Title
Representing Cognitive Maps in Parallel Networks

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 6(0)

Author
Zipser, David

Publication Date
1984

Peer reviewed

REPRESENTING COGNITIVE MAPS IN PARALLEL NETWORKS

David Zipser
Institute for Cognitive Science
University of California, San Diego

Maps are an important part of systems which represent world knowledge. They are used to provide information for guiding movement through the environment or for answering queries about the location of objects. The number of facts about the spatial relationships between objects, which maps must provide, is too great to be stored explicitly. What is needed is a representation that stores basic information in a manner that allows the use of inference to derive the appropriate facts when required. For example, in order to answer such queries as "am I headed towards home?", "how do I get from here to Carnegie Hall?", McDermott (1980) developed a spatial representation called fuzzy maps which has proven very useful in the symbol processing, serial computer environment. In this paper I will describe a different way to represent spatial knowledge called view-maps. The main motivation for developing view-maps was biological plausibility. I wanted a representation which was amenable to implementation in parallel networks of neuron-like units and whose properties corresponded to what we know about the neurobiology and psychology of spatial location. For a variety of reasons fuzzy maps do not suit this purpose.

To get a feel for what view-maps do, consider the problem of getting from your current location to an unseen goal, such as home. Assume that while you cannot see your home from your current location, you can see some familiar landmarks. Now suppose that at a previous time you had recorded in memory the location of your home relative to these landmarks. You can use this remembered information to locate your home and view-maps provide a mechanism to accomplish this. Roughly speaking, view-maps store a set of individual views, (Kuipers 1983) together with the locations relative to these views, of goals such as home, work, etc. When you want to go to one of these goals, you compare your current view with the set of stored views in the view-map. There will, in general, be no exact match, because you won't be at a precise point where you recorded a view. However, there will be similarity between your current view and views stored at nearby locations. These nearby views will differ from your current view only quantitatively in the values which indicate landmark location. It is possible to use these stored values of landmark location together with the values obtained from your current location to derive a transformation that will map the location of any object from the reference frame of a stored view to your current reference frame. By applying this transformation to the stored locations of a goal, you can find it in your current frame. View-maps use this general principle to guide an observer through the world.

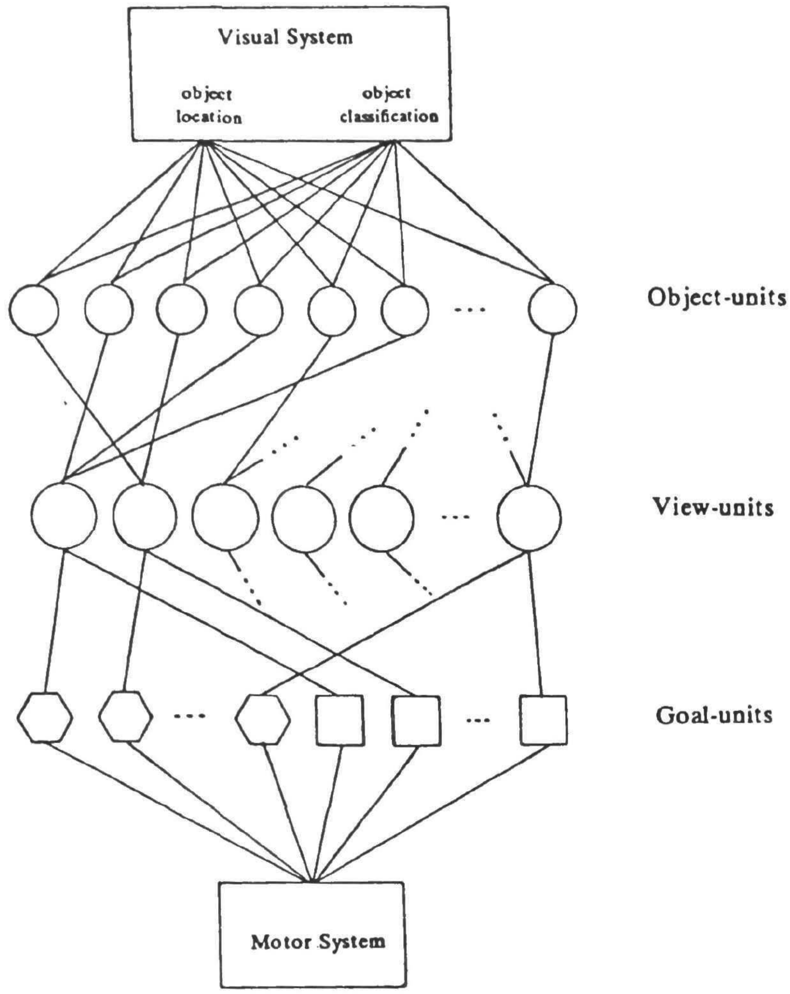
View-maps must first be built up by recording the location of goals at many viewpoints in the world. Then this information can be accessed by comparing the current view with the stored views to locate a desired goal. If an appropriate stored view does not exist in your view-map, it does not mean that you are lost. Often it is still possible to find a chain of views that overlap by several landmarks, so that a transformation can be generated to map the goal from a stored view to the current view, even though these two views do not contain the same landmarks. In this short paper I cannot discuss all the features of view-maps and the networks

