

Семантическая близость астрономических понятий, обозначающих небесные тела

Semantic Proximity of Astronomical Concepts Denoting Celestial Bodies

Получено: 05.07.2024 / Одобрено: 12.07.2024 / Опубликовано: 25.09.2024

Майер Р.В.

ФГОУ ВО «Глазовский государственный инженерно-педагогический университет им. В.Г. Короленко», Россия, 427621, г. Глазов, ул. Первомайская, д. 25, e-mail: robert_maier@mail.ru

Mayer R.V.

Glazov Korolenko State Engineering and Pedagogical University, 25, Pervomayskaya St., Glazov, 427621, Russia, e-mail: robert_maier@mail.ru

Аннотация. Рассмотрена проблема определения семантической близости астрономических понятий, обозначающих небесные тела в школьном курсе астрономии. Используется дистрибутивный метод, заключающийся в сравнении распределений понятий в тексте учебника астрономии. С помощью специальных компьютерных программ из текста удаляются знаки препинания, стоп-слова, определяется число упоминаний различных терминов, вычисляются косинусная мера близости и семантическое расстояние между понятиями. Это позволяет: 1) выявить часто используемые астрономические понятия, обозначающие небесные тела; 2) получить матрицы близости и семантических расстояний между понятиями «галактика», «звезда», «Солнце», «планета», «черная дыра», «белый карлик» и т.д.; 3) создать ментальную карту небесных тел, входящих в состав Солнечной системы с учетом расстояний между обозначающими их терминами; 4) построить облака понятий «виды звезд», «Земля – звезда – Вселенная». Для моделирования семантического пространства учебника используется компьютерная программа Graf.pas, которая случайным образом перемещает вершины в 2D- или 3D-пространстве, каждый раз вычисляя потенциальную энергию W соединяющих их упругих стержней и выбирая такое расположение вершин, при котором W минимально. Считается, что близко расположенные вершины соединены жесткими стержнями, а далеко удаленные – эластичными резиновыми нитями.

Ключевые слова: астрономия, дидактика, понятие, семантическое пространство, компьютерные методы, косинусная близость, учебник.

Abstract. The problem of determining the semantic similarity of astronomical terms denoting celestial objects in a school astronomy course is considered. A distributive method is used, which consists in comparing the distributions of concepts in the astronomy textbook. With the help of special computer programs, punctuation marks and stop words are removed from the text, the number of mentions of various terms is determined, the cosine measure of proximity and the semantic distance between concepts are calculated. This allows: 1) to reveal frequently used astronomical concepts denoting celestial bodies; 2) to obtain matrices of proximity and semantic distances between the concepts of "galaxy", "star", "Sun", "planet", "black hole", "white dwarf", etc.; 3) to create a mental map of the celestial bodies that make up the Solar system, taking into account the distances between the terms designating them 4) to build clouds of concepts "types of stars", "Earth – star – Universe". To model the semantic space of the textbook, the Graf.pas computer program is used, which randomly moves vertices in 2D or 3D space, each time calculating the potential energy W of the elastic rods connecting them and choosing such an arrangement of vertices that W is minimal. It is believed that the closely located vertices are connected by rigid rods, and the far removed ones are connected by elastic rubber threads.

Keywords: astronomy, didactics, concept, semantic space, computer methods, cosine proximity, textbook.

Введение

С позиций системного подхода школьный курс астрономии – сложный многомерный объект, который можно представить в виде совокупности ключевых концептов и понятий, связанных многочисленными семантическими связями. Анализ его содержания и структуры помогает лучше понять логику изложения учебного материала, построить ее модель и оптимизировать методику преподавания. Рассматривая учебник как модель курса, можно выявить самые важные и часто используемые научные понятия (НП), без понимания которых усвоение учебного материала невозможно [8]. Проблема

исследования семантического пространства (СП) учебника астрономии, предполагающая определение степени семантической близости между ключевыми астрономическими понятиями, является актуальной.

Рассмотрим идеального ученика, который воспринимает и усваивает всю учебную информацию, сообщаемую на уроке. Учитель, излагая учебный материал, придерживается содержания учебника астрономии, который можно рассматривать как модель изучаемого курса. При этом формируемые у ученика связи будут определяться: 1) содержанием учебника, связями между НП; 2) результатами взаимодействия с окружающим миром (наблюдений,

опытов); 3) информацией, поступающей из других источников (телевидение, Интернет, книги и др.). Определяющей деятельностью школьника является учебная деятельность, поэтому сообщаемые учителем знания сильно влияют на систему понятий, образующую СП в сознании ученика. Оно является «оттиском» СП учебника, представляющего собой часть концептосферы, совокупность взаимосвязанных научных терминов, отображающих рассматриваемый фрагмент реальной действительности и составляющих сложную систему значений, символов и концепций, которые формируют смысл излагаемого учебного материала. СП учебника отражает объективные закономерности изучаемого материала и позволяет выявить особенности СП, образуемого астрономическими НП в сознании ученика.

