Project

~~Meo~~ definition of a word refers to a ~~pt~~ an expression of multiple words, which have the same meaning.

Each word in the definition has its own definition.

Often meaning of one word consists of ~~an~~ a composition of multiple concepts:

book - [collection ~~of~~, sheets, papers, bound, together, hinge, one, edge, containing, printed, ~~written~~, material, pictures].
write

collection - [set, items, amount, material, procured, gathered, together].

~~sheets - [thin, bed, cloth, used, covering, mattress, layer, sleeper]~~

sheet - [Cover, wrap, cloth, paper, other, similar, material].

bound-[bind], boundary

How to tackle polysemy.

cases where the same word with the same sense is shown with multiple definitions.

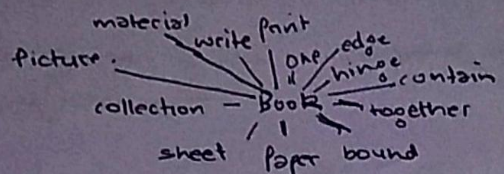
↓

- concatenate all definitions of the word and create one single definition array.
  - ~~implication will be words with many differen'~~
  - we will lose track of the senses.
  - Easier to construct graphs with.
  - This will not effect the structure of the graph - as words with similar meaning or relating to similar concepts with tend to cluster together. (think on the lines of how definitions are ~~cre~~ formulated.

- ~~create a tree~~ -

- ~~Take~~

- Take a ~~ran~~ word - ~~create~~ and start creating graph using the words from the definition. Go layer by layer
-



Book - root.
all words in the definition of words form layers. First layer will be the words that are in the definition of Book, second layer will be the words that are in the definition of words in layer 3.

→ The graph created will have a strongly connected component that contains all the words related to the word 'book'.

→ A random walk ~~on~~ such a graph can provide us with a sequence of words into which the concept of a book can be divided into.

- The subgraph extracted by such a random walk can be used for understanding how language models are used for word selection ~~in~~ in a language production task.

- When we want to retrieve ~~information about a gi~~ a word and know what it approximately means, then starting with ~~this approxi~~ a word that the approximate (similar) meaning defines, the desired word can be reached.

—

- But how will this random walk occur? - ~~How to decide~~ What will be the traversal algorithm?

- Mistakes made while retrieving words by people
    - retrieving words with similar phonological structure
    - retrieving words with almost similar meanings.
    -

- Determine the ~~o~~ maximum, minimum size of loops that occur in the dictionary, these will be required to determine the size of the ~~a~~ sub-graph that should be

    - Look into hypergraphs and how they are traversed. Might get insight into how to compute the meaning association ~~of the~~ between two words.
    - Hypothesis is that if relationships other than the defining relationships, ~~graphs with edges~~ different graphs will be obtained. ~~The F~~ For example,

        The part of speech in a definition can create a differentiation ~~between~~ between the meaning added to a word ~~by the~~ w by the words in its definition d(w).

    - Meaning association can be defined by the total number of incoming edges ~~to~~ from ~~from~~ the w ~~to~~ ~~from~~ to def(w)[i]. This can be used to compute a quantification of meaning association words. ~~To support this form~~

$$\cancel{m = \sum_{i} edo \cdot n_{edge}(def[i], w)}$$

$$\cancel{m_{(w, def(w)[i])} = \frac{N_E(def(w)[i], w)}{\text{indegree of } def(w)[i]}}$$

mc

$$w \longrightarrow def(w)[i]$$

$$M(w, def(w)[i]) = \frac{\left(N_E(w, def[w)[i])\right)^k \cdot K}{\text{outdegree}(w) * \text{indegree}(def(w)[i])}$$

M(w, def(w)[i]) ~~defines~~ ~~th~~ determines ~~how much~~ ~~the concepts defined by the two words are related.~~ meaning dependence.

To suppor this formulation, it can be argued that, Each word is ~~a~~ defined by a number of words.

Each word corresponds to a concept (A concept can be ~~repre~~ defined by a number of words in a language).

