CS 786G *Cognitive Modeling*
Week 6 Discussion Paper
Visual and Spatial Reasoning

Steve Woods
89800462

May 26, 1995

Ching (1993) describes a representational framework for visual information that addresses both the problems of computational problem solving and (human) cognitive modeling. This framework, attributed hypergraphs (AHGs), is derived from work in robot vision systems, and essentially is composed of a layering of graphs in which (for instance) at one level nodes represent corners of objects and arcs the spatial or other relationships of these corners and consequently a graph represents a surface; at the next level, the surfaces themselves are represented by nodes and relationships amongs these surfaces by arcs leaving the graph at this level to represent "characteristic" building blocks; at yet another level, a graph node represents a single building block, and the arcs the relations among the blocks, giving a graph which represents a complex object. In this way, scenes may be decomposed into their essential parts at "detail" levels, or composed into broader contexts at "breadth" levels of expansion. In addition, the graphs are essentially free-formed with respect to attribution, allowing labelings to range over the imagination, with arcs describing domain relationships such as *function* as might be the case for functionally related scence objects such as a key and a lock, or describing more general *spatial* relations such as distance between objects/points of a scene, or perhaps a shared orientation relationship.

Ching (1993) cites several psychological works that lend credence to a hierarchical imagery representation model for humans which is capable of creating imagery by extracting working-memory representations from long-term memory, essentially providing a mapping of generalized information to specific by overriding general defaults with specific information where appropriate. Attributed hypergraphs replace the previous two-model paradigm with a single model that essentially allows for instantiation, and presumably also generalization in the reverse stage of comprehension or learning rather than imagination or construction from memory.

Ching (1993) describes AHGs in the context of imagery and recognition of objects and observes that efficient graph morphism algorithms such as presented in Wong, You & Chan

(1990) may be exploited for computationally efficient reasoning with this representation. This mapping is not made particularly clear in this paper. It should be noted that the mapping of a "known" AHG representing a particular (possibly) generalized scene or object generally involves more than a graph to graph comparison for equality or correspondence, rather, it requires a potential mapping of many bits of uncorrelated image information into the "model" slots and a subsequent evaluation of the "accuracy" of this fit to a known model. Only once this context is gained may a morphism model be used to gauge the accuracy of the perceived "image" and the learned "model".

A closely related problem involving recognition of patterns according to spatial relationships of certain classes of objects has been proposed in (Woods 1993*b*, Woods 1993*a*). In this work generalized models of a certain object type are represented in an attributed graph structure which implicitly describes a constraint satisfaction problem (CSP) in terms of spatial and functional relations amongst the nodes or variables. A "situation" of domain objects exists in which one would like to recognize instances of the (general) template. Further, the identification of "critical" and "supporting" features is embedded in the representation through the explicit construction of abstraction hierarchies for the CSP which specify the alternatives for refining or generalizing a potential problem and consequently the potential solution set. Related CSP work (Mackworth, Mulder & Havens 1985, Sidebottom & Havens 1992) has focused on this type of hierarchical solution to difficult CSPs by expressing the hierarchy in terms of the potential range of the domains of the variables or node, for instance, generalizing the potential set of matches for each node restricting to a possible subset or class of the domain values at each level of abstraction.

The CSP approach discussed above has been applied to spatial recognition problems in which large quantities of noise is coupled with only partial identification and classification of situation objects, resulting in "images" that are highly confusing, even for domain experts. The identified matches are essentially highlighted in the complex situation as alternatives for an expert to consider, and essentially helps to focus attention on likely solutions. Partial solutions in such poorly perceived situation are highly desired and allow for an exploitation of the template knowledge for further inference about the existence of objects not currently recognized in the "image".

Ching (1993) observes that "graphs are particularly *natural* for capturing structural and topological information in an image", and this observation may be reinforced through the experience of experts (Woods 1993*b*) in the spatial template recognition domain, who although naive in creating or using graph formalizations, found them a quite easy-to-understand representation of the actual matching task at which they are experienced, both for the case of *spatial relations* and of *functional relations*. The high degree of actual correspondence between representation and object represented is certain to play a large part in this "naturalness". Many interesting questions remain about both the efficiency and usefulness of currently known algorithms which are able to reason using such an expressive representation, particularly in cases that approach anywhere near real-world complexity. Identifying and exploiting methods of human accomodation of extremely complex situations may be of great benefit, examples of which include the ways in which people are able to both abstract situations for important detail at a glance (not always successfully) and also the ways in which they can "focus" attention and take advantage of features of spatial locality in object identification.

References

Ching, J. (1993), Computer imagery: An attributed hypergraph framework for visual cognition, *in* P. Thagard, ed., 'CS786G Course Notes', University of Waterloo. Ph.D. thesis proposal, pp. 14-46, Permission to copy granted by John Ching.

Mackworth, A., Mulder, J. & Havens, W. (1985), 'Hierarchial arc consistency: exploiting structured domains in constraint satisfaction problems', *Computational Intelligence* **1**, 188–126.

Sidebottom, G. & Havens, W. (1992), 'Hierarchical arc consistency for disjoint real intervals in constraint logic programming', *Computational Intelligence* **8**(4), 601–623.

Wong, A., You, M. & Chan, S. (1990), 'An algorithm for graph optimal monomorphism', *IEEE Transactions on Systems, Man, and Cybernetics* **20**(3), 628–636.

Woods, S. (1993*a*), Interactive recognition of spatially defined model deployment templates, Research memorandum, Combat Intelligence Automation Group, Defence Research Establishment Valcartier, Courcelette, Quebec CANADA. unclassified.

Woods, S. (1993*b*), A method of interactive recognition of spatially defined model deployment templates using abstraction, *in* H. Merklinger, M. Farooq, P. Roberge & R. Grodski, J.J.and Dobson, eds, 'Proceedings of the Knowledge -Based Systems and Robotics Workshop', Department of National Defence, Government of Canada, pp. 665–675.

Woods, S. (1994*a*), Learning refinements to spatial template models based on false-positive and false-negative examples, advance copy, technical report.

Woods, S. (1994*b*), On the complexity of graph isomorphism : relationship to the spatial template recognition problem, advance copy, technical report.