Один из способов систематизации знаний и их представления в графическом виде заключается в создании ментальных карт. Обычно они выглядят как древовидные графы, в центре которых расположен центральный концепт, а вокруг него – второстепенные понятия первого, второго и третьего уровня, связанные между собой ребрами с написанными на них глаголами. Ментальные карты отражают логические связи между различными понятиями и могут использоваться в качестве графической модели лекции или рассказа [15]. При всех своих преимуществах ментальные карты имеют крупный недостаток, – расстояния между вершинами графа не соответствуют семантическому расстоянию между понятиями. Представим себе граф, удовлетворяющий этому требованию, – в каком-то смысле это была бы карта наших знаний!

Связанные между собой проблемы выявления ключевых терминов, построения тезауруса, изучения СП, установления семантической близости между понятиями анализировались различными учеными. Например, в [9] научный текст рассматривается как ментальное пространство, основанное на метафорических моделях. В [7] обсуждаются методы автоматического построения тезауруса некоторой предметной области и извлечения из текстов семантических отношений между терминами. В [12] анализируются различные алгоритмы определения степени семантической связанности понятий. В [10] обсуждаются способы построения семантических пространств различных предметных областей, основанные на методах дистрибутивной семантики. В статьях [2; 11] рассматриваются методы вычисления различных мер семантической близости понятий.

Цель исследования – определить степень содержательной связанности астрономических понятий,

обозначающих небесные тела; построить граф связей между понятиями с учетом их семантической близости. **Методологической основой** исследования являются работы следующих ученых: А.С. Ванюшкин и Л.А. Гращенко [3], С.О. Шереметьева и П.Г. Осминин [14] (извлечение ключевых слов); А.В. Анисимов, А.А. Марченко и В.К. Кисенко [1], М.И. Варламов и А.В. Коршунов [4], А.И. Панченко, С.А. Адейкин, А.В. Романов и П.В. Романов [12], М.С. Каряева [7], Д.С. Бранченко, Ю.А. Кравченко и Ю.С. Новикова [2], Н.В. Найханов и Б.А. Дышенов [11], А. Ashraf, A. Favez и A. Hani [16], D. Lin [17], P.D. Turney и P. Pantel [19] (смысловая близость терминов); Д.А. Ичкинеева [6], Н.А. Мишанкина [9], Ю.И. Морозова [10], С.В. Ракитина [13], С.Д. Manning, P. Raghavan и Н. Schütze [18] (семантическое пространство). Используются компьютерные программы, написанные в ABCPascal, которые обрабатывают текст, вычисляют «силу» семантических связей между понятиями и на основе этого моделируют семантическое пространство учебника.

Обсуждение проблемы исследования

В соответствии с идеей антропоцентризма система знаний, формируемая в сознании человека, строится на основе его особенностей восприятия мира. Это в полной мере относится к астрономической картине мира, формирование которой опирается на обыденное знание, включающее примитивные представления о Земле, Солнце, Луне и звездах. Развитие астрономических знаний в школе зависит от возможностей человека воспринимать окружающий мир органами зрения, а также от «расположения» человечества в Солнечной системе, галактике Млечный Путь и Вселенной. Это влияет на семантическую связанность важнейших астрономических понятий в школьном учебнике астрономии.

Проблема определения семантической близости ключевых понятий может быть решена методами дистрибутивной семантики, т.е. путем учета их распределения в тексте учебника. В основе лежит **дистрибутивная гипотеза**: лингвистические единицы (слова), часто встречающиеся в одинаковых фрагментах текста, связаны между собой семантически. Допустим, мы обсуждаем концепт «галактика». Очевидно, что в абзацах и предложениях, содержащих этот термин, также часто будет встречаться слово «звезда», а термины «метеорит» или «комета» – реже. Значит, семантическое расстояние между понятиями «галактика» и «звезда» существенно меньше, т.е. они ближе по смыслу. Чем больше семан-

тическая связанность слов, тем сильнее похожи их распределения в тексте, тем чаще эти понятия встречаются в схожих контекстах.

Сущность используемого метода состоит в следующем.

1. Полный текст учебника астрономии [5] конвертируют в текстовый файл *astronomiya.txt* и преобразуют все прописные буквы в строчные. С помощью специальной программы *Prog1.pas* удаляют знаки препинания, стоп-слова, двойные и тройные пробелы, пустые строки. К стоп-словам относятся предлоги, союзы, междометия; они не несут важной информации. Двойные термины преобразуют так, чтобы они воспринимались программой как одно слово, например: черная дыра → черн_дыра, белый карлик → бел_карлик, нейтронная звезда → нейтрон_звезда и т.д.
2. Текст разбивают на перекрывающиеся контекстные окна – фрагменты из 60 слов, с каждым шагом смещающиеся на 30 слов. С помощью компьютерной программы *Prog2.pas* для каждого термина A определяется координата x (номер слова от начала текста) и устанавливается контекстный вектор – одномерная матрица, элементы которой a_k равны количеству вхождений термина в k -ое контекстное окно ($k = 1, 2, \dots, N$). Множество контекстных векторов образует словесное векторное пространство.
3. Семантическая близость понятий A и B характеризуется косинусом угла между векторами $\vec{a}(a_1, a_2, \dots, a_N)$ и $\vec{b}(b_1, b_2, \dots, b_N)$ словесного пространства. Их компонентами являются количества этих понятий a_k и b_k в k -ом контекстном окне (некоторые из них равны 0). Программа *Prog2.pas* вычисляет косинус угла между этими векторами [1; 2]:

$$K(A, B) = \frac{a_1 b_1 + a_2 b_2 + \dots + a_N b_N}{\sqrt{a_1^2 + a_2^2 + \dots + a_N^2} \cdot \sqrt{b_1^2 + b_2^2 + \dots + b_N^2}},$$

где N – число контекстных окон. Когда вектора \vec{a} и \vec{b} совпадают, $K(A, B) = 1$. Если понятия A и B ни разу не попали в одно контекстное окно, то вектора и ортогональны, $a_k b_k = 0$ для любых k , $K(A, B) = 0$.