---

This research was supported by a grant from the System Development Foundation.

that implement them. Rather I will describe a simple network which implements some key features of view-maps, such as view recognition and goal location.

In practice, it is often possible to locate goals without actually computing transformations or even reading out the stored locations of the visible landmarks. The network shown below was developed to demonstrate how this can be done.



The first layer consists of object-units which receive input from the sensory system and become active when a particular landmark appears at a specific location in the current reference frame of the viewer. The second layer consists of view-units, each of which receive input from several object-units. View-units recognize the fact that the observer is in the vicinity of a particular place. The activity of the view units can be used to answer such questions as "am I at the site where I buried that stuff?" The output of the view units are connected to a layer of goal-units which, when activated, can tell the motor system how to get from the viewpoint of a view-unit to the goal they represent, i.e., "how do I get home from here?" There must be a separate goal-unit for each goal at each viewpoint. Sets of goal-units representing different goals are shown with different shapes.

Each class of units in this network has an associated function which determines how its output activity varies with its inputs. For example, the activity function for the view-units is just the thresholded sum of their input activities. The value of the threshold is chosen so that several landmarks must be recognized in approximately the expected locations for a view-unit to become active.

A more complex function is required for object-nodes because they must recognize the presence of a landmark and the degree to which its current location matches the expected location. This function should have a maximum when the viewer is in exactly the expected location and then fall off gradually as the viewer moves from this location. The function should be zero at all times when the expected object is not being viewed at all. A matching function I have used extensively, is:

$$\text{Object-unit activity} \begin{cases} = \exp\text{-} \left[(\text{current object location - expected object location})^2/\sigma^2\right] \\ ;\text{ if object detected.} \\ = 0;\text{ if object not detected.} \end{cases}$$

The values of the threshold of the view-unit activation function together with the σ of the object-function determine the size of the region in space in which a particular view-unit will become active. If this region is small the observer will have accurate knowledge of position but it will be available over a small area. If the region is large, less accurate information will be available over a wider area.

I have used computer simulations (Rabin & Zipser, 1983) in which the movement of an observer is guided by networks of the type shown above to investigate how view-maps deal with problems such as recognizing a location as previously visited and getting from the current location to a goal. The first of these problems requires quantitative matching of the current view to all previously recorded views. Because of the structure of the network used, this match occurs in parallel so that each view-unit always indicates, by its output activity, the degree to which the current view matches the view at its viewpoint. Of course, at any particular location most view-units are inactive. The viewer determines if return to the desired location has occurred by sensing the activity of the appropriate view-unit. This is how, for example, the location of previously buried stuff can be located. The computer simulation demonstrated that there was a "place-field" around the location represented by each view-unit in which the activity of the unit increases as the viewer approaches the viewpoint. This makes view-units similar to the spatial field neurons in the hippocampus (Muller, Kubie & Ranck, Jr., 1983).

