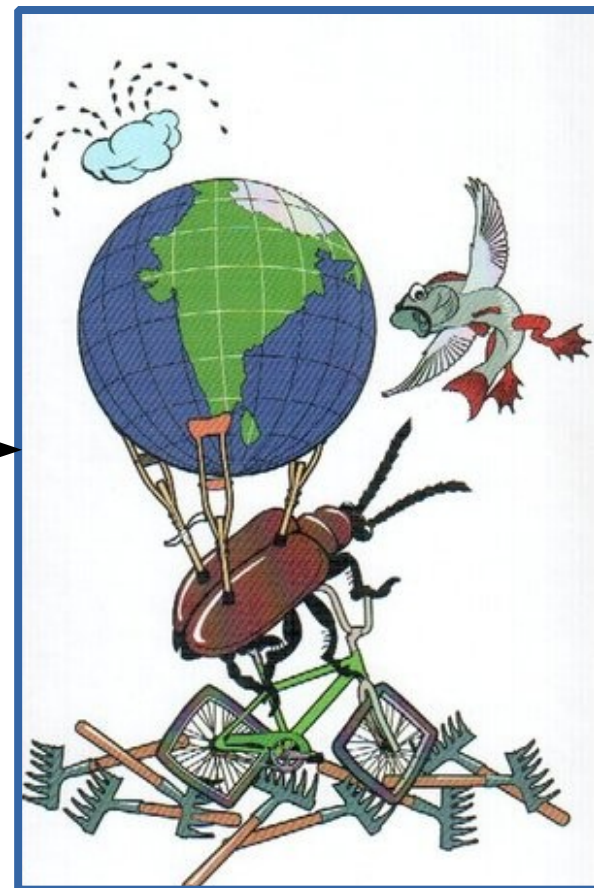
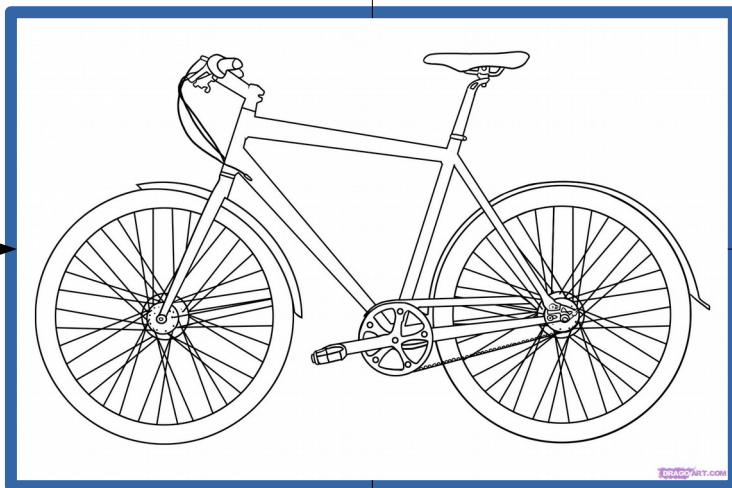


