Latent Semantic Analysis in Word Meaning Representation

Paper 3

Zhu Zhiyuan

July 28th, 2017

Introduction

Words play an important role in people’s lives. A normal person probably comes into contact with thousands words in a normal day.

Every day people are searching in the huge word bank in their mind. It is reasonable to suppose that words are in particular order in people's mind to make them easy to retrieve and use. Therefore the question is that in what format that words displayed in people's brains.

Broadly speaking, we have two approaches to the representation of meaning: *mental lexicon* approaches that list word meanings like a huge dictionary in the mind, and *generative* approaches that construct meaning out of certain elements (LSA vectors) according to some rules.

Human live with word

Few people regard words with the awe, since we use them all the time without a second of thinking. It is because words live in our minds silently. Just as Robert Mazy wrote in poem “Words”**[1]** :

“ We thought a day and night of steady rain was plenty, but it's falling again, downright tireless . . .

Much like words

But words don't fall exactly; they hang in there

In the heaven of language, immune to gravity

If not to time, entering your mind

From no direction, traveling no distance at all,

And with rainy persistence tease from the spread earth

So many ravishing scents . . .”

Words Are Stored In Order In Human’s Brains

Words cannot be heaped up randomly in the mind for two reasons.

First, there are so many of them. The number of words which an educated adult native speaker of English knows, and can use, is more than 50,000. **[2]**

Second, words can be found extremely fast in human’s minds. In a normal speech speed is six syllables a second, making three or more words in such a short period of time. Words can be located very fast. Most of times they can be pinpointed in a split second.

Third random facts and figures are extremely difficult to remember, but enormous quantities of data can be remembered and utilized, as long as they are well organized. Admittedly, some words are hard to seek out. Almost everybody has had the experience of not being able to think of the particular word they want, and they are sure they know it. This can be illustrated by a fictional but quite realistic dialogue from Douglas Adams's science-fiction satire Life, the Universe and Everything **[3]** :

Arthur shook his head in a sudden access of emotion and bewilderment.

"I haven't seen anyone for years," he said, "not anyone. I can hardly even remember how to speak. I keep forgetting words. I practice you see. I practice by talking to ... talking to ... what are those things people think you're mad if you talk to? Like George the Third."

"Kings?" suggested Ford.

"No, no," said Arthur. "The things he used to talk to. We're surrounded by them for heaven's sake. I've planted hundreds myself. They all died. Trees! I practice by talking to trees."

Arthur cannot remember the word "trees". But when he struggles to retrieve it, and he uses nearly 50 other different words without effort to describe it . Such fast and efficient retrieval must be based on a structured system, not on random system.

Therefore the large number of words known by humans and the speed with which they can be located lead the existence of a highly organized word structure in human minds.

Methods Of Word Storing

Words have different meanings, like bark in dogs bark and bark of a tree, that are unrelated semantically. It is impossible to list out every single meanings of each words like what a dictionary does in human mind. And it is also can be proved false with a simple experiment that words are not arranged alphabetically like dictionary in human minds. People occasionally make mistakes when they speak, selecting one word in error for another. If the words were organized in alphabetical order, one might expect speakers to accidentally pick an adjacent word in alphabet when making these kinds of errors. So, in error for the word “guitar” one might expect someone to accidentally pick guinea or gulf, all words which are near neighbors in standard dictionaries. But mistakes of this type are quite uncommon. If words are unlikely to be arranged in human minds as distinctly as dictionaries do, then the alternative is to define meaning implicitly, by its relations to other words. The core meaning of each word in a language is well defined in dictionary, but is modified in each context.**[4]** Word difference meanings emerge when words are used in different contexts. Indeed, every context generates its own word sense. The differences between the contextual meanings of a word may be small or large, but they do always exist. The word meaning without context is nothing but an abstraction and useless.

LSA is a relation among words. In such a relational system, one cannot talk about the meaning of a word in isolation; words have meaning only by virtue of their relations to other words—meaning is a property of the system as a whole. In LSA, the meaning of a word is situated with respect to all other words in the semantic space **[5]** .

LSA computes a context-free vector for the meaning of each word. Different meanings or senses of a word are all put together in this vector. The meaning of *bark*, in LSA, is right between the *tree* and the *dog*. Therefore, LSA is a model of the associative of semantics. The idea is simple in principle: Take the context-free word vector from LSA and embed it into its semantic context, generating a spreading activation network that selects appropriate information of context, calculating a appropriate word meaning in context. **[4]**

Specifically, suppose there is a simple sentences of the form Argument-Predicate (e.g., Noun-Verb, Noun-is-Adjective, or $Noun_1$-is-$Noun_2$). Let **A** be the vector of the argument, and **P** the vector of the predicate. What LSA do is to modify a sense of **P** that is appropriate for **A**. A network of words is construed by LSA, consisting of A, P, and the *n* (e.g., *n* = 100) closest neighbors (words) of **P**. Connect all the neighbors of P to both **P** and **A** by a link whose strength is equal to the cosine between the neighbor and **P** or **A**, respectively, measuring their relatedness to **A** and **P**. Let {S} be the set of all items in the semantic space except for P and A. The terms I in {S} can be arranged in a semantic neighborhood around P. Relatedness to P (the cosine between each item and P) determines how close or far a neighbor they are. Let cos(P, I) be the cosine between P and I in {S}. Furthermore, let cos(A, I) be the cosine between A and item I in {S}. The strength s(A,I) of these links is codetermined by how closely related they are to both A and P:

s(A,I) = f(cos(A,I), cos(P,I)). [6]