4. В таблицах *Excel* вычисляют семантическое расстояние между терминами: $L(A, B) = 1/K(A, B) - 1$. При $K(A, B) = 0$ (связь отсутствует) оно бесконечно велико, а при $K(A, B) = 1$ – равно 0. Получают матрицу расстояний между НП.
5. С помощью компьютерной программы *Graf.pas* на основе полученной матрицы расстояний стро-

ят двумерный или трехмерный граф, визуализирующий семантические связи между понятиями. Возможны два подхода: 1) выбирают 3–4 близко расположенных понятия, связи между ними моделируются жесткими стержнями (их длины практически не меняются), а остальные – упругими нитями; 2) находят центральное понятие, с которым тесно связаны остальные НП; инцидентные ему ребра считаются очень жесткими, а остальные – эластичными. Длинные ребра, соответствующие слабым связям между понятиями, могут быть изображены кривыми линиями, на которых указаны их веса (расстояния в СП).

Результаты исследования

Для анализа выбран учебник «Астрономия» (авторы И.В. Галузо, В.А. Голубев, А.А. Шимбалёв) [5], электронную версию которого можно найти в Интернете. В результате были выявлены следующие ключевые понятия [3; 14], обозначающие небесные тела: 1) **звезда** (1038 упоминаний): звезда, звездный, квазизвездный, межзвездный, протозвезда, созвездие; 2) **Солнце** (740): Солнце, солнцестояние, солнечный; 3) **Земля** (531): внеземной, землетрясение, Земля, земной, наземный; 4) **планета** (392): межпланетный, планета, планетезималь, планетный, планетология; 5) **галактика** (286): галактика, галактический, межгалактический; 6) **Луна** (207): Луна, лунный, полнолуние; 7) **спутник** (107); 9) **Марс** (81): Марс, марсианский; 10) **Юпитер** (50); 11) **астероид** (49); 12) **Венера** (41); 13) **Меркурий** (29); 14) **черная дыра** (24); 15) **Сатурн** (22); 16) **Уран** (21); 17) **Нептун** (17); 18) **квazar** (16); 19) **белый карлик** (13); 20) **нейтронная звезда** (12); 21) **пульсар** (6); 22) **звезда главной последовательности** (4). К этому списку добавим концепт **Вселенная** (98).

В табл. 1 представлены полученные с помощью программы *Prog2.pas* значения косинусной близости (слева внизу под диагональю) и расстояний (справа вверху) между понятиями в семантическом пространстве. Связи астрономических понятий с ключевым концептом можно упрощенно представить в виде списков: 1) **звезда**: Солнце – расстояние 3,1, Земля – 8,9, планета – 16,3, спутник – 23,2; 2) **Солнце**: Земля – 1,8, планета – 2,7, звезда – 3,1, Луна – 4,2, Меркурий – 8,4, Венера – 9,1, спутник – 10,1, астероид – 11, Юпитер – 11,2, Марс – 12,8, Уран – 12,9, Сатурн – 15,9, Нептун – 16,6; 3) **планета**: Земля – 2,2, Спутник – 2,4, Меркурий – 2,6, Марс – 2,6, Юпитер – 2,6, Солнце – 2,7, Венера – 2,8, астероид – 3,7, Уран – 4,1, Сатурн – 5,2, Нептун – 6, Луна – 8,2, звезда – 16,3; 4) **астероид**: Юпитер – 3,2, пла-

нета – 3,7, спутник – 7, Земля – 7, Марс – 7,7, Солнце – 11, Нептун – 12,4, Меркурий – 17,1.

На основе табл. 1 построены графы, визуализирующие семантическую близость между различными понятиями (рис. 1.1 и 1.2). Из них видно, что понятие «звезда» ближе к НП «Солнце», а НП «астероид» – к понятиям «Юпитер» и «планета». Недостатком подобного подхода является произвольность выбора центрального концепта и пренебрежение семантическими расстояниями между нецентральными понятиями.

Если для каждого НП вычислить среднее расстояние до остальных понятий, то удастся выявить центральные концепты, вокруг которых концен-

трируются остальные НП. К таким центральным концептам относятся понятия: «планета» $L_{cp} = 4,7$, «Земля» – 7,3, «Солнце» – 9,2 (табл. 1). Поэтому в центр графа поместим концепт «Планета», а все остальные понятия расположим так, чтобы их расстояния до центрального НП соответствовали найденным значениям $L_{i,j}$. Учтем, что НП «Луна» тесно связано с понятиями «спутник», «Земля», «Солнце» (табл. 1), а также, что L (галактика, звезда) = 3,8 и L (Вселенная, галактика) = 3,8. Получившийся граф (рис. 2) напоминает ментальную карту, отличаясь от нее тем, что был построен не произвольно, а с учетом наименьших семантических расстояний между входящими в него понятиями.