Значения слов

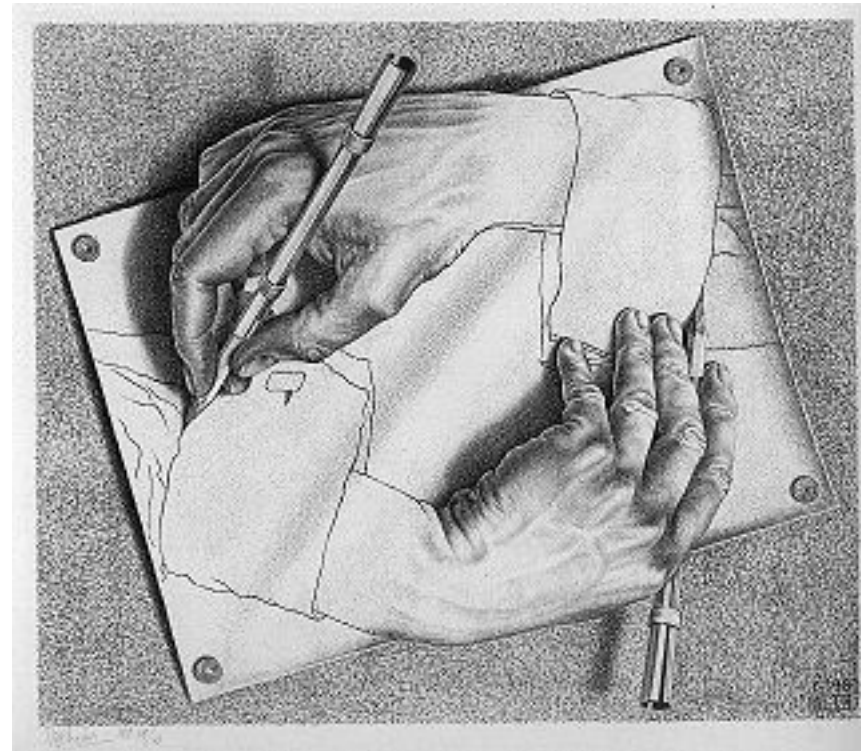
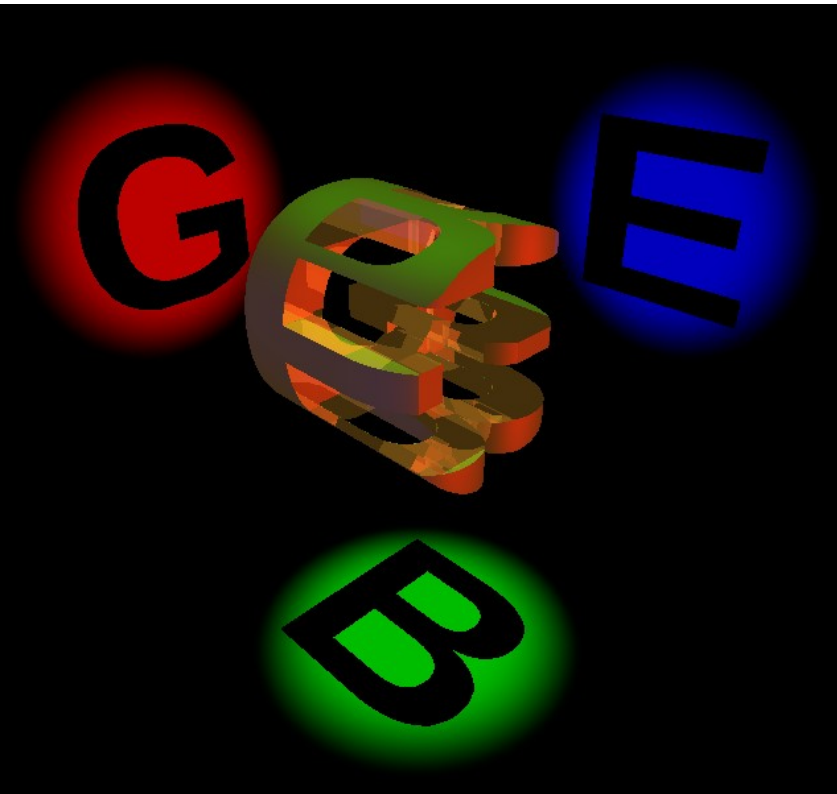
Велосипед



Студент 5-го курса
Химического факультета МГУ
Крылов Иван

Идея

- «Странные петли в сердце разума»
- Как можно взаимодействовать с «активными символами» в сознании человека?



Идея

- Упоминание концепции активирует связанный с ней символ
- Восприятие человеком значений слова показывает связи активных символов друг с другом
- Полученные данные можно использовать для восстановления и сравнения семантических сетей?