Activation spread in such a network will tend to concentrate in **A** and **P** and in those neighbors of **P** that have relatively strong links to **A**. Take the most strongly activated neighbors of **P** and combine their vectors with **P** to have a sense of **P** and in the context of **A**, $\mathbf{P}_A$. Specifically, the vector computed by the predication procedure is the weighted average most activated items in the net described above.

To make the predication model idea more clear, here is an example : **The horse runs**. What LSA do is to find a vector that represents the
meaning of *run* in the context of *horse runs*. What predication does it to let the argument *horse* select among the neighbors of the predicate *run* those that are related to *horse.* **[6]**

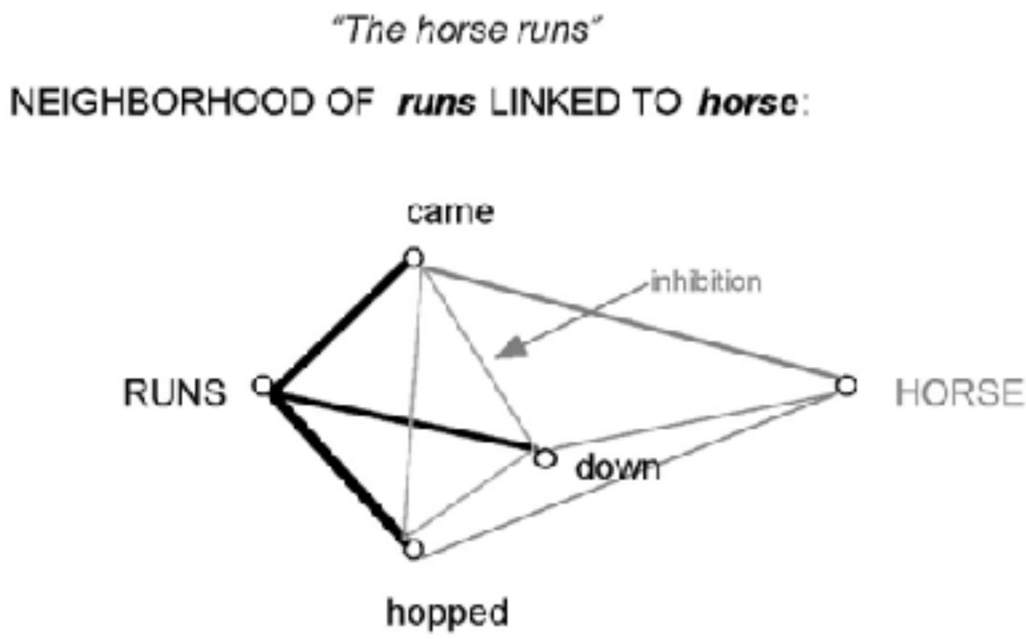
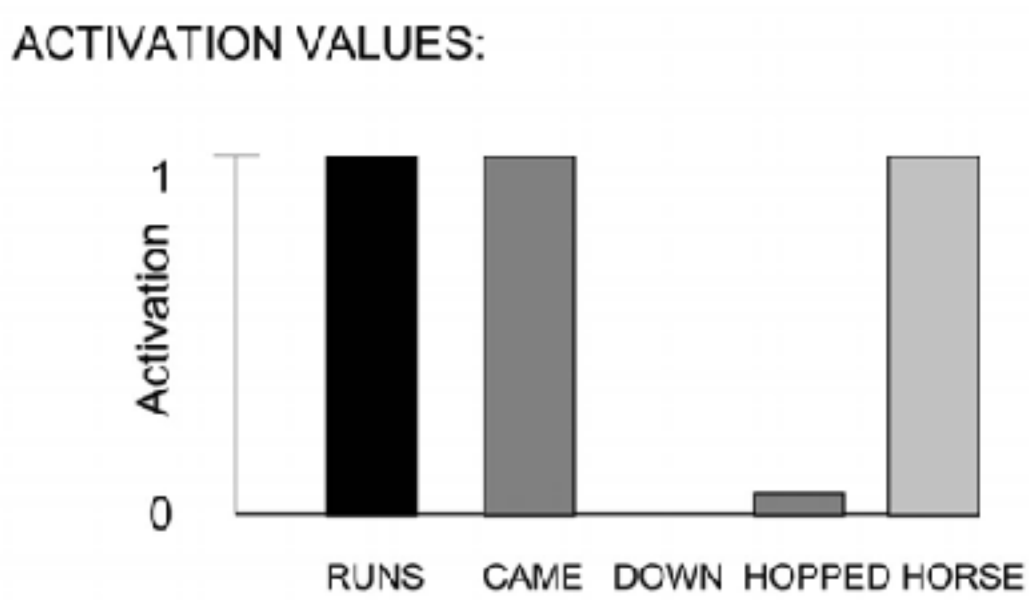


TABLE 1

In table 1 , "run" is split into three neighbor words : came, down, and hppped, but in actual the number would be much more higher. All neighbors are related to *run*, by definition, but only one, come, is related to *horse*. Thus, spreading activation in this network will result in the asymptotic activation values shown in table 2.

