

## ИСЧИСЛЕНИЕ ТРУДНОСТИ СОДЕРЖАНИЯ И ПОНИМАНИЯ ТЕКСТА

Гидлевский А. В., Здриковская Т. А.

*ФГБОУ ВПО «ОмГУ им. Ф. М. Достоевского», Омск, Россия (644077, г. Омск, проспект Мира, д. 55-А), e-mail: gidlevsky@yandex.ru*

В данном сообщении предлагается метод измерений трудности содержания и трудности понимания текста. В основе метода находится субъект-предикатный подход, согласно которому раскрытие содержания главного (начального) текстового субъекта представляет собой структурную сеть (граф), в узлах которой находятся подчиненные текстовые субъекты. При этом трудность содержания текста сводится к суммарной трудности определений текстовых субъектов. Трудность определения текстового субъекта равна произведению исходной трудности, величина которой может быть принята равной единице для всех субъектов, на ряд коэффициентов, зависящих от места субъекта в их последовательности, типа связей данного субъекта, трудности поиска в базах знаний и других характеристик. Затем суммируются показатели трудности для всех линий раскрытия, входящих в задачу определения трудности содержания соответствующего фрагмента текста. В свою очередь, трудность понимания текста вычисляется посредством суммирования трудностей определений нераскрытых текстовых субъектов и их модификатов.

Ключевые слова: субъект-предикатный подход, трудность содержания текста, трудность определения текстового субъекта, коэффициент иерархичности, коэффициент модифицируемости.

## CALCULATION OF TEXT CONTENT AND UNDERSTANDING DIFFICULTIES

Gidlevisky A. V., Zdrikovskaja T. V.

*FGBOU VPO "Dostoyevsky OMSU", Omsk, Russia (55A Mira Avenue, Omsk, 644077), e-mail: gidlevsky@yandex.ru*

The paper offers the method to calculate content difficulty and understanding difficulty of a text. The method is based on the subject-predicate approach according to which the disclosure of the main (initial) text subject is presented as a structural net (graph) in the junctions of which there are subordinate text subjects. The difficulty of the text content comes to the total difficulty of text subjects definitions. The difficulty of a text subject definition is equal to the product of the initial difficulty the value of which may be accepted as 1 for all subjects and the number of coefficients depending on the place of a subject in their sequence, types of connections of this subject, difficulty of search in knowledge databases and others. Then all the sums included in the task of the definition of content difficulty in a passage are summed up. In its turn the difficulty of text understanding is calculated by summarizing the difficulties of undisclosed text subjects and their modifiers.

Key words: subject-predicate approach, difficulty of text content, difficulty of text subject definition, hierarchy coefficient, modifiability coefficient.

### Введение

В нашем сообщении обсуждается универсальная концепция текста, позволяющая эффективно использовать современные модели представления знаний. Универсальность предлагаемой концепции определяется ее соответствием моделям переработки информации естественным и искусственным интеллектом и широкими эвристическими возможностями для создания эффективной системы проектирования и диагностики учебно-методического обеспечения образования.

Некоторым препятствием концептуального плана в развитии предлагаемой концепции служит имеющее место в современной дидактике разночтение понятий сложности и трудности содержания текста. Мы имеем в виду алгоритмическую, семантическую трудность содержания, которую можно условно назвать когнитивной сложностью. Однако, по нашему

мнению, как только произносится слово «когнитивный», сложность автоматически становится трудностью. Нами в данной работе используется термин «трудность содержания» по нескольким причинам: во-первых, этот термин используется рядом исследователей – А. Н. Колмогоровым, Р. А. Гильмановым, Н. П. Быковой, Н. Г. Рыженко и др. [1, 3, 5, 6]. Во-вторых, поскольку трудность понимания (решения) логико-гностической задачи связана с синтаксическими и семантическими характеристиками алгоритмов, целесообразно объединить эти характеристики в другом понятии – «трудности содержания». В-третьих, понятие сложности, по мнению ряда исследователей, включает в себя количества действий, объектов, процессов и др., то есть, существенно разноразмерных, разновесных и разнородных величин, что вызывает у авторов данной статьи ряд вопросов.

