Sparse neural representation for semantic indexing

Peter Foldiak
School of Psychology, University of St Andrews, St Andrews KY16 9JP, Scotland, U.K.
Peter.Foldiak@st-andrews.ac.uk
http://psy.st-and.ac.uk/foldiak

Abstract

1. Traditional models think of concepts as nodes and associations as links between nodes. Distributed neural representations suggest a model of relationships between concepts where explicit links are not even necessary. Similarity and multiple inheritance can be expressed as vector similarity. High dimensional multiple categorisation is implicit in the neural representation itself.

2. Sparse representations (low fraction of active neurons) are found at different levels of the sensory system, and can be assumed to be present in semantic representations. It is efficient to consider a sparse activity vector as a mathematical set: the set of active units.

3. Semantic relationships can be represented implicitly by different kinds of overlap between these sets of ‘features’, rather than with direct links. Set algebra can be used to combine representations for expressing and indexing new concepts or items. The advantages of neural representations can be rescued for practical applications where neural network learning would be currently unfeasible.

4. Efficient and precise ‘concept’ search can be based on sets.

Most of today’s catalogues and directories have a new structure where each item is assigned to a single category corresponding to a branch of the tree. Traditional cognitive models of the relationships between concepts are usually also captured in graphs (often also trees, e.g. semantic nets) where the nodes are the concepts and the links express relationships. The rich and multilayered structure of relationships between concepts is not encoded well by those models. This results in poor retrieval performance even in computerised systems. Distributed neural representations studied in both artificial and biological neural networks, however, suggest a different model.

Neural representations consist of a large set of neurons, only some of which are active when representing an item. For each distinguishable item, a different pattern of activity can be observed. This pattern can be described by a vector, whose components correspond to the activities of the neurons. In a high-level, semantic neural representation, two similar items would be represented at two similar, i.e. highly overlapping patterns of activity. But neural representations encode more than just the similarity of items. As each neuron, or subset of neurons, encodes one particular aspect or feature of the represented item, the pattern of activity corresponding to two items also reveals the nature of the relationship between the items. Note that the relationships between items are represented implicitly in the relationships of the patterns of activity, and need not be made as explicit links as in the graph models.

Representations for “cat” and ‘dog’. The amount of similarity is represented by amount of overlap (3 of the 5 active units). The common active units (4, 6, 16) correspond to “pet”, and describe the nature of the relationship.

A binary activity vector can be considered equivalently as a set of active units. If units (or subsets of them) stand for features, this is equivalent to a representation of an item by a set of features (or ‘microfeatures’).

If items are represented as mathematical sets, their relationships can be described as overlapping sets and visualised with Venn diagrams. Venn diagrams show sets as closed curves and their elements as items inside the curves graphically.

Hierarchical inheritance relationships can be mapped into feature sets. “animal” is a subset of the “dog” feature set, “living thing” is a further subset of “animal”.

Note the duality between the ‘sets of items’ and the ‘sets of features’. In the traditional ‘sets of things’ model, each concept is a set of the things that belong to that concept, so ‘animal’ is a subset of ‘living thing’ (as only some living things are animals). In the ‘sets of features’ model, the most general concept corresponds to the smallest set, with each specialisation adding more features to form super sets.

One of the drawbacks of semantic nets is the difficulty of inheritance. Representing all possible (combinatorially large number of) associations explicitly is impractical, ‘spreading activation’ was proposed to work out indirect relationships. The assumption was that association is transitive, i.e. if “A associates B” and “B associates C”, then “A associates C”. While this is sometimes true (e.g. “associate(chemistry, fire)” and “associate(chemistry, fire engine)” – because they are all red), it is not generally true: “associate(chemistry, fire engine)” and “associate(chemistry, lemon)” but NOT “associate(chemistry, lemon)”. This latter is not transitive because the association between chemistry and fire engine is along a different aspect (colour) than that between chemistry and lemon (function). Spreading activation fails in such cases.

Feature sets deal with transitivity appropriately. They represent not just association, but also the nature of the association. The association between chemistry and fire engine, and that between chemistry and lemon, correctly, does not imply that lemon and fire engine are associated. The overlap between the first pair involves colour-related features, while that between the second pair function-related features.

If concepts are represented as sets, set operations (set algebra) can be used to combine concepts to:
- define new concepts,
- express queries.

Oversimplified example:

<table>
<thead>
<tr>
<th>definitions:</th>
</tr>
</thead>
<tbody>
<tr>
<td>thing 1 =</td>
</tr>
<tr>
<td>penguin := (bird ( \land ) fly) ( \land ) (antarctic ( \lor ) place)</td>
</tr>
<tr>
<td>query:</td>
</tr>
<tr>
<td>cold ( \lor ) animal ( \rightarrow ) penguin</td>
</tr>
</tbody>
</table>

Comments:
- “Y is a set with a single newly generated feature “animal” is a set of two features (one in thing, one head)
- penguin contains the bird features except the flying related features, plus the antarctic specific features (but not place features in antarctic, as penguin is not a place).
- Question: “What is an animal related to cold?”
- answer: penguin

The query performs a search for the smallest stored sets of maximal overlap with the query set efficiently. The result can be depicted graphically in 2D. The best match is clue to origins of coordinate system (UD). Vertical axis is general/proficiency and (more general concepts above, more specific below). Horizontal axis is distance (more similar is placed closer to the left).

Example shows analogy:
“What is the equivalent of Persian in dogs?”

query: persian \( \land \) cat \( \lor \) dog
answer: Collie

Acknowledgements

The project was supported by a Leverhulme Trust Grant (SG3).