Таблица 1

Косинусная близость и расстояния между понятиями

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1 звезда		3,07	16,26	200,6	23,24	46,51	43,31	8,88	71,76	25,52	44,44	78,30	41,17	105,4
2 Солнце	0,25		2,66	10,99	10,07	8,44	9,13	1,81	12,83	11,21	15,92	12,94	16,57	4,20
3 планета	0,06	0,27		3,70	2,39	2,61	2,82	2,20	2,61	2,62	5,18	4,09	6,00	8,17
4 астероид	0,00	0,08	0,21		6,95	17,12	31,58	6,99	7,71	3,23	37,11	35,29	12,40	1,E+05
5 спутник	0,04	0,09	0,30	0,13		15,08	20,03	2,49	13,22	1,97	4,35	14,62	11,97	2,90
6 Меркурий	0,02	0,11	0,28	0,06	0,06		3,34	9,45	7,98	6,26	9,28	13,68	10,56	35,13
7 Венера	0,02	0,10	0,26	0,03	0,05	0,23		6,30	6,30	11,74	9,86	11,22	16,02	54,12
8 Земля	0,10	0,36	0,31	0,13	0,29	0,10	0,14		5,16	11,47	9,68	16,21	11,39	2,88
9 Марс	0,01	0,07	0,28	0,11	0,07	0,11	0,14	0,16		6,30	27,99	23,54	31,61	53,33
10 Юпитер	0,04	0,08	0,28	0,24	0,34	0,14	0,08	0,08	0,14		5,39	24,33	5,41	9,39
11 Сатурн	0,02	0,06	0,16	0,03	0,19	0,10	0,09	0,09	0,03	0,16		17,72	13,22	7,64
12 Уран	0,01	0,07	0,20	0,03	0,06	0,07	0,08	0,06	0,04	0,04	0,05		5,77	1,E+05
13 Нептун	0,02	0,06	0,14	0,07	0,08	0,09	0,06	0,08	0,03	0,16	0,07	0,15		37,87
14 Луна	0,01	0,19	0,11	0,00	0,26	0,03	0,02	0,26	0,02	0,10	0,12	0,00	0,03	