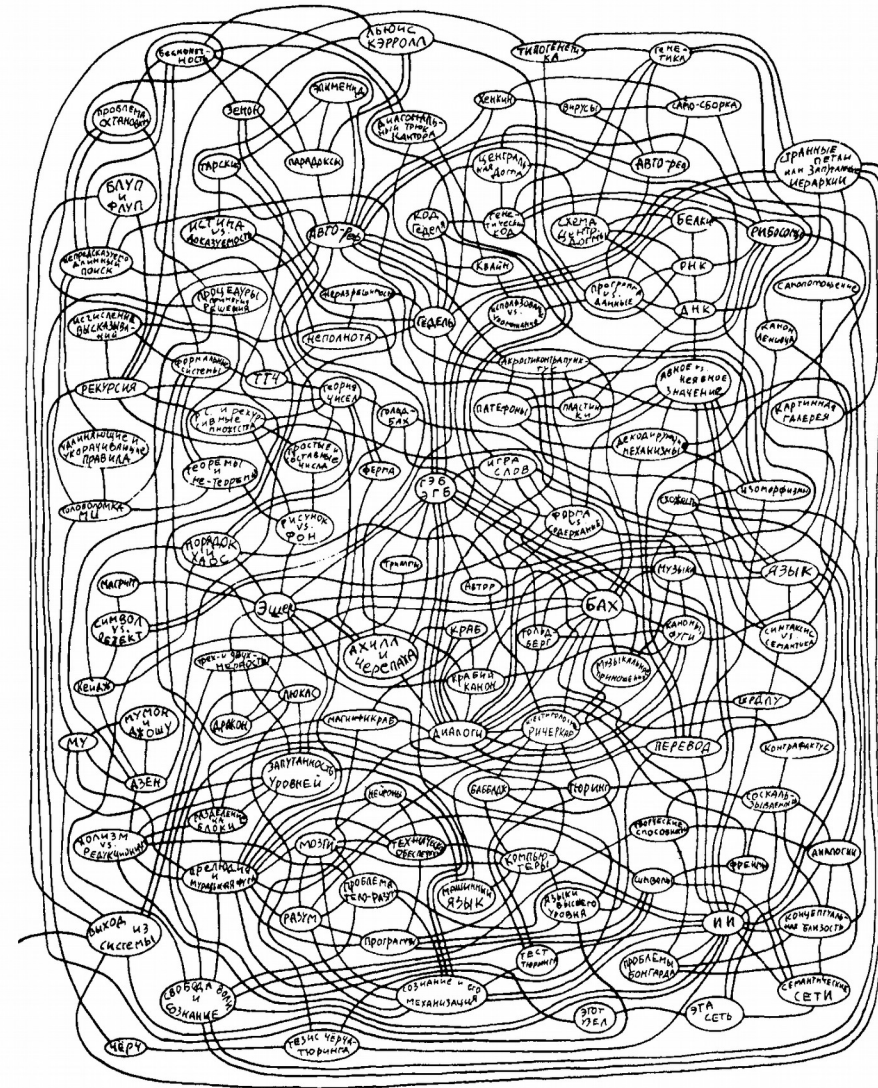
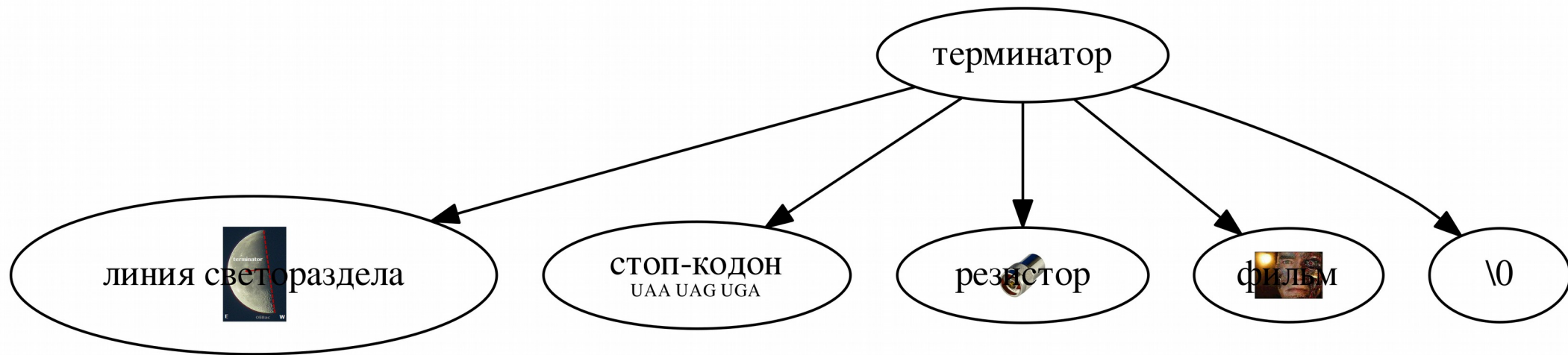


Рис. 70. Крохотный фрагмент семантической сети автора.

