6. Смирнов А.В. Формирование мотивации учебной деятельности у студентов технического вуза // Психология, социология и педагогика. 2012. № 12, URL: <http://psychology.snauka.ru/2012/12/1450> (дата обращения: 28.09.2017).

Помельников Ю.В., Попов В.В., Штельмах Т.В.

Об автоматизации формирования визуальных образовательных средств

Волгоградский государственный университет

(Россия, Волгоград)

doi: 10.18411/lj-25-12-2017-41

idsp: 000001:lj-25-12-2017-41

Аннотация

В статье обосновывается и предлагается к использованию в образовательном процессе методика автоматического конструирования карт понятий. Обычная практика разработки такого вида методических средств требует больших затрат времени на анализ учебного материала и накладывает существенные ограничения на объем анализируемой информации. Предлагаемая методика позволяет значительно сократить затраты времени и обрабатывать текст практически произвольного объема.

Ключевые слова: визуализация, карта понятий, образование, автоматическая обработка текста, Python GenSim.

Большие потоки информации, сопутствующие процессу обучения, побуждают искать такие способы изложения материала, которые позволяют обучаемому в ускоренном режиме знакомиться с основными понятиями изучаемого предмета, а также с их взаимосвязями. Возможности машинного обучения и искусственного интеллекта эффективно используются в различных областях жизнедеятельности. Закономерно их применение как прикладного рабочего инструмента и в образовательной практике. В работе предлагается создание определенного вида графических наглядных методических материалов, основанных на обработке текста (или текстов) с помощью указанных технологий.

Обоснования эффективности такого вида средств обучения следует связать с исследованиями в нейропсихологии и психологии познания. Здесь разработаны подходы и методы осознания текста и оперирования со смыслами: образы и пиктограммы. Л. С. Выготский [Ошибка! Источник ссылки не найден., стр.439] пишет Пиктографическое письмо особенно легко развивается у ребенка потому, что, как мы видели, детский рисунок, в сущности, и является своеобразной его графической речью...". Причем он обращает внимание на то, что при возникновении трудностей в передачи мысли целым изображением ребенок рисует его "части, схему, а иногда, наоборот, создает более своеобразный путь для изображения ...". В дальнейшем эта концепция пиктограмм была развита в работах А. Р. Лурия, А. Н. Леонтьева и была использована в диагностике психотипов личности. Об особой эффективности визуальных способов представления информации указывает Е. А. Макарова [Ошибка! Источник ссылки не найден.]: "Повышается мотивация, мышление становится более гибким, таким образом можно избавиться от стереотипов, превратить догматическое мышление в критическое. Образы, представления и понятия объединяются в смыслообразе, что дает развитие логическому мышлению". Собственно технологические моменты при анализе познавательной деятельности встречаем, например, у И. А. Трухан, Д. А. Трухан [Ошибка! Источник ссылки не найден.]: "...способ систематизации и визуального отображения учебной информации основывается на выявлении существенных связей

между элементами знания и аналитико-синтетической деятельности при переводе вербальной информации в невербальную (образную), синтезирование целостной системы элементов знаний».

В западной традиции приложения этих концепций нашли воплощение в огромном количестве решений, связанных с так называемыми картами понятий или Mind Map. Появилось множество инструментов и ресурсов, в которых можно воспользоваться готовыми картами понятий или создавать свои, уже возможно за плату [Ошибка! сточник ссылки не найден.]. В отечественной образовательной практике эти инструменты применялись П. М. Эрднеевым [Ошибка! Источник ссылки не найден.], крупнѐнные дидактические единицы, а также методика опорных сигналов — В. Ф. Шаталовым [Ошибка! Источник ссылки не найден.]. В настоящее время это ироко распространенные образовательные средства.

Используем указанные выше свойства этих образовательных средств в рассмотрении некоторых особенностей современного вузовского образования в связи с возможностями применения автоматических методов обработки текста. Если проанализировать методы работы студентов с информацией в сети [Ошибка! Источник сылки не найден.] и эффективность ее использования, то возникает вопрос о возможности отрегулировать процесс получения информации из электронных источников, сделать его более результативным. Закономерные проблемы с осмыслением текста, в частности математического [Ошибка! Источник ссылки не найден.], также ызывает необходимость искать способы, призванные развить умения работы с текстом. Причем здесь надо учитывать интенсивность процесса обучения и достаточно большой объем необходимого учебного материала, а, значит, и временные ограничения как у студента, так и преподавателя. Привлечение ментальных карт в учебный процесс позволит исключить некоторые конфликтные ситуации. Например (в соответствии с [Ошибка! Источник ссылки не найден.] и [Ошибка! Источник ссылки не найден.]):

1. Схематичное представление понятий как отдельных тем, так и всей дисциплины позволит в них сориентироваться и скорректировать необходимый объем информации: студент уже не "завязнет" в ней.

2. Группирование и направленность понятий в схеме позволит организовать последовательность изучения материала конкретных тем, выявить слабые места в знаниях и умениях студента.

3. Организация работы с картами поможет сформировать умения самостоятельно определять причинно-следственные связи между частями текста, выделять главные тезисы в общем объеме текста.

4. Привлечение автоматических способов работы с текстом дает возможность экономии времени преподавателя по составлению необходимых схем.

