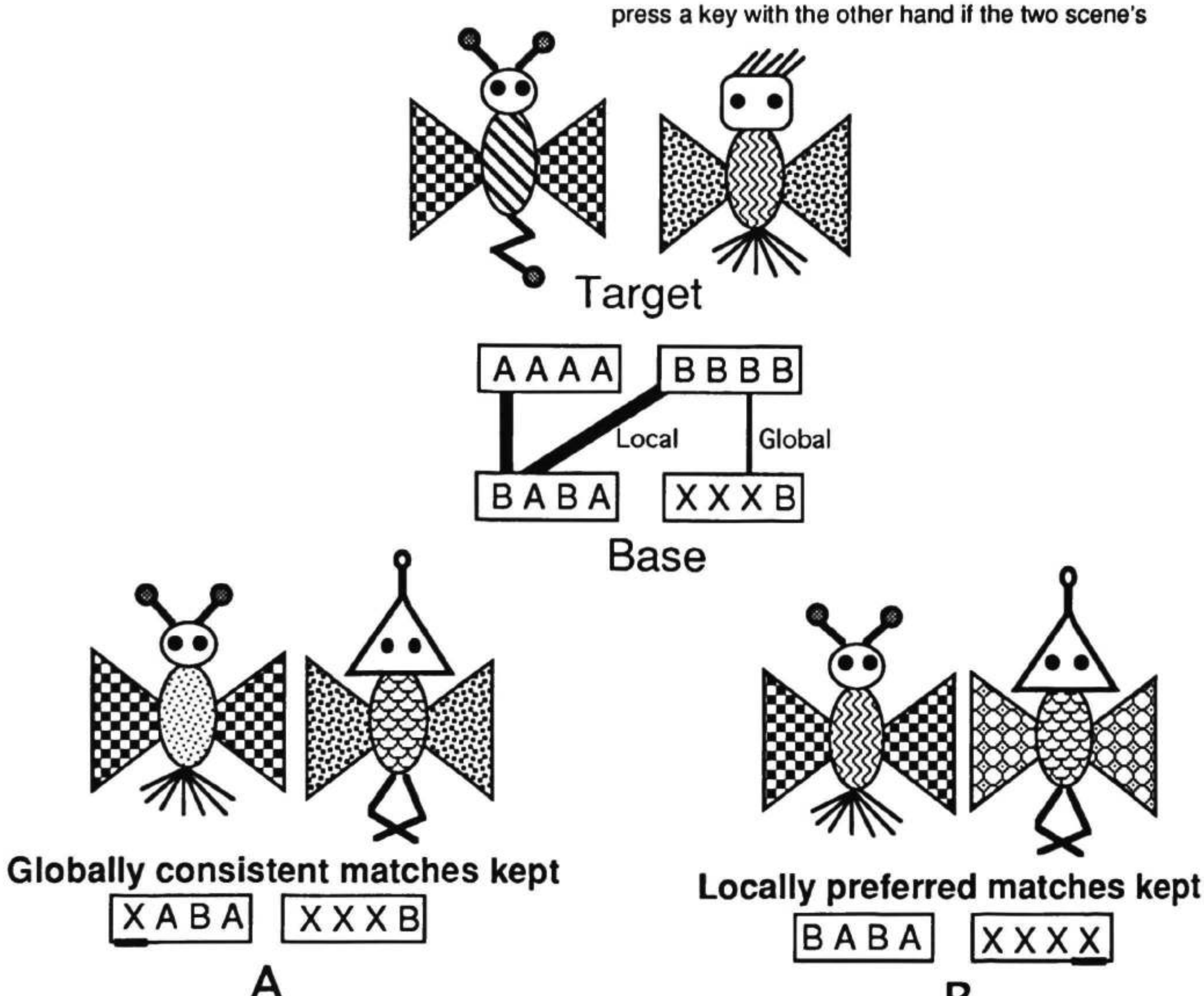

Figure 1. Sample scenes used in the experiment.

butterflies were different. It was stressed to subjects that the same butterflies did not have to be in the same positions in their scene in order to respond "same." The experiment was divided into 19 blocks. On each block, subjects were given a "very fast," "fast," or "fairly slow" deadline (1, 1.84, and 2.68 sec respectively). If a subject did not respond before the deadline passed, the message "OVERTIME" appeared on the screen.

The significant (F (4, 288) = 3.94, mse = 8, p<.05) cross-over interaction between deadline and type of display is shown in Figure 2. If subjects are forced to respond within a short deadline, the display that preserves the locally preferred match is more often incorrectly responded to as "same" than the display that preserves the globally consistent match. The opposite effect is found when subjects are given longer to respond. The four mean error rates of particular interest are: slow-deadline/global-matches-kept = 5%, slow-deadline/local-matches-kept = 3%, fast-deadline/global-matches-kept = 18%, and fast-deadline/local-matches-kept = 21%. A planned comparison of these four data shows a significant interaction between deadline and type of scene on error rate (F (1,288) = 3.46, mse=6.8, p<.05). The overall times to correctly respond "Different" to the different displays are not significantly different (base = 1.147 sec, global match kept = 1.137 sec, local match kept = 1.135 sec).

Implications

If similarity is measured by the percentage of trials that scenes with different butterflies are incorrectly judged to be the same, then the obtained results are consistent with SIAM's prediction. More incorrect "same" judgments are found for short deadlines when local matches are preserved. More incorrect "same" judgments are found for the longest deadline when global matches are preserved. This is consistent with SIAM's dynamic account of similarity. The influence of one object-to-object mapping on another takes time to develop, and until it is developed, object-to-object mappings will be largely determined by feature-to-feature matches. Locally consistent matches are more important than globally consistent matches for similarity early in processing (fast deadline). Later in processing, globally consistent matches gain in importance relative to local matches. At first, both butterflies of the target are mapped onto one butterfly of the other scene, but with time the influence of one mapping redirects the other mapping.