Ниже мы рассмотрим метод количественной оценки трудности содержания и трудности понимания текста. В основе метода находится субъект-предикатный подход к установлению смысловой структуры содержания текста. При этом структура содержания текста отображается с помощью графа, узлами которого являются текстовые субъекты. Трудность содержания текста на том или ином уровне раскрытия текстовых субъектов, находящихся в узлах графа, вычисляется через сумму величин трудности определений субъектов, занимающих соответствующие уровни раскрытия. Трудность понимания текста оценивается суммой величин трудности определений нераскрытых индивидом текстовых субъектов и их модификатов.

### **Трудность содержания текста**

В соответствии с субъект-предикатным подходом структура содержания текста представляется иерархией субъектов, в которой раскрытие главного (начального) текстового субъекта – предмета высказывания происходит посредством его замещения (модификации) системой субъектов нижележащих рангов [4]. Некоторые текстовые субъекты не имеют модификаций и являются конечными (терминальными). Таким образом, система субъектов текста оказывается системой раскрытия начального (главного) субъекта.

Согласно нашей концепции, следует различать трудность содержания текста и трудность его понимания. Расчет трудности содержания текста может быть проведен следующим образом. На основании структурной схемы (графа) выделяются линии модификации, для каждой из которых определяется трудность как сумма трудностей (трудностей определений) учитываемых субъектов. Трудность определения субъекта равна произведению исходной трудности, величина которой может быть принята равной единице для всех субъектов, на ряд коэффициентов, зависящих от места субъекта в их последовательности, типа связей данного субъекта, трудности поиска в базах знаний и других характеристик. Затем суммируются по-

казатели трудности для всех линий раскрытия, входящих в задачу определения трудности содержания соответствующего фрагмента текста.

Таким же образом определяется трудность содержания узлов структурной сети, в качестве которых фигурируют текстовые субъекты.

Пусть имеем субъект  $S_0$ , структура раскрытия которого показана на рис. 1. Как видно из рисунка 1, субъект  $S_0$  является главным, и трудность его содержания будет равна сумме трудностей его модификатов (субъектов нижележащих рангов). К примеру, субъект  $S_1$  принадлежит к первому рангу. Главному субъекту, таким образом, нами назначен нулевой ранг.

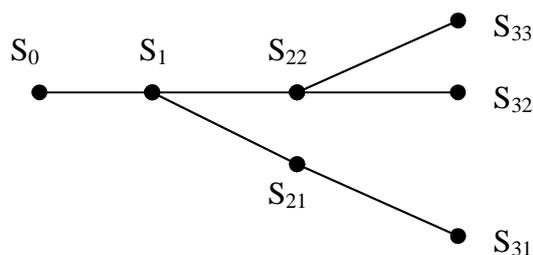


Рис. 1. Структура содержания текстового субъекта  $S$

Искомая трудность содержания приведенного на рис. 1 условного текста равна сумме трудностей определений всех субъектов, кроме главного. Как мы уже говорили выше, трудность того или иного субъекта (трудность определения субъекта)  $T_i$  может быть вычислена как произведение исходной трудности  $T_0$ , которую удобно положить раной единице, и ряда коэффициентов, характеризующих субъект с точки зрения его «когнитивных» (когнитивно-алгоритмических) характеристик – места субъекта в иерархии, характера его связи с другими субъектами, степени знакомства читателю и др. То есть, здесь мы имеем дело именно с «когнитивной» трудностью раскрытия субъекта. Для простоты примера определим два коэффициента, характеризующие субъекты, – коэффициент иерархичности ( $k_i$ ) и коэффициент модифицируемости ( $k_m$ ). Прежде всего, отметим, что на рисунке 1 субъекты  $S_{31}$ ,  $S_{32}$  и  $S_{33}$  являются терминальными, трудность определения которых минимальна. В ряде случаев трудностью их определения можно пренебречь. В том случае, когда мы не хотим пренебрегать терминальными субъектами, положим трудность их определений равной единице для любого, в том числе, составного терминального субъекта. При этом для терминальных субъектов величины коэффициентов  $k_i$  и  $k_m$  должны быть равны единице. Для субъектов более высокого ранга (не терминальных), например  $S_{21}$  на рисунке 1 коэффициент иерархичности равен двум; для субъекта  $S_1$  – трем. Подобная иерархическая неравноценность субъектов объясняется просто. Последовательность раскрытия главного субъекта – это последовательность действий в решении мыслительной задачи. Известно, что при решении задачи первый шаг –

