

Е.В. Сенницкая

ВАРИАНТ РЕШЕНИЯ ПРОБЛЕМЫ ОПИСАНИЯ И ИЗМЕРЕНИЯ СЛОВЕСНОЙ ИНФОРМАЦИИ

Аннотация. Объектом исследования является возможность создания метода описания и измерения информации на естественном языке (в том числе с включением математических и иных формул), который позволил бы решить следующие актуальные задачи: 1) точно указать количество информации, имеющейся в речи или тексте; 2) сделать точными педагогические рекомендации, связанные с учётом сложности изучаемого материала; 3) обеспечить дозирование информации для максимально лёгкого восприятия; 4) исследовать особенности мышления людей, основываясь на информационном анализе их речи или текстов. Создание такого метода моделирования осуществлялось за счёт выделения определённой единицы информации и нахождения способа описания связей между отдельными единицами. В результате анализа работ предшественников был сделан вывод о том, что решение вопроса шло двумя путями: за счёт выделения единицы описания (преимущественно визуально представимого характера) и за счёт представления информации в виде иерархической структуры. Неудачи в решении проблемы были обусловлены тем, что эти подходы не были совмещены. Автором был разработан метод описания словесных массивов информации, названный объектно-иерархическим. Объектным он является потому, что единицей информации служит визуально представимый объект в грамматическом смысле, то есть предмет или явление, о котором можно спросить «что это?». Иерархическим же метод можно назвать по той причине, что единицы информации в связной речи или тексте взаимосвязаны и образуют иерархическую систему. В качестве единицы информации автор предлагает принять описываемую словами, математической или иной формулой визуально представимую картину, заключающую в себе: предмет в грамматическом смысле либо предмет и совершаемое им действие (т.е. явление) либо предмет, совершаемое им действие, и предмет, на который это действие направлено. Предлагаемый метод позволяет дозировать информацию и обеспечивать максимально лёгкое её восприятие. Также он является удобным инструментом для исследования особенностей мышления разных людей за счёт структурного анализа их речи или текстов.

Ключевые слова: структура речи, структура текста, информационная структура, словесная информация, моделирование информации, описание информации, измерение информации, единица информации, моделирование структуры текста, описание структуры текста.

Abstract. The object of study in this article is the ability to create a method for describing and measuring the information in natural language (including texts containing mathematical and the other formulas) that would allow: 1) to determine the amount of information in a speech or text; 2) to make more accurate pedagogical recommendations concerning the complexity of the studied materials; 3) to dose information for readability; 4) to explore the features of people's thinking, based on the information analysis of their speech or text. Creation of such a method was carried out by determination of the information units in a text as well as finding the ways of describing the relations between these units. The analysis of works written by predecessors had concluded that the solution of this issue was going in two ways: by production of the unit of description (especially visually presentable) and by presentation of information in a form of hierarchical structure. Failures in resolving this problem were caused by the fact that these two approaches have not been combined. With this in mind, the author offers an object-hierarchical method for description of amounts of verbal and textual information, which allows to dose the information and is a convenient tool for studies of the features of different people's thinking due to the structural analysis of their speech or text.

Key words: modelling of text structure, unit of information, structure of a speech, structure of a text, information structure, verbal information, information modelling, description of information, measurement of information, description of text structure.

Практическое использование описываемого метода осуществляется в рамках авторского проекта, финансируемого по гранту фонда Бортника. Проект заключается в дозировании информации в учебных программах для максимального облегчения восприятия.

Одним из предметов изучения в психологии является восприятие информации на естественном языке (иногда с включением математических и иных формул) при слушании речи или чтении текста. Тем не менее наука не располагает удобным методом описания и, соответственно, измерения информации на естественном языке, что особенно актуально для педагогической психологии, поскольку педагогическим рекомендациям свойственна расплывчатость. Практически в каждом методическом пособии говорится о том, что материал не должен быть слишком сложным, но, с другой стороны, не должен быть примитивным, но подобные советы не сопровождаются чёткими количественными характеристиками, и непонятно, каким же образом можно обеспечить должное внимание и запоминание.

Чтобы решить эту проблему, необходимо прояснить вопросы о том, что именно следует принять в качестве единицы информации и как описать смысловые взаимосвязи между этими единицами и, соответственно, общую структуру текста или речи человека.

