How might an artificial neural network represent metric space?

Patricia M. Boechler and Michael R.W. Dawson

Department of Psychology, University of Alberta

ABSTRACT

An artificial neural network was trained to rate the distances between pairs of cities on the map of Alberta, given only place names as input. The issue of interest was the nature of the representations developed by the network's hidden units after it successfully learned to make the desired responses. A number of different analyses indicated that the network's hidden units had developed metric representations of space. The manner in which this network completed this task has implications for the representation of spatial relationships in biological systems, specifically, how place cells in the hippocampus may represent spatial information.

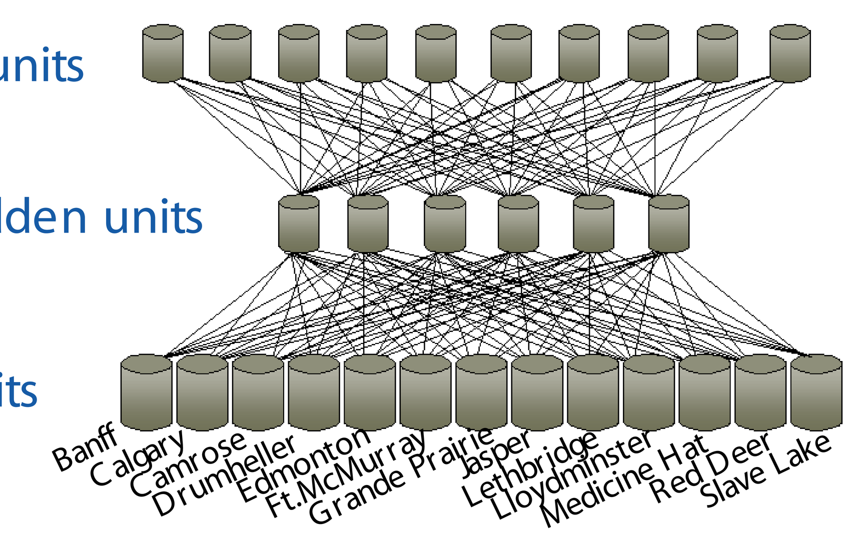
METHODS

From a map of Alberta, we determined the shortest distance in kilometres between each pair of locations. These distances were then converted into ratings.

	BANFF	CALGARY	CAMROSE	DRUMHELLER	EDMONTON	FORT MCMURRAY	GRANDE PRAIRIE	JASPER	LETHBRIDGE	LLOYDMINSTER	MEDICINE HAT	RED DEER	SLAVE LAKE
BANFF	0	2	4	3	5	9	7	3	4	7	5	3	7
CALGARY	2	0	3	2	3	8	8	5	3	6	3	2	6
CAMROSE	4	3	0	2	1	6	6	5	5	3	5	2	4
DRUMHELLER	3	2	2	0	3	8	8	6	3	5	3	2	6
EDMONTON	5	3	1	3	0	5	5	4	6	3	6	2	3
FT. MCMURRAY	9	8	6	8	5	0	8	8	10	6	10	6	5
GR. PRAIRIE	7	8	6	8	5	8	0	4	10	8	10	6	4
JASPER	3	5	5	6	4	8	4	0	7	7	8	5	5
LETHBRIDGE	4	3	5	3	6	10	10	7	0	7	2	4	8
LLOYDMINSTER	7	6	3	5	3	6	8	7	7	0	5	4	5
MEDICINE HAT	5	3	5	3	6	10	10	8	2	5	0	5	8
RED DEER	3	2	2	2	2	6	6	5	4	4	5	0	4
SLAVE LAKE	7	6	4	6	3	5	4	5	8	5	8	4	0

NETWORK ARCHITECTURE

The network was a network of value units that had 10 output units, 6 hidden units and 13 input units. Each input represented one of the 13 place names. Pairs of place names were presented as stimuli by turning two of the input units on (activated at a value of 1). All of the other input units were turned off (activated by a value of 0). The input units themselves did not provide any metric information to the network.



RESULTS

The hidden units of the network did indeed develop metric representations of space as indicated by the following analyses:

- 2 dimensional multi-dimensionsal scaling analyses accounted for almost all of the variance in the activation matrix for each hidden unit.

- assuming each hidden unit occupies its own position on the map, the position for each hidden unit produces a high correlation between the distances , the origin of the MDS plot and the Gaussian-transformed distances between cities and the location of the hidden unit.

HIDDEN UNIT	Correlation between Transformed Map Distances and Distances from MDS Solution Origin	HIDDEN UNIT LATITUDE	HIDDEN UNIT LONGITUDE	GAUSSIAN MEAN
HO	-0.9288	50.09979	114.2014	-0.52913
H1	0.656923	49.82874	112.38	0.583262
H2	0.794037	53.82825	116.5319	0.528801
H3	-0.77478	49.95480	113.7699	-0.12785
H4	-0.82919	53.63924	113.439	-0.6091
H5	-0.54974	51.16697	115.567	1.7

- for each hidden unit, one could find a vector that passed through the MDS plot such that when city locations are projected onto this vector, there were very high correlations between these projections and the connection weights feeding into the hidden unit.

HIDDEN UNIT	Correlation between Projections and Connection Weights	X- Coordinate of Unit Vector	Y - Coordinate of Unit Vector
HO	0.947	0.681	-0.732
H1	0.857	0.935	-0.354
H2	0.877	-0.925	0.380
H3	0.905	0.996	-0.087
H4	0.954	0.996	-0.085
H5	0.963	0.992	0.124

- assuming each hidden unit occupies its own position on the map, one could find a location for each hidden unit that produced a high correlation between the connection weights feeding into the hidden unit and the distances on the map between cities and the position of the hidden unit.

HIDDEN UNIT	HIDDEN UNIT LATITUDE	HIDDEN UNIT LONGITUDE	GAUSSIAN MEAN	Correlation between Map Distances and Incoming Weights	Correlation between Transformed Map Distances and Distances from MDS Solution Origin
HO	51.44895	113.8842	0.596687	0.872964	0.874018
H1	50.83889	113.6293	0.650555	0.586887	0.606727
H2	50.87556	117.6973	0.563113	0.717306	0.652394
H3	53.39033	115.0497	0.567472	-0.54432	-0.52623
H4	53.91175	113.4247	-0.5956	0.790458	-0.82248
H5	51.167	115.567	0.233109	-0.48462	0.449098

CONCLUSIONS

Although this artificial neural network was not explicitly given metric input, after training, it produced metric responses. Our analyses indicate that the connection weights to the hidden units are responsible for this metric behavior. Due to coarse coding, each hidden unit itself does not produce activity that reflects the configuration of the map of Alberta but the activation of the entire network produces an accurate simulation of the place positions. This demonstrates that a system with only six processing units can represent the spatial relationships between thirteen different landmarks. The implication for biological systems is that they need not be organized topographically. Instead, locations of landmarks in the environment could be represented as a pattern of activity distributed over a number of different place cells.

NSERC CRSNG

This research was supported by NSERC Research Grant A2038 awarded to Dr. M.R.W. Dawson.

Biological Computation Project
Department of Psychology, University of Alberta