TABLE 2



Predication then computes the vector for *run* in the context of *horse* as the vector sum of *run* and *come* weighted by their activation values. (Table here is simplified)

Now, consider the sentence *The color runs*. Table 3 and 4 illustrate how predication calculates the vector for *run* in the context of *color runs*. The result is different this time, and *down* is the neighbor of *run* who modifies its meaning.**[6]**

ACTIVATION VALUES: **TABLE 3**

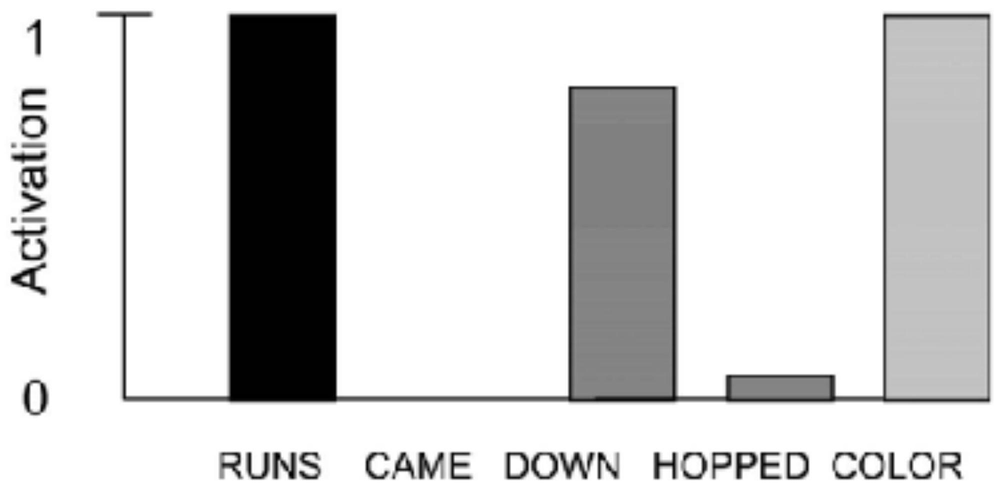
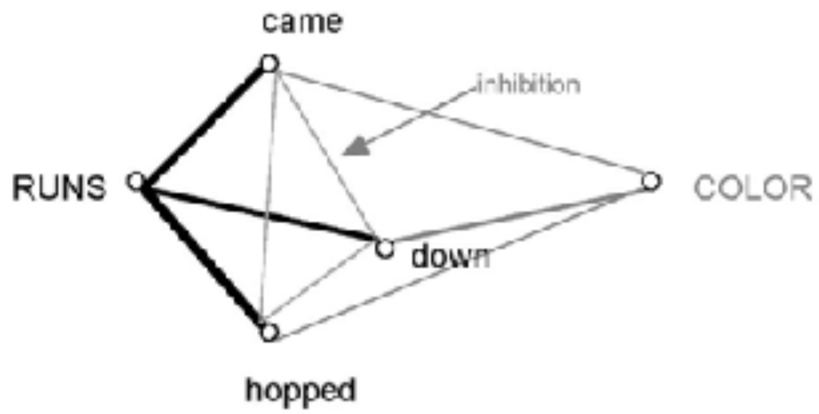


TABLE 4

"The color runs"

NEIGHBORHOOD OF ***runs*** LINKED TO ***color***.



How much the meaning of a word changes in context greatly depends on exactly how word and context are related. For example, consider *house yard* the sense of *house* that is constructed in the context of yard. According to predication, there is almost no change at all: First, the closest neighbors of *house* are the same as for *yard*, and therefore the sense of *house* in the context of *yard* remains pretty much unchanged. However, consider *house of representatives*. If house is predicated about *representatives*, its meaning is no longer the same, because *representatives* selects items far from the neighborhood of *house.*

Conclusion

Humans acquire knowledge in almost the same way as LSA does: by keeping track of the events in their environment and getting from it a high-dimensional semantic space. This semantic space serves as process of cognition . LSA can provide a good model of human cognitive processes.

References

[1] Robert Mazy poem Words

[2] Vocabulary size Lexical facts, Johnson, May 29th 2013, by R.L.G. | NEW YORK

[3] Douglas Adams, Life, the Universe and Everything

[4] Kintsch, W. (1998). Comprehension: A paradigm for cognition. New York, NY: Cambridge University Press.

[5] The Construction of Meaning Walter Kintsch & Praful Mangalath

Walter Kintsch & Praful Mangalath ... (Klein & Murphy, 2001).

[6] Predication 1 Predication Walter Kintsch University of Colorado W. Kintsch

Institute of Cognitive Science