К попыткам выделения единиц информации следует отнести идеи И.-Ф. Гербарта [1], Г. Эббингауза [2], Л. Витгенштейна [3], Д.-А. Миллера [4], Б.-Ф. Скиннера [5], В.П. Беспалько [6]. В информатике единицей информации считается 1 байт (= 8 бит). Для машины количество информации, заключённой в каком-либо тексте, соответствует количеству символов и пробелов, из которых он состоит. Компьютер не учитывает количество смысла в тексте.

В структурных языках программирования, например, Алголе [7], Си и особенно HTML, возникла идея представления информации в иерархическом виде. Первая попытка такого представления учебной словесной информации принадлежит Т. Бьюзену [8]. Правда, подход Бьюзена не содержит ни определения единицы информации, ни каких-либо правил её подсчёта и дозирования [9].

Автором был разработан метод описания словесных массивов информации, названный *объектно-иерархическим*. Объектным он является потому, что единицей информации служит визуально представимый объект в грамматическом смысле, то есть предмет или явление, о котором можно спросить «что это?». Иерархическим же метод является по той причине, что единицы информации в связной речи или тексте взаимосвязаны и образуют иерархическую систему.

При решении вопроса, что же принять за единицу информации, в основу было положено следующее обстоятельство: наш мир сам по себе

является множеством визуально представимых объектов, а также действий, которые эти объекты производят, и за каждым словесным сообщением стоят объекты и их действия, что согласуется с определением информации в ГОСТ 7.0-99: «Информация – сведения, воспринимаемые человеком или специальными устройствами как отражение фактов материального мира в процессе коммуникации» [10].

Например, фраза о том, что «в XIX в. барыни ездили в каретах» включает объект «XIX век», который, в свою очередь, содержит в себе два объекта: «Барыни» и «кареты». Значит, в этой фразе три объекта. Один из них («XIX век») является из-за своей визуальной непредставимости *абстрактным*. Поясним, что *абстрактным* понятием мы называем понятие, требующее его определения (раскрытия для понимания) с помощью других понятий. Введём также такой термин, как *предметное понятие*, то есть понятие, определяющее (раскрывающее смысл) абстрактного понятия. (Само предметное понятие в свою очередь может быть абстрактным по отношению к понятиям, определяющим его). Как показано на рис. 1, абстрактный объект «XIX век» раскрывается, конкретизируется за счёт двух визуально представимых объектов – «барыни» и «кареты».

Теперь рассмотрим, как описать смысловые связи между единицами информации. Обратим внимание, что между абстрактным объектом «XIX век» и предметными объектами «барыни» и «кареты» существует иерархия («барыни» и «кареты» являются элементами, характеризующими объект «XIX век»).

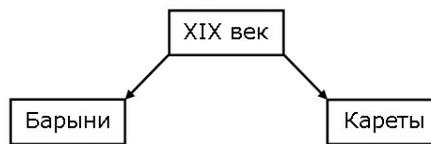


Рис. 1. «XIX век» – объект 1-й ступени, «барыни» и «кареты» – 2-й ступени

«XIX век» – это *объект первой ступени иерархии или, иначе, корневой объект*, а «барыни» и «кареты» – *объекты второго уровня вложенности*.

Спрашивается, куда же делось действие, ведь было сказано, что «барыни ездили»? Ответ на этот вопрос следующий. Каждое действие можно представить в виде последовательности статичных положений объектов, как в мультфильме. Движение в мультфильме создаётся последовательностью картинок, отображающих различные положения тел персонажей. Таким образом, движение как таковое не является самостоятельной единицей ин-

формации, а информация – это набор статичных картинок.

Особо отметим, что чисто оценочные суждения с точки зрения нашего подхода информацией не являются. Оценивая, человек просто «окрашивает» факты в тот или иной цвет: это «чёрное», а это «белое», это хорошо, а это – плохо, однако ничего конкретного не сообщается.

Поскольку в психологии не существует определённого мнения о том, что такое «единица информации», за единицу информации мы принимаем описываемую словами, математической или иной формулой визуально представимую картину, заключающую в себе:

- а) предмет (предметом в грамматике называется всё то, о чем можно спросить «кто это? что это?»),
либо
- б) предмет и совершаемое им действие (т.е. явление),
либо
- в) предмет, совершаемое им действие, и предмет, на который это действие направлено.

В единицу информации может также входить дополнительное наименование предмета (*понятие-синоним*). Примеры: «рок» = «фатум» = «судьба», «стол» = «the table» = «la table», «парламент РФ» = «Федеральное Собрание РФ», «глубокое заседание» = «гран-плие».