Each concept is ~~made up of a n~~ composed of a number of concepts according ~~to~~ to a given set of rules of the language.

Thus a dictionary can be thought of as a document that contains all the possible combination of concepts, each represented by one or more words. ~~This docume.~~

A graph constructed from a sub-set of such a document, ~~will also~~ should also have the relationship of a ~~di~~ the dictionary.

~~The formulation defined uses of the total number of immediate words in the definition neighbourhood of w, how many of them reach def(w)[i] and, of or of the total on, the number of edges)~~

The formulation is an attempt to quantify ~~how related are the concepts defined by each word.~~

~~Numbe.~~

$N_E(w, def(w)[i])$ is to total possible paths ~~between~~ from w to def(w)[i]. This brings the total possible ways w and def(w)(i) are related into the formulation.

Division outdegree(w) brings the ~~total number~~ total ~~number~~ ratio of the total number of concepts used to define w and the number of concepts that that define w and are also used to define def(w)[i] directly or indirectly (through more than one connections).

Division by Indegree(def(w)[i]) bring the ratio of the total number of concepts that def(w)[i] is defining and the total number of concepts that are ~~defining~~ being defined by def(w)(i) are also being used to define w directly or indirectly.

K = {1,2} - need to figure out which will give a more meaningful formulation.

~~[illegible]~~

The formulation computes the ~~ratio of~~ meaning dependence of w on def(w)[i].

~~The formula~~ $N_G(w, def(w)[N])$

→ ~~The ratio obtained~~

There are more factors F that need to be included this, and their contribution is added in K, which also has a normalizing coefficient N

$$K = F/N$$

<u>We need to find F.</u>

Meaning dependence is how much the meaning of w was dependent on def(w)[i]. This can tell us how much the two words are related.

w ←

~~For every~~

All of this hypothesis is based on the fact that meaning is a network property and that concepts are composed of concepts. $s(w_1, w_2)$ can be used to find the relatedness of the concepts defined by w_1 and w_2.

~~The~~ M means the total number of ways a word w_2 can be used to define w_1. This can be a count of the ~~tot~~ fraction of all ways of composition of concepts using w_2, how many ways can be used to composed w_1 from w_2

$$s(w_1, w_2) = M(w_1, w_2) * M(w_2 w_1)$$

~~T and T~~

$$M(w_1, w_2) = \frac{[N_E(w_1, w_2)]^2}{T_1(w_1) \times T_2(w_2)}$$

Total number of paths that can be travesed. that can start from w_1

$T_1(w_1) = \text{Outdegree}(w_1) * F_1$

$F_1 = \text{outdegree}(\text{def}(w_1)[i])\,\text{outdegree}(\text{def}(\text{def}(w_1)[i]))\ldots$

Till d_1 recursions.

Total number of paths that can be traversed that end into w_2

$T_2(w_2) = \text{Indegree}(w_2) * F_2$

$F_2 = \text{indegree}(\text{def}^{-1}(w_2)[i])\,\text{indegree}(\text{def}(\text{def}(w_2)[i]))\ldots$

till d_2 recursions

de

d_1, and d_2 are the number of recursions till there are no more vertices to visit to and from ~~word~~.

d_1 & d_2 are different ~~from~~ for all vertices

For ~~ever~~ every word w_1 there will be a ~~distribution~~ function

$$S(w_1) \implies S(w_1)(w_2) = S(w_1, w_2)$$

That will show how ~~mixture~~ ~~anything~~. how all the words in the graph are conceptually related to w_1.

~~If S can be learnt, then this can be thought of as a model of meanings depth in the language.~~

S can then be used to differentiate two words and also be utilized with ~~grammar~~ ~~text parser~~ language understanding system.

This model can be used to

$\boxed{\text{No water}}$

How do people store words in the mind?

In what order are words stored in the brain.

アイウエオ
カキクケコ

あいうえお
かきくけこ
さしすせそ
たちつてと
なにぬねの
はひふへほ
まみむめも
らりるれろ
やゆよ
わをん

がぎぐげご

アイウエオ
カキクケコ
サシスセソ
タチツテト
ナニヌネノ