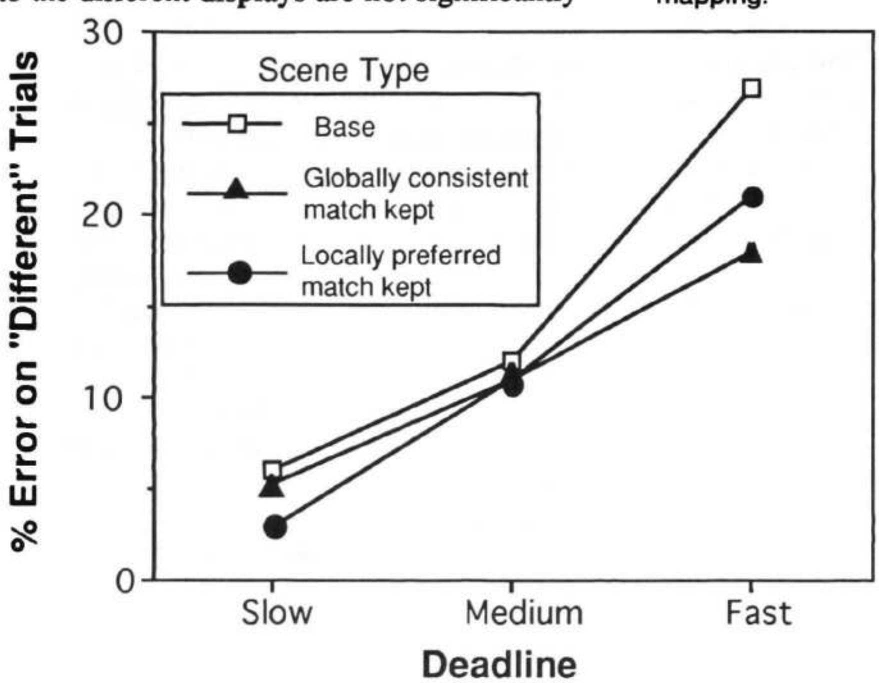

Figure 2. Results showing an interaction between deadline and type of scene.

The experiment indicates that featural similarity cannot completely predict object correspondences. Objects will tend to be aligned if they share many features, however object alignment also depends on the similarity of other objects pairs. Butterfly BBBB from the target scene of Figure 1 is most similar to butterfly BABA of the base scene, but it is placed in proper alignment with butterfly XXXB. BBBB corresponds to XXXB and not BABA because BABA is also similar to the target scene's other butterfly, AAAB. By aligning BBBB with XXXB and BABA with AAAB, the number of matching features between consistently (one-to-one mapping) aligned objects is maximized. With increased processing time, SIAM and subjects both seem to base object correspondences more on global consistency than on the local similarity of objects.

The experiment supports a notion of similarity as constructed over time. In fact, the results are problematic for any model that hypothesizes that two entities have a single process-independent similarity value. We cannot assign single estimates for the similarity of the target scene and Figure 1A, and the target scene and Figure 1B, because at different times each is more similar than the other. Figure 1A is more similar to the target scene on slow deadlines, but Figure 1B is more similar to the target scene on fast deadlines. The similarity of two entities seems to depend on the particular mechanisms of the comparison process. In the current case, comparisons seem to involve a process in which locally determined correspondences give way to globally consistent ones. More generally, the outcome of a comparison seems to depend not just on the things compared, but also on the process that is doing the comparing.

Acknowledgements

This paper has benefitted greatly by many useful suggestions by Douglas Medin, Dedre Gentner, Ed Smith, Keith Smith, and Keith Holyoak. The research was supported by NSF Grant BNS-87-20301 awarded to Dedre Gentner and Douglas Medin.

References

Carroll, J. D., & Wish, M. (1974). Models and methods for three-way multidimensional scaling. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.) Contemporary developments in mathematical psychology (Vol. 2, pp. 57-105). San Francisco:Freeman.

Falkenhainer, B., Forbus, K.D., & Gentner, D. (1990). The structure-mapping engine: Algorithm and examples. Artificial Intelligence, 41, 1-63.

Farell, B. (1985). "Same"- "Different" judgments: A review of current controversies in perceptual comparisons. Psychological Bulletin, 98, 419-456.

Goldstone, R. L. (1991). Similarity, Interactive Activation, and Mapping. Unpublished doctoral dissertation, University of Michigan.

Goldstone, R.L., & Medin, D.L. (forthcoming). Interactive Activation, Similarity, and Mapping. in K. Holyoak and J. Barnden (Eds.) Advances in Connectionist and Neural Computation Theory, Vol. 2: Connectionist Approaches to Analogy, Metaphor, and Case-Based Reasoning. Ablex

Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constrain satisfaction. Cognitive Science, 13, 295-355.

Marr, D. (1982). Vision. San Francisco: Freeman.

Marr, D., and Poggio, T. (1979). A computational theory of human stereo vision. Proceedings of the Royal Society of London, 204, 301-328.

McClelland, J. L., & Rumelhart, D.E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. Psychological Review, 88, 375-407.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. Cognitive Psychology, 9, 353-383.

Ortony, A., Vondruska, R. J., Foss, M. A., and Jones, L. E. (1985). Salience, similes, and the asymmetry of similarity. Journal of Memory and Language, 24, 569-594.

Tversky, A. (1977). Features of similarity. Psychological Review, 84, 327-352.