**Проект solver**

**Размещение слов из словаря в памяти:**

При парсинге для каждого слова вычисляется свой хэш код, и этот код будет являться индексом для доступа к массиву из слов загруженного словаря. Вариант использовать вместо массива структуру типа map. Это увеличит время доступа к элементу, но развяжет руки с функцией хеширования. Она тогда сможет быть любой.

**Получить словарь из word**

Для использования словарей от OpenOffice необходимо переименовать расширение oxt на rar и открыть обычным архиватором. Файл с расширением aff содержит принципы изменения слов. Файл с расширением dic – сами слова.

Возможно придется распарсить какой-то(какие-то) словари под собственные нужды.

Для начала необходимо определить, что именно, из данных в словарях мне интересно, и лишь потом приводить имеющиеся слова к необходимости. Для начала составить собственные данные для словаря. Такой экспериментальный словарь.

Открыть Базу Ильи Козиева и посмотреть как он это делал! Структуру.

**Еще посмотреть примеры нейронных сетей…**

При составлении словарей может рассмотреть случай, когда в словарь добавляется часть слова и правила формирования полных слов из этой части.

Например, слово – мечта

мечт, сущ., ж., им.-а, род.-ы, дат.-ой,… мн. им.-ы, мн.род.-ы, мн.дат.-ами, глагол – мечтать, я – аю, ты – ешь, вы – ете, они –ют,… прил.им. - ательный,… ср.род им.-ательное

Количество окончаний и приставок не бесконечно!!! Можно всем приставкам и суфиксам и окончаниям раздать свои ID. И уже после этого собирать слова.

ДЛЯ ОПИСАНИЯ ЗНАЧЕНИЙ СЛОВ В СЛОВАРЯХ используется ограниченный набор слов. Необходимо после всего определить этот набор. И сделать определение смысла слов только на основании комбинаций этого набора

БЫТЬ МОЖЕТ для разбора слов и получения их хеш-кода придется давать различные ВЕСА для приставок, корней, суффиксов и окончаний.