II. MODEL-BUILDING IN THE VISIONS SYSTEM[1]

Thomas Williams, John Lowrance, Allen Hanson , and Edward Riseman
Computer and Information Science Department
University of Massachusetts
Amherst, Massachusetts 01003

Abstract

The semantic Interpretation system of VISIONS receives a segmentation containing symbolically encoded regions, boundaries and endpoints. The model-builder consists of four major components: 1) multiple levels of representation for both the image-specific model and long-term general knowledge, 2) a set of modular knowledge sources, 3) a hierarchical modular strategy to control application of the knowledge sources, and 4) a tree representing the current state of search through the space of possible models.

Multiple Levels of Representation

A model can be viewed as a set of short-term instantiations of long-term general concepts relative to the current environment [Williams and Lowrance, 1977]. This suggests that visual knowledge should be subdivided into a short-term-image-specific model and long-term general knowledge components. Visual knowledge also includes information at many levels of abstraction [Erman and Lesser, 1975; Hanson and Riseman, 1976]. The primary levels of representation include frames of stereotypical situations [Minsky, 1975], the objects in those scenes, the volumes and surfaces of those objects, and the regions, boundary segments and endpoints of those surfaces. An example of an instantiated element (short-term) in the model would be a node (0BJI) representing an association between a region (RI) and a long-term general concept (TRUNK).

Processes as Modular Sources of Knowledge

The model-building processes are viewed as modular Knowledge Sources [Erman and Lesser, 1975]. KS's are responsible for the generation/verification of hypotheses. Typical KS's include attribute matcher, perspective, shadow, occlusion, shape, etc. KS application is controlled by the model-building strategy in the search for a satisfactory model.

Hierarchical Modular Control Strategy

The complexity of the image interpretation task, in general, necessitates the integrated application of many different processes. A hierarchical modular control structure (Figure 2) provides a powerful tool for exploring the large number of potentially interesting strategies for using these processes.

A model-building strategy is a selected configuration of modules from the model-builder library. The strategy must determine which model in the model search space to expand, which level of representation within the model is to be selected, and which theory (i.e., node) at the selected level will be expanded.

At each level of this hierarchy the hypothesis-test paradigm is implemented as a sequence of three types of control modules: 1) FOCUSing on an appropriate aspect of the task, 2) EXPANDing that aspect by generating new hypotheses, and 3) VERIFYing those hypotheses. Each control module makes strategic decisions at its level determining which lower level control modules and/or KS's to apply. For example, the model search space expander calls: 1) a model focusser (strategy to select a partially developed model from the search space), 2) a model expander (strategy to hypothesize an incremental change to the partial model), and then 3) a model verifier (strategy to determine the correctness of a hypothesis).

Model Search Space

The current state of the search is represented as a tree of partially constructed models, where each node contains an incremental change from its parent node. In addition, a sketch of the history of construction for each model will be stored. This history might include an indication of the KS's and the strategy modules primarily responsible for each hypothesis, as well as any portions of the model on which the decision was dependent. This information will be used to carefully analyze and understand the model construction process during system development. Eventually it will be used by the strategy in making its decisions.

Implementation

The representation and manipulation of both information and processes in the high-level portions of VISIONS is supported by a system of data and procedural primitives. This system currently consists of a library of default control modules, a graph processing language [Lowrance, 1977], and a relational data base [Konolige, 1977]. All of these are implemented in ALISP [Konolige, 1975], the UMass version of LISP.

References

Erman, L. and Lesser, V. [1975] "A Multi-Level Organization for Problem-Solving Using Many, Diverse, Cooperating Sources of Knowledge," Proc. of Fourth IJCA1, September 1975, pp. 483-490.

Hanson, A. and Riseman, E. [1976] "A Progress Report on VISIONS: Representation and Control in the Construction of Visual Models," COINS TR 76-9, Univ. of Mass., July 1976.

This work was supported by NSF grant DCR75-16098.

[2]A. Hanson is affiliated with Hampshire College, Amherst, Massachusetts 01002.

Hanson, A., Riseman, E, and Williams, T. [1976] "Constructing Semantic Models in the Visual Analysis of Scenes," Proc. of IEEE MSAC[2], April 1976, pp. 97-102.

Konolige, K. [1975] "The ALISP Manual," Univ. Computing Ctr., Univ. of Mass., August 1975.

Konolige, K. [1977] "The ALISP Relational Database," COINS TR 77-9, Univ. of Mass., July 1977.

Lowrance, J. [1977] "GRASPER Reference Manual," COINS Tech. Report, Univ. of Mass., in preparation.

Minsky, M. [1975] "A Framework for Representing Knowledge," The Psychology of Computer Vision (P. Winston, Ed.), McGraw-Hill, pp. 211-277.

Williams, T. and Lowrance, J. [1977] "Model-Building in the VISIONS High Level System," COINS TR 77-1, Univ. of Mass., January 1977.