При изучении движений единицей информации является статичная фаза движения, которую должен представить обучаемый, чтобы оно было выполнено. Таким образом, единицы информации здесь те же, что и в пунктах а), б), в).

В случае, если в тексте присутствуют числа, формулы, математические расчёты, в качестве единицы информации выступают:

- а) отдельное число или переменная в качестве статичных объектов;
или
- б) математическое действие, поскольку, например, « $x+y$ » – это ни что иное как объект « x », совершающий действие, направленное на другой объект « y ».

Иными словами, если речь идёт о решении математических задач, роль единицы информации играет картина расположения чисел и переменных в каждом действии.

Для оценки сложности как текстового материала, так и математической или иной задачи, содержащей формулы, строится иерархическая структура текста (см. рис. 2) или задачи и производится учёт следующих параметров:

- 1) количество ступеней (глубину иерархии);

- 2) количество предметных объектов, раскрывающих каждый абстрактный объект (степень разветвлённости);

- 3) количество абстрактных объектов, оставшихся не раскрытыми с помощью предметных объектов (для любой задачи, в том числе математической, такими объектами являются неизвестные, которые следует вывести самостоятельно).

Объектом исследования с помощью объектно-иерархического метода является информация, содержащаяся в речи или тексте, в том числе с включением формул. Предметами исследования могут послужить различные параметры восприятия информации слушателем или читателем:

1. Комфортность восприятия учебного материала (т.е. отсутствие стресса при его изучении) в зависимости от иерархической структуры речи или текста.
2. Внимание (и, следовательно, дисциплина).
3. Понимание информации. Поскольку понимание – это включение новой информации в систему известного ранее, то связь понимания с иерархической структурой того материала, который раскрывает новую информацию, очевидна. Качество понимания можно оценить, исходя из степени включённости нового материала в другой, ранее известный материал. Иначе говоря, хорошо понятый факт – это тот, который имеет большое количество связей с уже известными фактами, между тем количество этих связей легко подсчитать, используя объектно-иерархическое представление учебного материала (см. рис. 2).
4. Оценка сложности информации, в частности, оценка сложности задач (в том числе содержащих формулы). Сложность можно оценить с помощью следующих показателей: 1) глубины иерархической структуры (т.е. количества ступеней иерархии); 2) разветвлённости иерархической структуры (т.е. количества ветвей); 3) количества предметных понятий, раскрывающих каждое абстрактное понятие; 4) наличия или отсутствия абстрактных понятий на нижнем уровне иерархической структуры (т.е. понятий, которые остались не раскрыты). Оценка сложности математической задачи осуществляется с помощью тех же показателей, поскольку условие задачи тоже представляет собой иерархическую структуру объектов, один из которых (или несколько) остаются неизвестными (в терминах объектно-иерархического подхода это абстрактные объекты, которые пока не раскрыты). Как и

условие задачи, так и её решение представляет собой иерархическую структуру с определённой глубиной (многоступенчатостью) и разветвлённостью.

5. Оценка степени интересности информации в зависимости от того, как она структурирована. Такой параметр как интересность (увлекательность) близок понятию «внимание», но имеет следующий смысловой оттенок: это степень желанности самостоятельно углублять свои знания в какой-то области или степень желанности решить задачу.
6. Запоминание и забывание. Объектно-иерархический метод является инструментом для исследования вопроса о том, какая структурная организация речевого или текстового материала способствует наилучшему его запоминанию, а какая ведёт к забыванию тех или иных фрагментов, имеющих определённое положение в иерархически организованной информации. Объектно-иерархический метод, в частности, способен помочь в разработке учебных пособий и поурочных планов. Что касается изучения процесса забывания отдельных фрагментов информации, выпадения их из сознания, это может быть интересно для исследования механизмов внушения.
7. Различия в восприятии информации людьми с разными способностями (например, классом коррекции, обычным и продвинутым).
8. Различия в восприятии гуманитарной и математической информации.