самый трудный, все остальные шаги фактически следуют из первого. То есть, коэффициент иерархичности – это сугубо «трудностный», когнитивный параметр. То же самое мы можем сказать и о коэффициенте модифицируемости для того или иного субъекта. Коэффициент модифицируемости субъекта определяется количеством его непосредственных модификатов: чем их больше, тем выше значение данного коэффициента. Если мы учитываем один модификат, то величина коэффициента модифицируемости данного субъекта равна двум (единица уже занята терминальными субъектами). Если терминальных модификатов (терминалов) у данного субъекта несколько, то коэффициент модифицируемости данного субъекта также можно положить равным двум. В таком случае модификация в несколько терминалов – это модификация в более сложный терминал. В случае двух непосредственных модификатов (не терминальных) величина  $k_m$  равна четырем (например, для субъекта  $S_1$  на рис. 1) и т.д.

Для вычисления трудности показанной на рисунке системы субъектов составим таблицу 1.

Таблица 1 к рис. 1

Субъекты	$k_i$	$k_m$	$T_i$
$S_1$	3	4	12
$S_{21}$	2	2	4
$S_{22}$	2	2	4
$S_{31}$	1	1	1
$S_{32}$	1	1	1
$S_{33}$	1	1	1

Таблица 1. Значения коэффициентов и трудности определения для субъектов, показанных на рис. 1

Суммируя значения в последнем столбце таблицы 1, получим для трудности содержания выбранного условного фрагмента текста значение  $T = 23$ .

### Уровни раскрытия содержания текста

Уровни раскрытия содержания текста мы связываем с уровнями раскрытия содержания концептов, формирующих текст. Для иллюстрации сказанного обратимся к другому примеру – линии модификации  $S_0S_1S_2S_3$ , дополненной субъектами, раскрывающими содержание субъектов  $S_1$  и  $S_2$  (рис. 2).

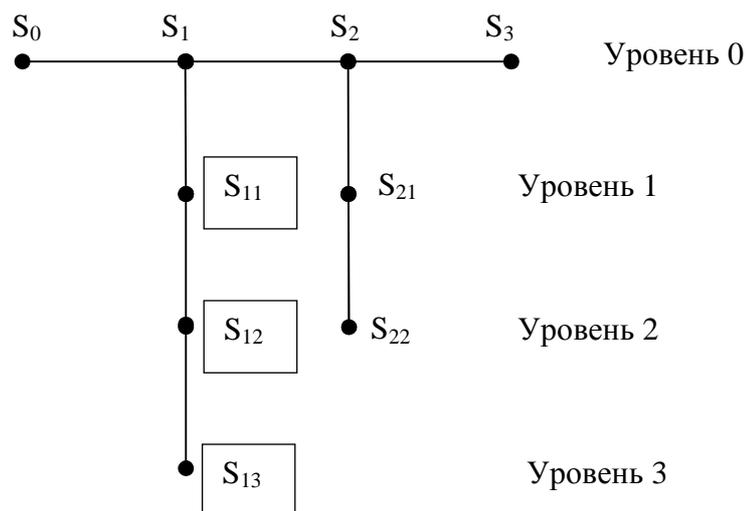


Рис. 2. Граф условного текста с учетом содержания субъектов  $S_2$  и  $S_1$ . В рамки заключены обозначения нераскрытых текстовых субъектов

Как видно из рис. 2, субъекты  $S_{11}$  и  $S_{12}$  соответствуют первому уровню раскрытия, субъекты  $S_{21}$  и  $S_{22}$  – второму и т.д.

В таблице 2 приведены значения коэффициентов, трудности определений и уровней раскрытия для субъектов, показанных на рис. 2.