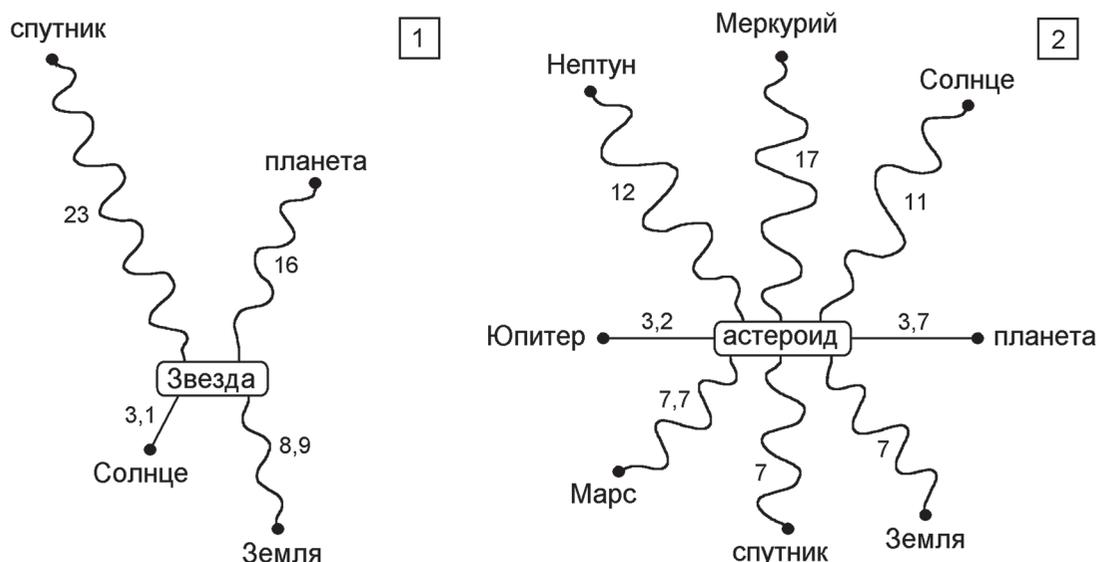


Рис. 1. Семантические расстояния между некоторыми понятиями

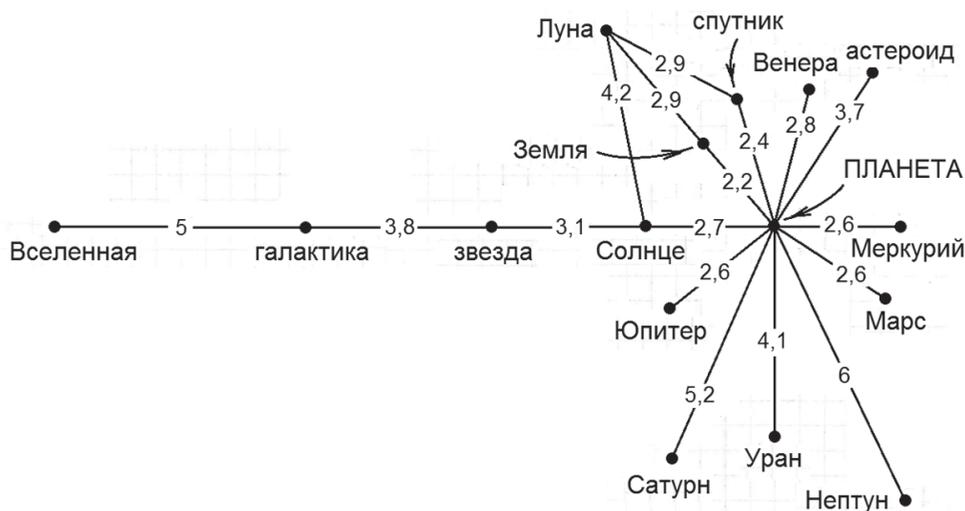


Рис. 2. Ментальная карта, учитывающая семантическую близость понятий

Для оптимального расположения вершин графа в 2D-пространстве используется компьютерная программа *Graf.pas*. Алгоритм ее работы состоит в следующем: вершины графа случайным образом располагаются на плоскости, затем вычисляется «потенциальная энергия» W всех упругих связей между ними. При этом считается, что близко расположенные вершины соединены жесткими стержнями, а далеко удаленные вершины – эластичными резиновыми нитями (у них жесткость $k_{i,j}$ меньше). Используются формулы:

$$W = \sum_{i=1}^{N-1} \sum_{j=i+1}^N k_{i,j} \frac{(L_{i,j} - s_{i,j})^2}{2},$$

$$s_{i,j} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2},$$

где $s_{i,j}$ — расстояние между i -ым и j -ым понятиями на данном шаге оптимизации. Затем случайно выбранная вершина сдвигается на небольшое расстояние и снова вычисляется потенциальная энергия W' . Если потенциальная энергия уменьшилась ($W' < W$), то новые значения координат вершин принимаются, а если нет – отвергаются. После этого все повторяется снова. В результате десятков тысяч итераций получается граф, близкий к оптимальному. Как правило, программа застревает в одном из локальных минимумов W , поэтому следует несколько раз запустить программу и выбрать результат с наименьшим W .

На рис. 3 показано, как распределены в семантическом пространстве термины, обозначающие небесные тела Солнечной системы. В качестве центрального выбрано НП «Солнце». В первую очередь

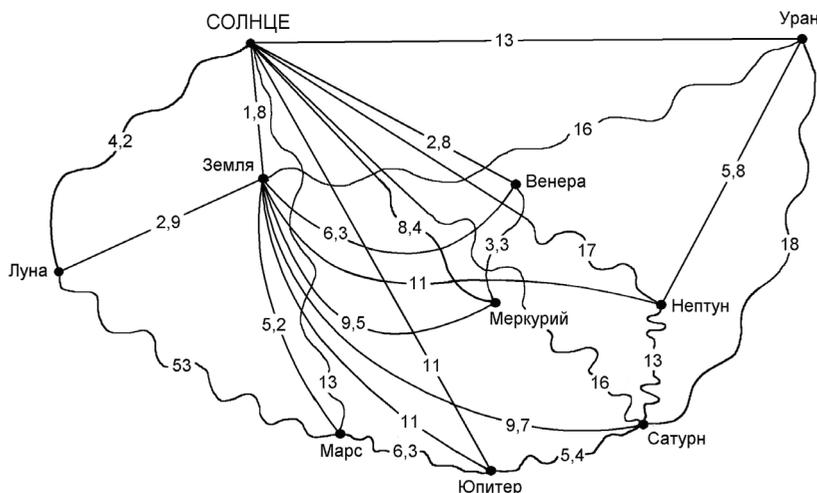


Рис. 3. Понятия, обозначающие тела Солнечной системы, и связи между ними

учитывались короткие (сильные) связи; они показаны прямыми линиями. Это связи «Солнце – Земля – Луна», связи между планетами земной группы, ледяными и газовыми гигантами.

Найденные значения косинусной меры близости и семантического расстояния между понятиями «Вселенная», «галактика», «звезда», «Солнце», «Земля» представлены в табл. 2. На ее основе построен граф (рис. 4). Программа расположила вершины графа (понятия) так, чтобы длины ребер приблизительно равнялись $L_{i,j}$. Короткие сильные связи моделируются жесткими стержнями и изображаются прямыми отрезками. Те вершины, расстояние между которыми велико, соединяются кривыми ребрами.

Таблица 2

Понятия из множества «Земля – звезда – Вселенная»

	Вселенная	галактика	звезда	Солнце	Земля
Вселенная		4,95	20,41	56,11	34,57
галактика	0,168		3,81	11,25	48,73
звезда	0,047	0,208		3,07	8,88
Солнце	0,018	0,082	0,246		1,81
Земля	0,028	0,020	0,101	0,356	

При решении этой задачи на плоскости (в 2D-пространстве) получается результат, представленный на рис. 4. В качестве центрального выбрано понятие Вселенная; программа «старалась» в первую очередь

выдержать расстояние от него до остальных четырех НП. Обсуждаемые понятия образуют множество точек в многомерном искривленном пространстве, а на рисунках изображаются их проекции на некоторую плоскость, получается двумерный граф. Теоремы геометрии Евклида не выполняются, некоторые ребра изображаются отрезками меньшей длины, одна сторона треугольника может превышать сумму двух других сторон.