В случае, если изучается не процесс восприятия информации, а, напротив, некий интеллек-

туальный продукт, можно выделить следующие предметы исследования:

1. Особенности речи и письменного творчества психически больных людей, характерные для каждого типа заболевания. Такими особенностями могут быть преобладание линейных структур или, напротив, чрезмерно сложных иерархических конструкций, избыток абстрактных (т.е. не раскрываемых в тексте или в речи) понятий или, наоборот, диспропорция в сторону предметных понятий и т.д. Обучение методу структурирования информации может послужить новым средством реабилитации больных с нарушениями логики, т.к. появляется возможность показать им наглядно, чем их речь и тексты отличаются от таковых у здоровых людей.
2. Особенности мышления людей, живших в другие исторические эпохи или людей, относящихся к разным национальным культурам, исходя из объектно-иерархического анализа их речи и письма.
3. Творчество известных писателей с точки зрения особенностей объектно-иерархической структуры их произведений.

По некоторым из указанных направлений автором проведены исследования, конкретные количественные результаты которых будут впоследствии опубликованы.

Таким образом, объектно-иерархический метод имеет не только важное практическое, но и теоретическое значение, поскольку служит новым удобным инструментом для изучения психики и сближает психологию с точными дисциплинами.

Информационная структура урока зоологии (см. рис. 2)

Корневой объект «Особенности млекопитающих»	раскрыт с помощью: 5 объектов уровня 1
Объекты 1-го уровня:	
объект 1 «Особенности внешнего строения млекопитающих»	5 объектов уровня 2
объект 2 «Кожа млекопитающих»	5 объектов уровня 2
объект 3 «Роговое вещество у млекопитающих»	4 объектов уровня 2
объект 4 «Скелет млекопитающих»	3 объектов уровня 2
объект 5 «Основные мышцы млекопитающих»	5 объектов уровня 2
Объекты 2-го уровня:	
объект 2.4 «Потовые железы млекопитающих»	2 объектов уровня 3
объект 4.1 «Отделы скелета млекопитающих»	4 объектов уровня 3
объект 4.2 «Особенности скелета млекопитающих»	5 объектов уровня 3
объект 4.3 «Функции скелета»	3 объектов уровня 3
Объекты 3-го уровня:	
объект 4.1.2 «Отделы позвоночника млекопитающих»	5 объектов уровня 4
объект 4.1.3 «Пояс передних конечностей»	3 объектов уровня 4
объект 4.1.4 «Пояс задних конечностей»	4 объектов уровня 4
Объекты 4-го уровня:	
объект 4.1.2.2 «Грудной отдел позвоночника»	2 объектов уровня 5
Объекты 5-го уровня:	
объект 4.1.2.2.2 «Ключицы»	2 объектов уровня 6