To answer such questions as "how do I get home from here?" requires that the location of a landmark not currently visible, i.e., home, be determined. If view-units which connect to goal-units for the desired unseen location have been recorded fairly evenly over the environment, then at any location the viewer will activate several view-units that refer to the goal. An estimate of the location of the goal can be made by forming an average of the location stored in each of these goal-units, weighted by the activity of its connected view-unit. Computer simulations showed that the use of this weighted average value of the goal location is an effective way to determine where the goal is when the viewer is far away from the goal. However, when the observer gets very close to the goal, there are serious difficulties because some of the viewpoints used are beyond the goal and thus give the wrong sign to goal direction. When the observer reaches the vicinity of the goal, motion becomes erratic. However, sooner or later the simulated observer generally finds a path into the goal. Behavior of the sort that occurred in these simulations is not too unreasonable since, when far away from a goal, an observer cannot see it and is then forced to use landmarks. When the observer gets close enough to actually see the goal, landmarks are no longer needed.

A more formal analysis of view-maps, not presented here, has been carried out which shows that they can serve as a robust representation of the spatial organization of objects. Several important issues have been analyzed to varying degrees and also studied with computer simulation (Zipser, 1983). Among these are: what is a landmark? How are object-units, goal-units and view-units learned? What features are used by the sensory system to localize landmarks? How is it possible to get to a goal whose location is not known to any view-unit currently active?

The object-unit to view-unit hierarchy provides a very general mechanism for representing the spatial location of objects. It has been shown here how it can be used to construct view-maps which represent the spatial organization of landmarks in the environment. It can also be used to represent the locations of features within a single object and in this way they may be useful for the recognition of objects. View-units can also be considered as schemata in which information about location is given major prominence (Brewer & Treyens, 1981). Viewed as schemata or frames, view-units can be thought of as having additional information besides that already discussed. For example, view-units might connect to units indicating the suitability of place fields, for example, dangerous or safe. The activation of a particular view-unit as an observer moves through the environment would then immediately give access to information about the desirability of remaining at that location. In general, any variable quantity connected to view-units would be accessible when the observer returned to a location in which the view-unit was activated. This corresponds to the common experience of recalling events or even thoughts that occurred at specific locations (Nigro & Neisser, 1983).

REFERENCES

Brewer W. F., & Treyens, J. C. (1981). Role of schemata in memory for places. *Cognitive Psychology, 13,* 207-230.

Kuipers, B. (1983). Modeling human knowledge of routes: Partial knowledge and individual variation. *Proceedings of the National Conference on Artificial Intelligence, August 22-26, 1983 (AAAI-83)* (pp. 216-219). Los Altos, CA: William Kaufmann, Inc.

McDermott, D. (1980). *Spatial inferences with ground, metric formulas on simple objects* (Res. Rep. No. 173). New Haven: Yale University, Department of Computer Science.

Muller, R. U., Kubie, J. L., & Rank, J. B., Jr. (1983). High resolution mapping of the 'spatial' fields of hippocampal neurons in the freely moving rat. *Society for Neuroscience* Abstract Number 191.4.

Nigro, G., & Neisser, V. (1983). Point of view in personal memories. *Cognitive Psychology, 15,* 467-482.

Rabin, D., & Zipser, D. (in preparation). *P3 - Parallel process programmer*. (Tech. Rep.). La Jolla: University of California, San Diego, Institute for Cognitive Science.

Zipser, D. (1983). *The representation of maps* (ICS Rep. No. 8303). La Jolla: University of California, San Diego, Institute for Cognitive Science.