Рассмотренный выше метод позволяет получить распределение понятий, обозначающих различные виды звезд. На рис. 5 показан двумерный граф, центральным понятием которого является НП «Вселенная». Видно, что НП «пульсар», «черная дыра» (ЧД), «нейтронная звезда» (НЗ), «белый карлик» (БК) достаточно тесно связаны друг с другом. Семантическое расстояние между НП «Вселенная» и «звезда главной последовательности» (ГП) довольно велико.

Внесение небольших изменений в программу *Graf.pas* позволило перейти от плоских (двумерных) графов к трехмерным. В результате оптимизации расположения вершин в 3D-пространстве получены следующие отцентрированные координаты понятий: Вселенная (–8,93; –1,12; –1,4), галактика (–6,33; –0,62; –1,72), звезда ГП (4,43; –1,22; 1,41), белый карлик – БК (1,57; –1,02; 0,08), нейтронная звезда

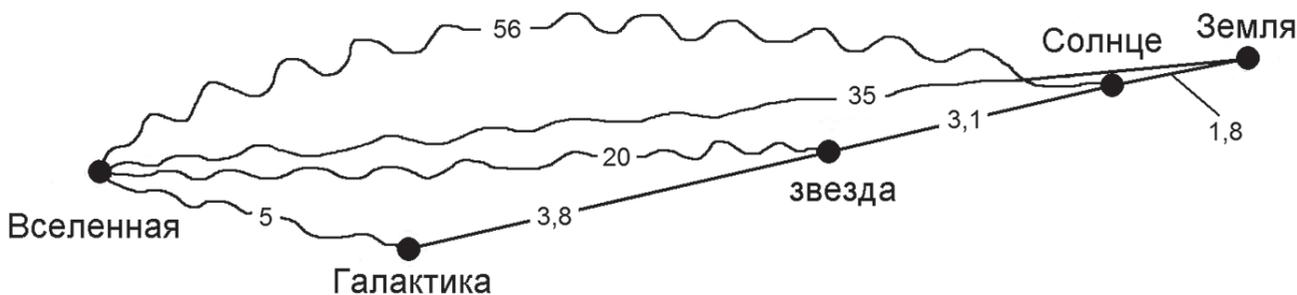


Рис. 4. Связи между понятиями из множества «Земля – звезда – Вселенная»

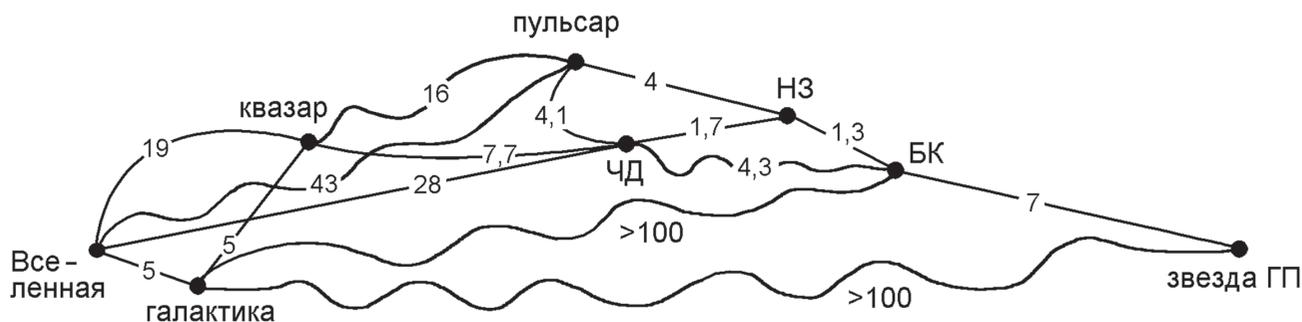


Рис. 5. Семантическое 2D-пространство «Виды звезд»

– НЗ (1,57; 0,08; 0,74), квазар (–4,43; –0,12; –1,72), пульсар (0,67; 1,58; 0,38), черная дыра – ЧД (–0,47; –0,57; 0,23), гигант (1,17; –1,22; 0,06), новая звезда (–2,23; 1,73; –0,29), переменная звезда (12,97; 2,54; 2,25). С помощью онлайн-ресурса *Geogebra* были поставлены вершины в трехмерном пространстве и нарисован трехмерный граф (рис. 6). Близкие вершины, для которых $L_{i,j}$ меньше 9, соединены отрезками, символизирующими короткие жесткие связи.