Методика



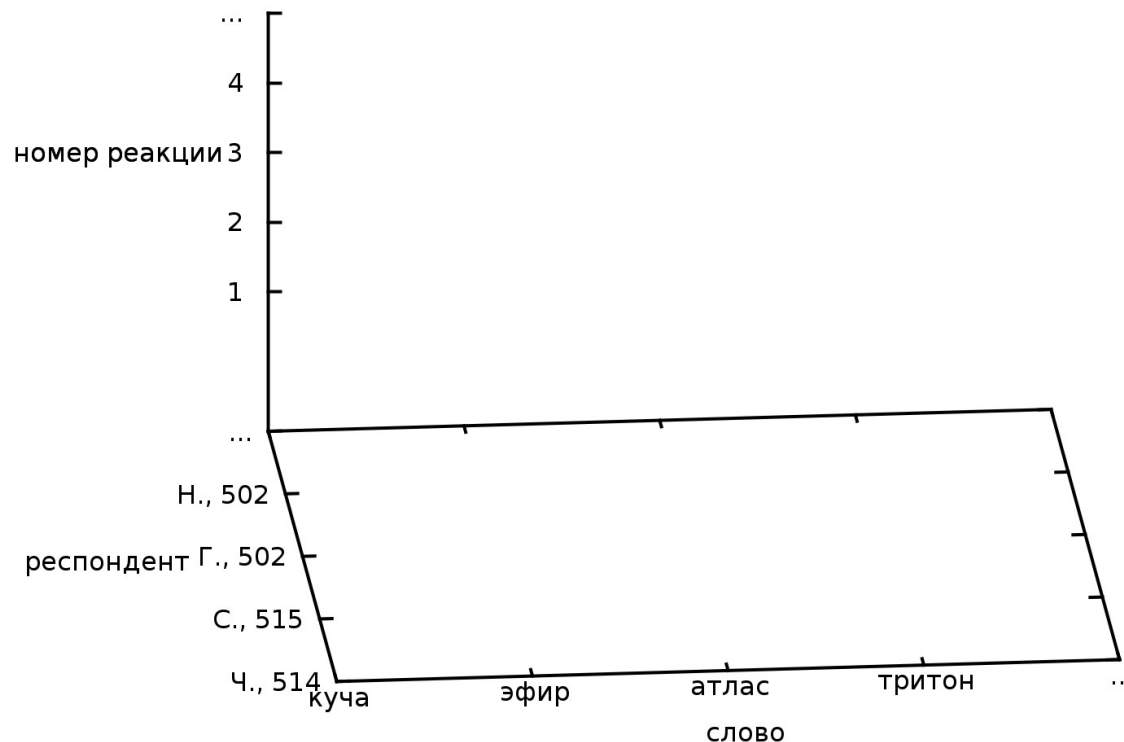
- 20 слов, подобранных интуитивно
 - «хорошими» считались многозначные слова
 - «матрица»
 - «сегмент» из ВМС
- 30 секунд на слово
- Глядя на надпись на экране, записать значения в порядке, в котором они приходят в голову
 - Значения, ассоциации, ...

Методика

- Отказ от интуитивного подбора слов
 - Гипотеза Г.Л.Васильева: «поток», «русло»
 - Пересечения существительных из MyVocabulary.com Alphalary for physics, chemistry, biology друг с другом и с UWL level 1 с последующим переводом
 - Опрос не-(физиков, химиков, биологов): «Какие слова приходят вам в голову при мыслях о химии/физике?»

Обработка: подготовка

- Семантическое сведение одинаковых по смыслу ответов в одинаковые строки
 - «Компьютер не знает русского»
- Помещаем строки в массив



