**CELEST NUGGET FOR NSF**

**LEARNING A KNOWLEDGE HIERARCHY**

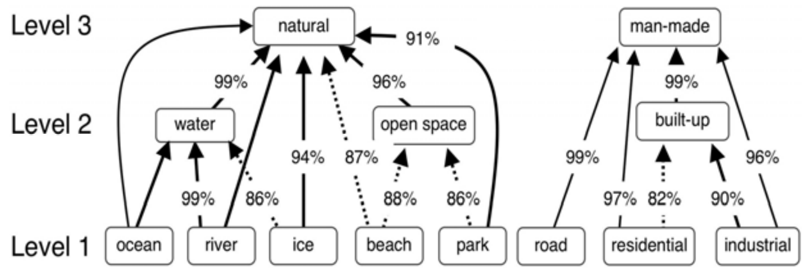


Image fusion has been defined as "the acquisition, processing and synergistic combination of information provided by various sensors or by the same sensor in many measuring contexts." (Simone et al., 2002, p. 3) When multiple sources provide inconsistent data, such methods are called upon to select the accurate information components. As quoted by the International Society of Information Fusion: "Evaluating the reliability of different information sources is crucial when the received data reveal some inconsistencies and we have to choose among various options." For example, independent sources might label an image pixel *beach* or *road* or *river*. A fusion method could address this problem by weighing the confidence and reliability of each source, merging complementary information, or gathering more data. In any case, at most one of these answers is correct. Methods developed by Gail Carpenter and CELEST colleagues address a complementary and previously unexamined aspect of the information fusion problem, seeking to derive consistent knowledge from sources that are paradoxically both inconsistent and accurate. This is a problem that the human brain solves well. A young child who hears the family pet variously called *Spot*, *puppy*, *dog*, *dalmatian*, *mammal*, and *animal* is not only not alarmed by these conflicting labels but readily uses them to infer functional relationships. An analogous problem for information fusion methods seeks to classify the terrain and objects in an unfamiliar territory based on intelligence supplied by several reliable sources. Each source labels a portion of the region based on sensor data and observations collected at specific times and based on individual goals and interests. Different sources might label a given pixel *beach*, *open space*, and *natural*. A human mapping analyst would, in this case, be able to apply a lifetime of experience to resolve the paradox by placing objects in a knowledge hierarchy, and a rule-based expert system could be constructed to codify this knowledge. Alternatively, an analyst might be faced with complex or unfamiliar labels, and the structure of label relationships might vary from one region to the next. Current research has shown how an ARTMAP neural network can derive hierarchical knowledge structures from apparently inconsistent training data. The system learns that disparate pixels map to the output class *beach*; but, if similar or identical pixels are later labeled *open space* or *natural*, the system learns to associate multiple classes with a given input. The overall pattern of distributed predictions reveals a knowledge hierarchy which guides the production of consistently layered maps. Even though no inter-class information is provided with the training inputs, the system readily derives knowledge of relationship rules, confidence estimates, and hierarchical structures. CELEST projects are now extending ARTMAP knowledge discovery methods beyond the image domain. For example, a pilot study has created a hypothetical set of relationships among protease inhibitors, based on resistance patterns from genome sequences of HIV patients.