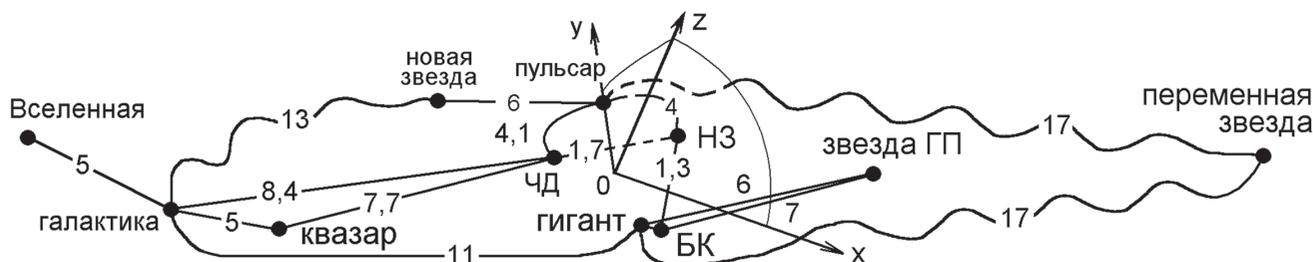


Рис. 6. Семантическое 3D-пространство «Виды звезд»

Выводы

В статье изучается семантическое пространство школьного курса астрономии, образованное понятиями, обозначающими небесные тела. С помощью компьютерных программ анализируются распределения понятий в тексте учебника, вычисляются косинусная мера близости и семантическое расстояние

между понятиями. В результате: 1) получен список часто используемых астрономических понятий, обозначающих небесные тела; 2) рассчитаны матрицы близости и семантических расстояний между понятиями типа «Земля», «звезда», «Солнце», «Юпитер», «черная дыра», «белый карлик», «нейтронная звезда» и т.д.; 3) создана ментальная карта небесных тел, входящих в состав Солнечной системы, учитывающая семантическую близость обозначающих их терминов. В результате оптимизации

распределения понятий в семантическом пространстве с помощью компьютерной программы получены следующие 2D- и 3D-графы: 1) небесные тела Солнечной системы; 2) Земля – Солнце – звезда – галактика – Вселенная; 3) звезды и их виды. Установленные связи и закономерности носят объективный характер.

Литература

1. Анисимов А.В. Метод вычисления семантической близости-связности между словами естественного языка [Текст] / А.В. Анисимов, А.А. Марченко, В.К. Кисенко // Кибернетика и системный анализ. – 2011. – № 4. – С. 18–27.
2. Бранченко Д.С. Меры семантической близости на основе онтологических размерностей понятий в информационных системах [Текст] / Д.С. Бранченко, Ю.А. Кравченко, Ю.С. Новикова // Информатика, вычислительная техника и инженерное образование. – 2017. – № 1. – С. 46–58.
3. Ванюшкин А.С. Методы и алгоритмы извлечения ключевых слов [Текст] / А.С. Ванюшкин, Л.А. Гращенко // Новые информационные технологии в автоматизированных системах. – 2016. – № 19. – С. 85–93.
4. Варламов М.И. Расчет семантической близости концептов на основе кратчайших путей в графе ссылок Википедии [Текст] / М.И. Варламов, А.В. Коршунов // Машинное обучение и анализ данных. – 2014. – Т. 1. – № 8. – С. 1107–1125.
5. Галузо И.В. Астрономия: учеб. пособие для 11-го кл. общеобразоват. учреждений с рус. яз. обучения с 11-летним сроком обучения [Текст] / И.В. Галузо, В.А. Голубев, А.А. Шимбаев. – Минск: Нар. асвета, 2009. – 216 с.
6. Ичкинеева Д.А. О закономерностях дискретизации семантического пространства текста [Текст] / Д.А. Ичкинеева // Вестник ОГУ. – 2010. – № 11. – С. 98–100.
7. Каряева М.С. Лингвостатистический анализ терминологии для построения тезауруса предметной области [Текст] / М.С. Каряева // Моделирование и анализ информационных систем. – 2015. – Т. 22. – № 6. – С. 834–851.
8. Майер Р.В. Сложность учебных понятий и текстов [Текст]: монография / Р.В. Майер. – Глазов: Изд-во ГИПУ, 2024. – 132 с.
9. Мишанкина Н.А. Ментальное пространство научного текста: метафорические модели [Текст] / Н.А. Мишанкина // Вестник Томского государственного университета. – 2007. – № 297. – С. 7–11.
10. Морозова Ю.И. Построение семантических векторных пространств различных предметных областей [Текст] / Ю.И. Морозова // Информатика и ее применения. – 2013. – Т. 7. – Вып. 1. – С. 90–93.
11. Найханов Н.В. Определение семантической близости понятий на основе использования ссылок Википедии [Текст] / Н.В. Найханов, Б.А. Дышенов // Программные системы и вычислительные методы. – 2016. – № 3. – С. 250–257.
12. Панченко А.И. Извлечение семантических отношений из статей Википедии с помощью алгоритмов ближайших соседей [Текст] / А.И. Панченко, С.А. Адейкин, А.В. Романов, П.В. Романов // Открытые системы. – 2012. – № 16. – С. 18–27.
13. Ракитина С.В. Концептосфера и семантическое пространство научного текста [Текст] / С.В. Ракитина // Альманах современной науки и образования. № 8: в 2 ч. Ч. I. – Тамбов: Грамота, 2009. – С. 125–126.
14. Шереметьева С.О. Методы и модели автоматического извлечения ключевых слов [Текст] / С.О. Шереметьева,