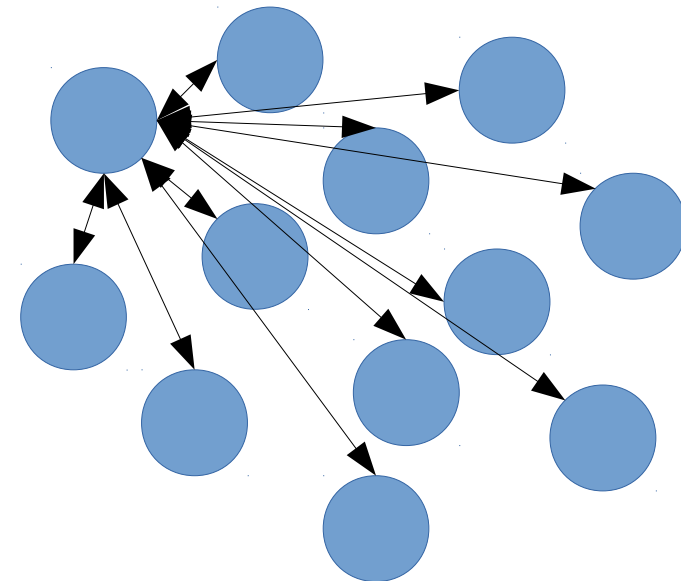
	Малютин				Евсюнина Мария, 514		
1	хвороста	рядовой "куча"			тетрадей	снега	листва
2	вещество	у греков	связь		вещество	телевизионный	
3	картографический	анатомический			картографический	ткань	
4	животное	спутник Нептуна	чудовище	ядро трития	животное		
5	кантилевер	атмосферный			космический		
6	лицо	соц.сети		преобладающее знание	соц	сети	
7	растения	из тюрьмы			из тюрьмы	растения	
8	товаров	автомата	журнал		товаров		
9	LEGO				LEGO		
10	растения	суть	мат	операция	слова	растения	суть
11	животного	тетрадная	Фарадея	грудная	биологическая	тетрадная	животного
12	таблица чисел	фильм	подложка		фильм	таблица чисел	полимерная
13	часть речи				часть речи	повод	
14	Куна	сферы	часть		часть	Куна	
15	многогранник	туловище	твердое		человека	небесное	
16	часовой	реакции			реакции	часовой	
17	поэт	в боевом искусстве			блок-сополимер	в общеаге	кирпич
18	функции	женские формы			кривая		
19	значок	знамение	шрифт		признак	кодирующий элемент	
20	дерева				дерева		

Обработка: расчёты

- Как получить числа?
 - Насколько различается пара ответов на данное слово?
 - Идея: пусть каждое отдельное значение в реакции играет роль буквы в слове
 - Применим расстояние Левенштейна
 - Поделим на сумму количеств значений в реакции, чтобы для малых количеств значений различия играли бóльшую роль