В данной работе для автоматического построения карты понятий предложено использование метода обработки текста, основанного на следующем правиле: два понятия близки, если они находятся во множестве предложений в близком соседстве [Ошибка! сточник ссылки не найден.]. Алгоритм воплощен в библиотеке Python GenSim (generating similarity — выработка сходства) в котором каждый терм, выраженный словом, преобразуется в вектор в многомерном пространстве, в котором обычно от 50 до 300 компонент. Эти векторы нумеруют единицы текста (термы) по отношению к единицам корпуса документов и положению терма в самом документе. Этот метод называется W2V (word to vector). Сходство этих векторов определяется как степень близости направлений,

измеряемой с помощью расстояния между двумя векторами с концами на единичной сфере:

$$\cos(v,u) = (v,u)/(\|v\| \cdot \|u\|).$$

Кроме этого, параметр сравнения или сходства по контексту термина предполагается как настраиваемое число: так называемая энграмма (n-gram) — количество слов в окрестности данного. В GenSim для определения сходства термов есть специальные методы, например, `most_similar`. С векторами можно работать как с элементами многомерного пространства, к которым применяются обычные операции сложения и вычитания. Например, необходимо найти слова, близкие к 'woman', 'king', без рассмотрения слова 'man', т.е. складываются 'woman', 'king' и вычитается 'man'. Вызов этого метода может быть таким:

```
model.wv.most_similar(positive=['woman', 'king'], negative=['man']).
```

Результат: [('queen', 0.50882536), ...], где числа в списке — расстояние между выбранным понятиями и найденным связанным с ними соответствующим термом.

Нужно отметить, что построение словаря и векторного представления для него основывается на обучении нейронной сети. Материалом для обучения являются файлы в текстовом формате, объем которых может быть достаточно большим, вплоть до всей Википедии. Сам словарь, построенный в результате обучения нейросети, используется в нашей работе для построения семантического дерева понятий. Пользователь может выбрать коренное понятие (терм) и в дальнейшем получить все связанные с этим понятием термы. В виде графа или дерева, развертывающегося от вершины коренного понятия к остальным.

С помощью изложенного выше метода можно построить следующий пример:

```
model.wv.most_similar('программа')[0:4]
[('язык', 0.9999394416809082), ('проект', 0.9999381303787231), ('цель',
0.9999366402626038), ('использование', 0.9999361038208008)]
[t[0] for t in model.wv.most_similar('программа')[0:4] ]
['язык', 'проект', 'цель', 'использование']
[[t1[0] for t1 in model.wv.most_similar(t0[0])[0:4] ] for t0 in
model.wv.most_similar('программа')[0:4] ]
['программирование', 'проект', 'система'], ['разработчик', 'система', 'программист',
'проектирование'], ['проект', 'система', 'вид'], ['система', 'проект', 'помощь', 'язык']]
```

Для графического представления результата была использована программа Graphviz [Ошибка! Источник ссылки не найден.].

Граф близких понятий для слова "программа"

На диаграмме изображен граф понятий, полученный из текста учебника Б. Старуструпа "Язык программирования C++". Все понятия, связанные с корневым словом "программа" изображены в порядке их связанности с корнем. Он строится следующим образом: сначала выбираются первые 4 термина второго уровня, затем каждый из термов второго уровня связан с четырьмя терминами третьего уровня и, возможно, далее. Такие схемы дают представление о тематической связанности понятий в данном тексте.

Граф близких понятий для слова "транслятор"

Кроме используемого нами подхода существуют решения исследователей в области информационных технологий и русскоязычного семантического анализа, так называемого дистрибутивного тезауруса русского языка. Однако наличие такого проекта не отменяет необходимости широкого внедрения интерактивных методов обучения, основанных на автоматическом анализе текста и построения его семантических схем.

1. Выготский Л. С. Психология развития ребенка. – М.: Изд-во Смысл, Изд-во Эксмо, 2004. – 512с. (Серия «Библиотека всемирной психологии»).
2. Макарова Е. А. Визуализация как способ структурирования знаний и формирования ментального пространства [Электронный ресурс] URL: http://sosh3.oprb.ru/data/partner/6/message/RR9f14_3049.pdf (дата обращения: 17.12.2017).
3. Трухан И. А., Трухан Д. А. Визуализация учебной информации в обучении математике, ее значение и роль [Электронный ресурс] // Успехи современного естествознания. – 2013. – № 10. – С. 113-115. URL: <https://www.natural-sciences.ru/ru/article/view?id=32992> (дата обращения: 17.12.2017).
4. Эрдниев П. М. Укрупнение дидактических единиц как технология обучения, ч. 1. – М.: Просвещение, 1992.
5. Шаталов В. Ф. "Эксперимент продолжается" – М. Педагогика, 1989. 336с.
6. Штельмах Т. В. О некоторых особенностях классического обучения в сопровождении интернета [Текст] / Т. В. Штельмах// Международное научное периодическое издание "Новая наука: теоретический и практический взгляд" по итогам международной научно-практической конференции (Стерлитамак, 14.12.2015 г.). – Стерлитамак: АМИ, 2015.
7. Штельмах Т. В. О проблемах осмысления студентами математических текстов [Текст] / Т. В. Штельмах// Международное научное периодическое издание «Новая наука: проблемы и перспективы» по итогам международной научно-практической конференции (Стерлитамак, 26.12.2016 г.). – Стерлитамак: АМИ, 2016.
8. Graphviz [Офиц. сайт]. URL: <https://www.graphviz.org/> (дата обращения: 17.12.2017).
9. MindMeister [Офиц. сайт]. URL: <https://www.mindmeister.com/> (дата обращения: 17.12.2017).
10. Tomas Mikolov, Ilya Sutskever, Kai Chen, Gregory S. Corrado, and Jeffrey Dean. Distributed representations of words and phrases and their compositionality. In Advances in Neural Information Processing Systems 26:

27th Annual Conference on Neural Information Processing Systems 2013. Proceedings of a meeting held December 5-8, 2013, Lake Tahoe, Nevada, United States, pages 3111–3119, 2013.