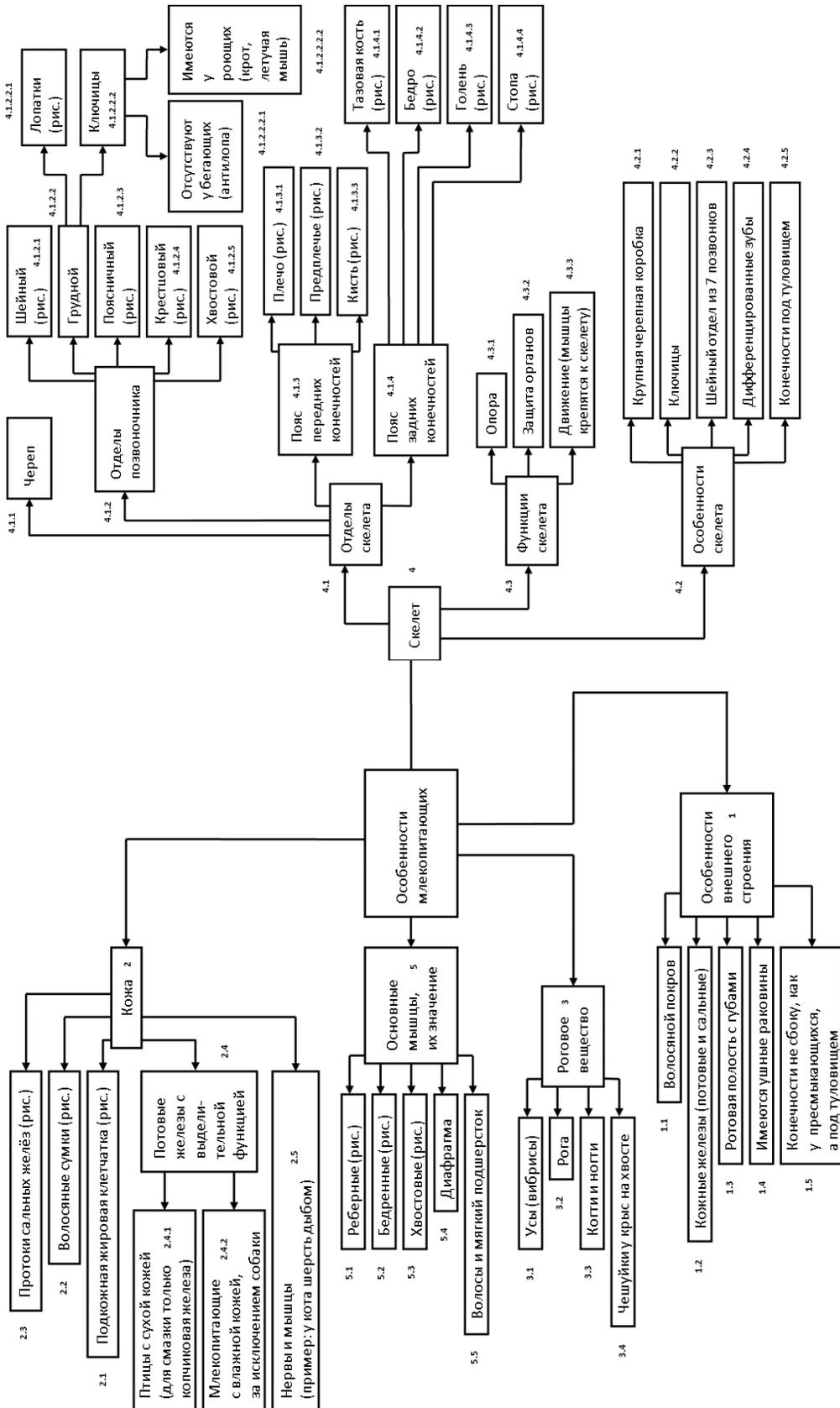


Рис. 2. Схема урока зоологии. Составлена автором

Список литературы:

1. Гербарт И.-Ф. Психология. М.: Территория будущего, 2007. 280 с.
2. Эббингауз Г. Основы психологии. СПб., 1912. 268 с.
3. Витгенштейн Л. Логико-философский трактат. М.: АСТ, Астрель, Мидгард, 2010. 192 с.
4. Миллер Д.-А. Магическое число семь плюс или минус два. Инженерная психология. М.: Прогресс, 1964. С. 192-225.
5. Программированное обучение // Available at. URL: <https://ru.wikipedia.org>.
6. Беспалько В.П. Программированное обучение. Дидактические основы. М.: Высшая школа, 1970. 300 с.
7. Алгол // Available at. URL: <https://ru.wikipedia.org>.
8. Бьюзен Т., Бьюзен Б. Супермышление. Минск: Попурри, 2008. 208 с.
9. Buzan, T. iMindMap8 // Available at. URL: <http://thinkbuzan.com/download/>.
10. ГОСТ 7.0-99. Система стандартов по информации, библиотечному и издательскому делу. Информационно-библиотечная деятельность, библиография. Термины и определения. М., 1999. 28 с.

References (transliterated):

1. Gerbart I.-F. Psikhologiya. M.: Territoriya budushchego, 2007. 280 s.
2. Ebbingauz G. Osnovy psikhologii. SPb., 1912. 268 s.
3. Vitgenshtein L. Logiko-filosofskii traktat. M.: AST, Astrel', Midgard, 2010. 192 s.
4. Miller D.-A. Magicheskoe chislo sem' plyus ili minus dva. Inzhenernaya psikhologiya. M.: Progress, 1964. S. 192-225.
5. Programmirovannoe obuchenie // Available at. URL: <https://ru.wikipedia.org>.
6. Bepal'ko V.P. Programmirovannoe obuchenie. Didakticheskie osnovy. M.: Vysshaya shkola, 1970. 300 s.
7. Algol // Available at. URL: <https://ru.wikipedia.org>.
8. B'yuzen T., B'yuzen B. Supermyshlenie. Minsk: Popurri, 2008. 208 s.
9. Buzan, T. iMindMap8 // Available at. URL: <http://thinkbuzan.com/download/>.
10. GOST 7.0-99. Sistema standartov po informatsii, bibliotechnomu i izdatel'skomu delu. Informatsionno-bibliotechnaya deyatel'nost', bibliografiya. Terminy i opredeleniya. M., 1999. 28 s.