Bilingual Processing On Neural Networks

Marc Tucker, Jie Li

Motivation

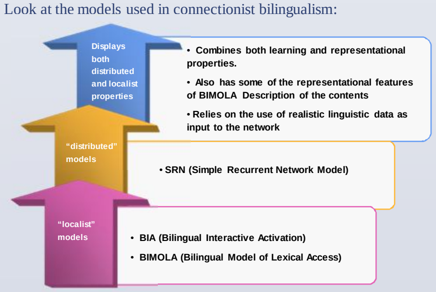
Neural networks **have significantly influenced research in the cognitive sciences in the last fifteen years.**

Language**, as one of the most important human cognitive components, has received in-depth treatments since the beginning of connectionist research.**

Objectives

The objective is to construct Artificial Neural Networks (ANN) which are used to acquire bilingual languages simultaneously. More specifically, the system is undergoing the process of bilingual acquisition by the use of inductive reasoning. We try to use this to model the cognitive process of the human mind in the event of acquiring two or more languages.

Look at the models used in connectionist bilingualism:



As you can see, we would like to build the model which represents both distributed and localist properties. To use this model, the main objective is to find out the way to obtain both the lexical and semantic representations of the words in two languages in the same way.

For the semantic representations, we need to prepare the realistic linguistic data to be the input. After that, we should find a way to generate the semantic representations for the input data and to see whether the words meaning the same in two languages have similar representation values.

After we have both representations, we learn the bilingual languages through the artificial neutral work and to see how the lexical and semantic representations in two languages will be clustered in the training process. According to the results, we get the conclusion of how the bilingual acquisition supposed to be when a human is in bilingual environment.

Finally, we want to test our program to make sure about the correctness of our conclusion and consider how to improve the program to get more information from the model we build.

Materials and Methods

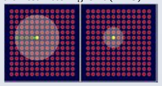
Materials:

- CHILDES – Child Language Data Exchange System
  From this system, we get the words used most often to generate the input sentences for generating the semantic representations.
- Phonological representations – Syllable based template (MacWhinney and Leinbach)
  Considering that when human acquire the languages, even we do not know how to spell the words, we can still understand and communicate with others, we decide to use the phonological representations rather than the orthographic representations to be the lexical representations.
  In our representation, we use two groups of CVVCCVVC and t1, t2 for the tones to represent the words both in English and Chinese.
  Each C or V is represented by a set of 5 feature units, and the feature values.
- Semantic representations – FGREP in the case-role assignment task and stand for distinct meanings.
  We know that in different languages, there are different sentence structures for the same meaning. However, the components (cases: subject, verb, object) in the sentences are similar.
  We take the original right order sentences and the case-roles sentences as input to generate the semantic representations.
- Input pairs – To help with the categories of different words.
  Wolf – predatr
  Sheep – prey

Methods:

Procedure:

- Setup lattice for each respective map
  - All weights are initialized 0 < w < 1
    - Calculate the Euclidean distance between each node's weight vector and the current input vector.
      - The node with the shortest distance will be the Best Matching Unit (BMU)
- Determine the Best Matching Unit's Local Neighborhood.
  - Calculate the radius of the neighborhood
  - Pythagorean Theorem
- Over time, the area of the neighborhood shrinks over time.
- Exponential decay function



- Eventually, the neighborhood will shrink to the size of just one node... the BMU.
- Every node within the BMU's neighborhood has its weight vector adjusted by:

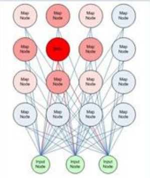
$$W(t+1) = W(t) + \Theta(t)L(t)(V(t) - W(t))$$
$$t = 1,2,3\ldots$$

$\Theta(t)$ given by: $\Theta(t) = \exp\left(-\frac{dist^2}{2\sigma^2(t)}\right) \quad t = 1,2,3\ldots$

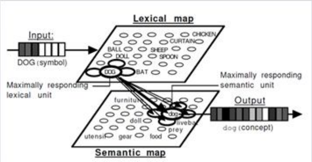
- Both maps now display localized patterns of activity.
- Hebbian Learning
  - The weights between active units are increased proportional to their activity.