- П.Г. Осминин // Вестник ЮУрГУ. Серия: Лингвистика. — 2015. — Т. 12. — № 1. — С. 76–81.
15. Яковлева С.С. Использование ментальных карт в обучении студентов вуза [Текст] / С.С. Яковлева // Scientific Review. — 2019. — № 4. — С. 134–137.
 16. Ashraf A. Semantic similarity measures between words: a brief survey / A. Ashraf, A. Fayez, A. Hani // Scientific International Journal. 2018. № 30(6). Pp. 907–914.
 17. Lin D. An information-theoretic definition of similarity // ICML. 1998. T. 98. Pp. 296–304.
 18. Manning C.D. An Introduction to Information Retrieval / C.D. Manning, P. Raghavan, H. Schütze. Cambridge University Press. 2008.
 19. Turney P.D. From frequency to meaning: Vector space models of semantics / P.D. Turney, P. Pantel // J. Artificial Intelligence Research. Menlo Park, California: AAAI Press, 2010. № 37. Pp. 141–188.
- References**
1. Anisimov A.V. Metod vychislenija semanticheskoy blizosti-sv-jaznosti mezhdu slovami estestvennogo jazyka / A.V. Anisimov, A.A. Marchenko, V.K. Kisenko // Kibernetika i sistemnyj analiz. 2011. № 4. Pp. 18–27.
 2. Branchenko D.S. Mery semanticheskoy blizosti na osnove ontologicheskikh razmernostej ponjatij v informacionnyh sistemah / D.S. Branchenko, Ju.A. Kravchenko, Ju.S. Novikova // Informatika, vychislitel'naja tehnika i inzhenernoe obrazovanie. 2017. № 1. Pp. 46–58.
 3. Vánjushkin A.S. Metody i algoritmy izvlechenija ključevyh slov / A.S. Vánjushkin, L.A. Grashhenko // Novye informacionnye tehnologii v avtomatizirovannyh sistemah. 2016. № 19. Pp. 85–93.
 4. Varlamov M.I. Raschet semanticheskoy blizosti konceptov na osnove kratchajshih putej v grafe ssylok Vikipedii / M.I. Varlamov, A.V. Korshunov // Mashinnoe obuchenie i analiz dannyh. 2014. T. 1. № 8. Pp. 1107–1125.
 5. Galuzo I.V. Astronomija: ucheb.posobie dlja 11-go kl. obshhe-obrazovat. uchrezhdenij s rus. jaz. obuchenija s 11-letnim srokom obuchenija / I.V. Galuzo, V.A. Golubev, A.A. Shimbaev. Minsk: Nar. asveta, 2009. 216 p.
 6. Ichkineeva D.A. O zakonomernostjah diskretizacii semanticheskogo prostranstva teksta // Vestnik OGU. № 11. 2010. Pp. 98–100.
 7. Karjaeva M.S. Lingvostaticeskij analiz terminologii dlja postroenija tezaurusa predmetnoj oblasti // Modelirovanie i analiz informacionnyh sistem. 2015. T. 22. № 6. Pp. 834–851.
 8. Mayer R.V. Slozhnost' uchebnyh ponjatij i tekstov: monografija. Glazov: GIPU, 2024. 132 p.
 9. Mishankina N.A. Mental'noe prostranstvo nauchnogo teksta: metaforicheskie modeli // Vestnik Tomskogo gosudarstvennogo universiteta. 2007. № 297. Pp. 7–11.
 10. Morozova Ju.I. Postroenie semanticheskikh vektornyh prostranstv razlichnyh predmetnyh oblastej // Informatika i ee primenenija. 2013. T. 7. Vyp. 1. Pp. 90–93.
 11. Najhanov N.V. Opredelenie semanticheskoy blizosti ponjatij na osnove ispol'zovanija ssylok Vikipedii / N.V. Najhanov, B.A. Dyshenov // Programmnye sistemy i vychislitel'nye metody. 2016. № 3. Pp. 250–257.
 12. Panchenko A.I. Izvlechenie semanticheskikh otnoshenij iz statej Vikipedii s pomoshh'ju algoritmov blizhajshih sosedej / A.I. Panchenko, S.A. Adejkin, A.V. Romanov, P.V. Romanov // Otkrytye sistemy. 2012. № 16. Pp. 18–27.
 13. Rakitina S.V. Konceptosfera i semanticheskoe prostranstvo nauchnogo teksta // Al'manah sovremennoj nauki i obrazovanija. № 8 (27): v 2-h ch. Ch. I. Tambov: Gramota, 2009. Pp. 125–126.
 14. Sheremet'eva S.O. Metody i modeli avtomaticheskogo izvlechenija ključevyh slov / S.O. Sheremet'eva, P.G. Osminin // Vestnik JuUrGU. Serija: Lingvistika. 2015. T. 12. № 1. Pp. 76–81.
 15. Jakovleva S.S. Ispol'zovanie mental'nyh kart v obuchenii studentov vuza // Scientific Review. 2019. № 4. Pp. 134–137.
 16. Ashraf A. Semantic similarity measures between words: a brief survey / A. Ashraf, A. Fayez, A. Hani // Scientific International Journal. 2018. No. 30(6). Pp. 907–914.
 17. Lin D. An information-theoretic definition of similarity // ICML. 1998. V. 98. Pp. 296–304.
 18. Manning C.D. An Introduction to Information Retrieval / C.D. Manning, P. Raghavan, H. Schütze. Cambridge University Press. 2008.
 19. Turney P.D. From frequency to meaning: Vector space models of semantics / P.D. Turney, P. Pantel // J. Artificial Intelligence Research. Menlo Park, California: AAAI Press, 2010. No. 37. Pp. 141–188.