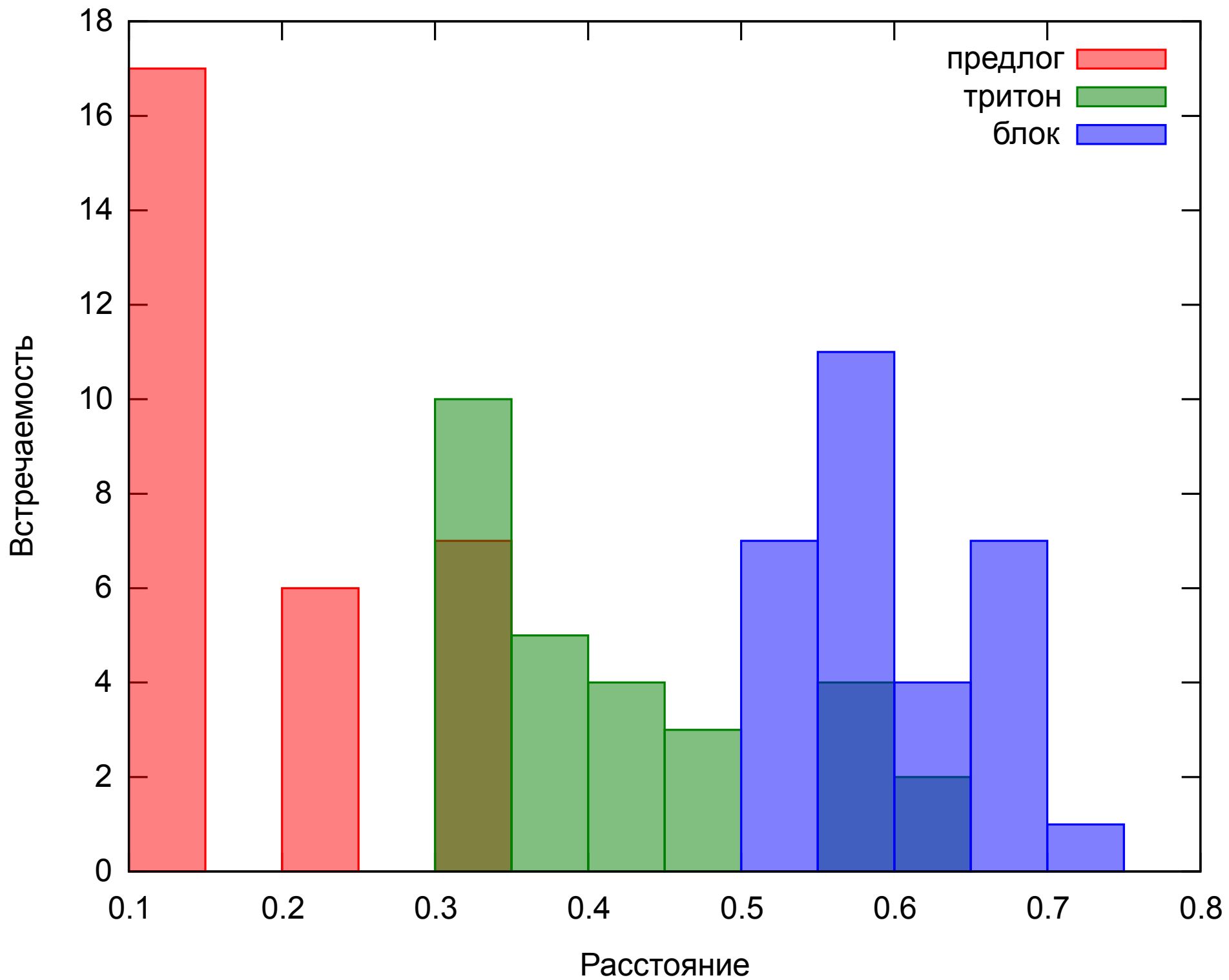
Обработка: величины

- Для каждого слова
 - Среднее отклонение каждого ответа на слово от всех остальных: как распределено?
- Для каждого респондента:
 - Среднее отклонение каждого ответа респондента от ответов всех остальных респондентов на то же слово



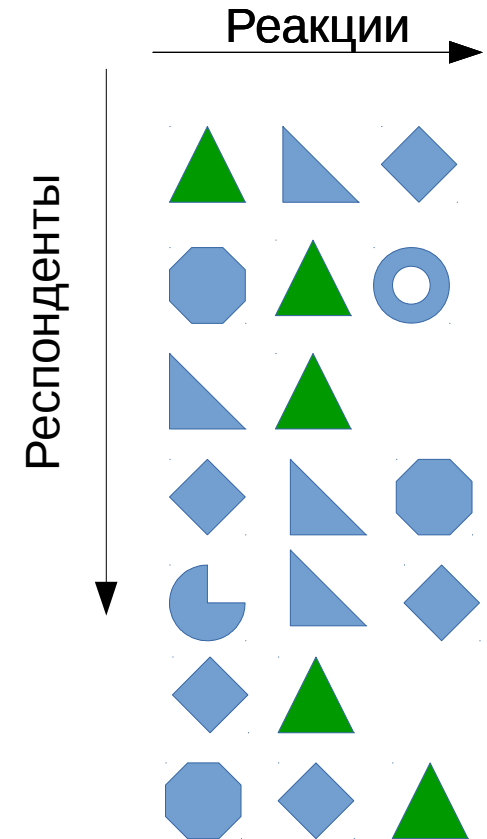
Отклонение слова от среднего

- Крайности:
 - «Предлог»: почти всегда «часть речи, повод»
 - Расстояние $0,19 \pm 0,05$
 - «Блок»: кто во что горазд, от поэта до боевых искусств
 - Расстояние $0,60 \pm 0,04$
- «Тритон»
 - В основном, «животное» и «чудовище»
 - Но ещё много других значений, от музыкального интервала до ПАВ
 - Расстояние $0,42 \pm 0,08$



О.В.Кукушкина: лидирующие реакции

- Для каждого слова
 - Было подсчитано количество уникальных индивидуальных реакций
 - Найдена наиболее частотная из них
 - Посчитана встречаемость наиболее частотной реакции в разных позициях