$$\Delta l_{ij,kh} = \Delta s_{kh,ij} = \alpha(t)\eta_{L,ij}\eta_{S,kh}$$

SOM Example:



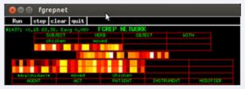
Training process:



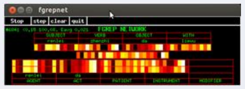
Results

Semantic representations generated test:

For English:



For Chinese:



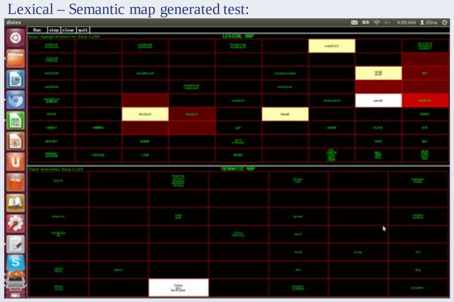
The representations for dog and gou (狗）

dog 0.725955 0.175367 0.172209 0.373092 0.330487 0.000900 0.570818 0.769120 0.090259 0.218430 0.342084 0.750827

gou 0.281397 0.744260 0.446485 0.036788 0.588638 0.585699 0.368603 0.611384 0.408005 0.708740 0.416559 0.840438

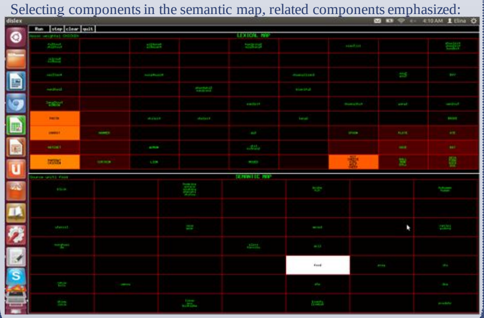
Even they have the same meaning, the English representation is very different from Chinese representation.

Lexical – Semantic map generated test:



Selecting components in the lexical map, related components emphasized:

Selecting components in the semantic map, related components emphasized:



Conclusions

As you can see from the results, for the semantic representations, even the words with the same meaning in two languages, they have very different representations, which shows that when we acquire two languages at the same time, for the different words of same meaning in different languages, we can still distinguish them through the different usage. What is more, see that in the semantic map, the English words and Chinese words are not combined.

For lexical representations, we see that in most condition, the words in the same language are clustered together except some special cases such as "window" & "lang2tou0". We think this is reasonable because imagine when we hear a word in language is very similar to that in another language, we will not be surprised, and possibly we associate them.

Another big founding is that look care about the distribution of the words in both English and Chinese in both lexical and semantic representations, we can see that there is a general separation between them, so we more support that in our brain, we have different areas to deal with different languages.

References

Ping Li and Igor Farkas. *A Self-Organizing Connectionist Model of Bilingual Processing.* University of Tichmond.

Risto Miikkulainen. *Dyalexic and Category-Specific Aphasic Impairments in a Self-Organizing Feature Map Model of the Lexicon.* Department of Computer Sciences, The university of Texas at Austin.

Risto Pekka Miikkulainen. *DISCERN: A Distributed Artificial Neural Network Model of Script Processing and Memory.* University of California.

Jordi Poveda and Alfredo Vellido. *Neural Network Models for Language Acquisition: A Brief Survey.* Technical University of Catalonia

Wei Xu. *Can Artificial Neural Networks Learn Language Models?* Carnegie Mellon University.

Claire D'Este, Michael Towsey, Joachim Diederich. *Sparsely-connected recurrent neural networks for natural language learning.* Machine Learning Research Centre, Queensland University of Technology.

Altarriba, J., Kroll, J.F, Sholl, A., & Rayner, K.. *The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times.* Memory & Cognition, 24, 477-492.

Contact

Jie Li:
Cell: 352-283-2703
Email: jili@cise.ufl.edu

Marc Tucker
Cell: 718-916-8169
Email: mtucker@ufl.edu

Welcome to contact us if you want to discuss!