Таблица 2 к рис. 2

Субъекты	$k_i$	$k_m$	$T_i$	Уровень раскрытия
$S_1$	3	2	6	0
$S_2$	2	2	4	0
$S_3$	1	1	1	0
$S_{11}$	3	2	6	1
$S_{21}$	2	2	4	1
$S_{12}$	2	2	4	2
$S_{22}$	1	1	1	2
$S_{13}$	1	1	1	3

Таблица 2. Значения коэффициентов, трудности определения и уровней раскрытия для субъектов, показанных на рис. 2

Из таблицы 2 видно, что трудность содержания на последующем уровне раскрытия равна трудности содержания на предыдущем уровне плюс трудность определения субъектов на исследуемом уровне.

### Трудность понимания текста

Трудность понимания текста индивидом является скрытым параметром и поэтому непосредственно измерена быть не может. Однако известно, что трудность в решении той

или иной мыслительной задачи может быть определена через неуспешность в решении задачи. Например, чем больше задач из какого-либо задания не решил учащийся, тем труднее для него оказалось это задание. Аналогично, чем меньше действий в решении задачи мог совершить учащийся, тем труднее для него эта задача. Понимание текста также сводится к решению ряда мыслительных задач по раскрытию содержания текста. Раскрыть содержание текста в таком «задачном» подходе – значит, в первую очередь, раскрыть содержание текстовых субъектов. Другими словами, трудность понимания текста индивидом пропорциональна трудности нераскрытых им текстовых субъектов. Если не раскрыты те или иные субъекты, то не раскрыты и их модификаты (субъекты нижележащих уровней). В самом простом случае трудность понимания текста вычисляется посредством суммирования трудностей определений нераскрытых текстовых субъектов и их модификатов.

Обратившись к таблице 2, мы видим, что нераскрытыми являются текстовые субъекты  $S_{11}$ ,  $S_{12}$  и  $S_{13}$ , суммарная трудность определений которых равна 11. Это и есть значение искомой трудности понимания выбранного фрагмента текста.

### **Заключение**

Предлагаемый метод исчисления трудности содержания и понимания текста применим как для конструирования текстов, так и для их диагностики на предмет глубины раскрытия содержания, понимаемости и др. Данный метод представляется эффективным и для конструирования диагностических материалов в образовании. В частности, метод может быть успешно применен к тестам по слабо формализованным дисциплинам – географии, биологии, русскому и иностранным языкам, литературе, истории и др. Тестовые задания для указанных дисциплин содержат пропущенные концепты (задания открытого типа), по успешности определений которых можно оценить трудность понимания текста тем или иным индивидом.

Что же касается тестов по высоко формализованным дисциплинам – математике, физике, химии и информатике, то трудность их содержания может быть оценена как сумма трудностей решений соответствующих задач [2].

### **Список литературы**

1. Быкова Н. П. Методика формирования обобщенных умений на основе моделирования задач физики и математики // Омский Научный Вестник. – 2004. – № 3 (28). – С. 225–229.
2. Гидлевский А. В. Универсальный метод оценки трудности учебных тестовых заданий // Вестник Омского университета. 2011. – № 2. – С. 271–276.

3. Гильманов Р. А. Проблемы конструктивной дидактики. – Казань: Изд-во Казан. ун-та, 1994. – 179 с.
4. Доблаев Л. П. Смысловая структура учебного текста и проблемы его понимания. – М.: Педагогика, 1982. – 176 с.
5. Колмогоров А. Н. Теория информации и теория алгоритмов. – М.: Наука, 1987. – 304 с.
6. Рыженко Н. Г. Сложность и трудность структуры решения текстовой задачи // Электронный научный журнал «Вестник Омского государственного педагогического университета». – Выпуск 2006. URL: <http://www.omsk.edu/volume/2006/methodics/> (Дата обращения: 11.12.2012).

**Рецензенты:**

Орлова Наталья Васильевна, доктор филологических наук, профессор кафедры русского языка ОмГУ им. Ф. М. Достоевского, г. Омск.

Бутакова Лариса Олеговна, доктор филологических наук, зав. кафедрой русского языка ОмГУ им. Ф. М. Достоевского, г. Омск.

Стернин Иосиф Абрамович, доктор филологических наук, профессор, чл.-корр. РАН, зав. кафедрой общего языкознания и стилистики, директор Центра коммуникативных исследований ВГУ, Воронежский государственный университет, г. Воронеж.