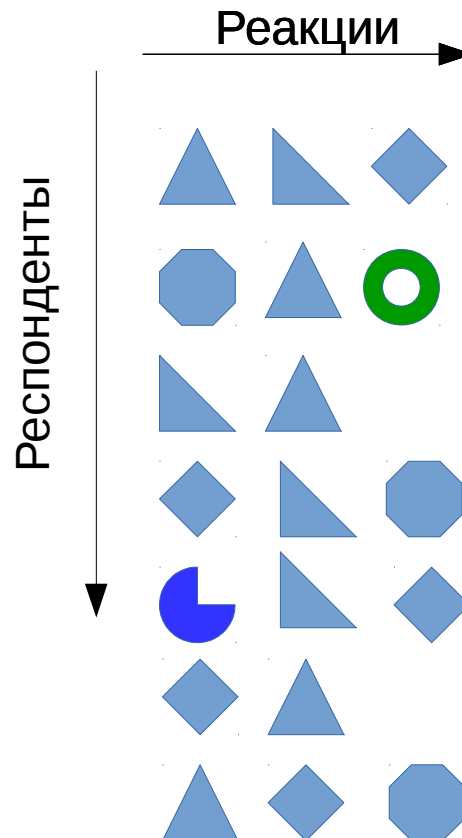


Слово	Всего реакций	Основная реакция	1	2	3	4
куча	25	вещей	7	5	1	1
эфир	13	вещество	23	3	1	0
атлас	12	картографический	22	2	0	0
тритон	18	животное	20	4	0	0
зонд	14	космический	8	6	1	0
профиль	15	соц.сети	8	4	1	0
побег	8	растения	7	11	1	1
магазин	6	товаров	24	3	1	0
конструк тор	9	профессия	7	11	0	0
корень	14	мат.операция	9	1	4	1
корень	14	растения	8	6	1	0

Слово	Всего реакций	Основная реакция	1	2	3	4
корень	14	слова	7	6	2	0
клетка	12	животного	16	4	2	0
матрица	13	таблица чисел	12	9	5	0
предлог	2	часть речи	21	6	0	0
сегмент	17	Куна	12	2	0	0
тело	30	человек	9	1	0	0
механизм	10	реакции	14	10	2	0
блок	34	поэт	9	2	0	0
выпуклость	13	неровность	11	1	0	0
символ	29	знак	9	4	0	0
ветвь	14	деревя	22	5	0	0

О.В.Кукушкина: уникальные реакции

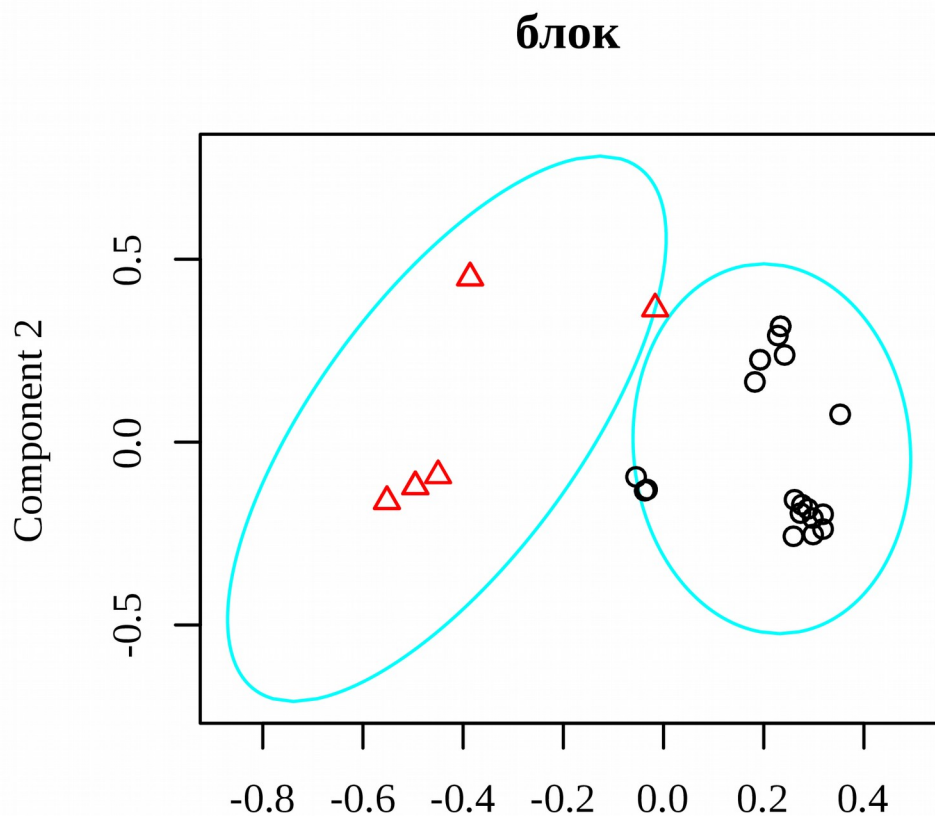
- Для каждого респондента были найдены реакции, которые не встретились у других