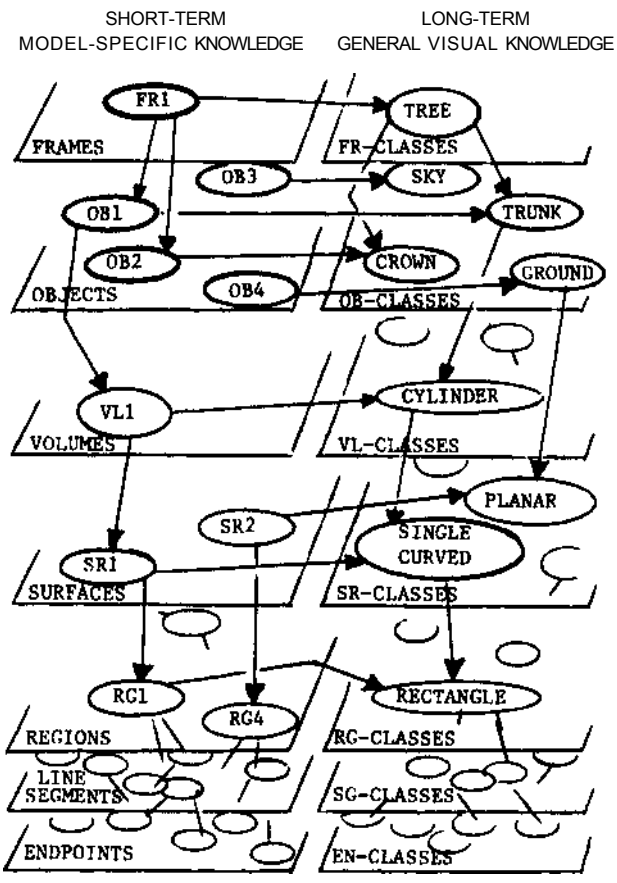

Figure 1 Declarative Knowledge is divided into: 1) Short-Term and Long-Term, and 2) levels of abstraction. A model consists of associations between Long-Term concepts and Image specific entities.

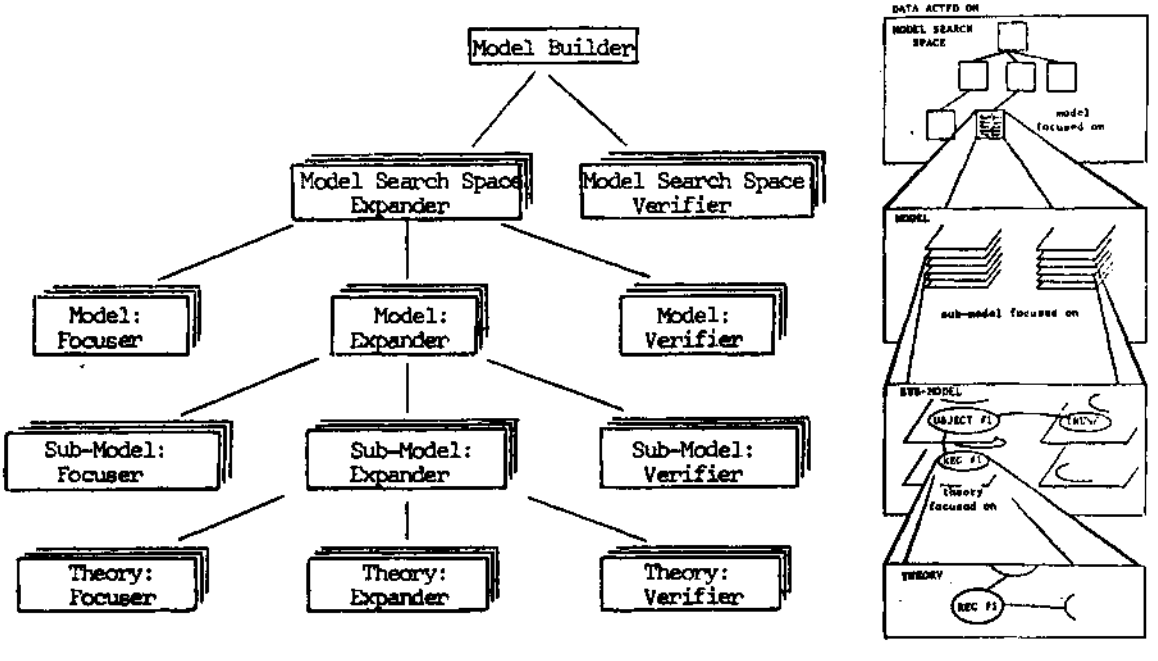

Figure 2 The Structure of Control. A modular control strategy is responsible for the integrated application of knowledge sources. The control modules are organized according to type (FOCUS, EXPAND, VERIFY) and data acted upon (Search Space, Model, Sub-Model, and Theory).