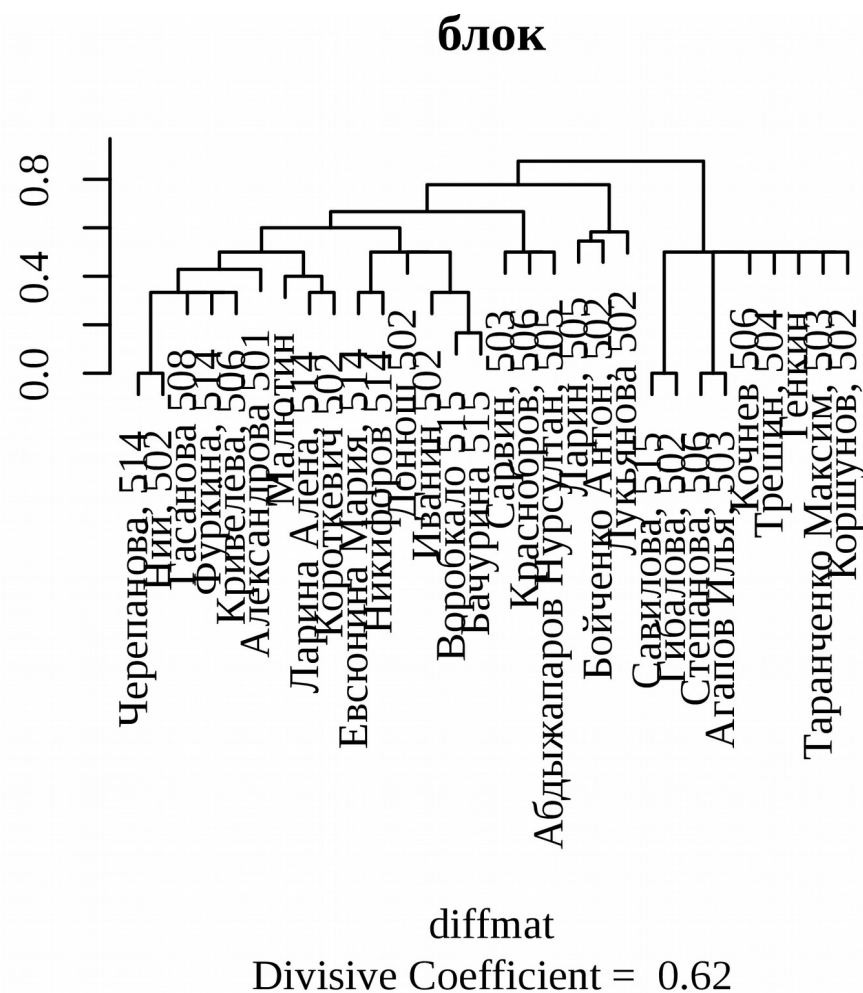
Фуркина, 514	7	куча	грязи	тритон	отец Ариэль
Ларин, 505	14	зонд	спец.подразделен ия	корень	волоса
Малютин	11	куча	хвороста	куча	рядовой "куча"
Евсюнина Мария, 514	6	куча	тетрадей	куча	снега
Ларина Алена, 514	5	побег	фильм	матрица	матрица
Коршунов, 502	16	куча	большой объем	куча	сложить
Абдыжапаров Нурсултан, 505	7	атлас	Атлант	тритон	тритон
Агапов Илья, 503	5	матрица	упорядоченное основание кристалл. решетки	выпуклость	метод выпуклых оболочек в термодинамике
Бойченко Антон, 502	21	куча	сахара	куча	(нецензурно!)
Таранченко Максим, 503	6	сегмент	кусок чего-либо	тело	подвижное
Трешин, 504	6	куча	навалено	механизм	
Никифоров 514	3	профиль	чего-либо	конструктор	машина

Обработка: кластеризация?

- Попарные расстояния между ответами на данное слово можно использовать для распределения ответов/респондентов по группам
- Для многих слов кластеры получаются плохо



These two components explain 35.83 % of the pair



Выводы

- Слабое место: семантическое сведение
 - Целиком и полностью завязано на экспериментатора
 - Выделение корня + подсчёт попарных расстояний между словами при помощи существующей базы данных + методики для работы с попарными расстояниями?
- Респонденты разделяются на уровне ответов на отдельные слова
 - Нужно считать попарные расстояния между респондентами по ответам на все слова