УДК 519.767.2

ОЦЕНКА ПОНИМАНИЯ СМЫСЛА ЕЯ КОМПАРАТОРНЫМ МЕТОДОМ

КАЛИНОВСКИЙ А.С., РУБЛИНЕЦКИЙ В.И., РЯБОВА Н.В.

(*Компьютерные науки*)

Рассматривается проблема оценки понимания смысла ЕЯ различными компьютерными системами, предназначенными для обработки ЕЯ текстов. Излагается методика компараторной идентификации на основе тест-текстов (ТТ) для оценивания меры смысловой схожести текстов. Предлагаются три алгоритма, позволяющие на множестве простых ТТ надежно распознавать тексты с одинаковым смыслом.

1. Постановка проблемы

Когда Бог решил пресечь строительство вавилонской башни, он сделал так, что строители начали разговаривать на разных языках. Сегодняшняя вавилонская башня - самая трудная задача, решаемая человеческой наукой и техникой, - это проблема понимания, точнее – проблема понимания компьютером человеческого естественного языка (ЕЯ). Со времени постановки задачи прошло полстолетия, быстродействие и память машин увеличились на много порядков, во многих областях интеллектуальной деятельности машина давно обогнала человека, а в башне проблемы понимания едва выстроен первый этаж. Причина, в основном, состоит в трудности самой задачи, но также и в том, что строители упорно говорят на разных языках.

Специальная литература предлагает много программных систем с благозвучными названиями (АИСТ, ПОЭТ, FAUSTUS, Элиза и т.п.), которые в какой-то мере понимают текст. Было бы очень удобно знать, в какой действительно *мере* они это делают. Тогда было бы легче ориентироваться, какие системы сильнее и какие подходы обещают больше. Не зря отец современной науки Галилей советовал: «Измеряй все измеримое и делай неизмеримое измеримым».

Многие из разработанных в интересующей нас области систем узко направлены на решение специальных задач. Так, система Винограда [1] понимает подъязык, описывающий манипуляции с несколькими фигурами на экране компьютера; она разумно уточняет неясные и отвергает невыполнимые команды, правильно манипулирует с фигурами. Программа Элиза, созданная Вейценбаумом [2], подражает речам психотерапевта. Программа Командина [3] извлекает смысл из объявлений о купле, продаже, съеме и сдаче квартир, представляя смысл объявления точкой в многомерном пространстве признаков. Полный список таких программ спецназначения имел бы длину средней статьи. Многие из этих систем узкого назначения вполне удачны. Так, девушки, которые сами писали отдельные процедуры системы Элиза, оставались после работы, чтобы, переговорив с Элизой, облегчить душу. К сожалению, непонятно, как сравнивать силу систем разной специализации.

Однако существует много систем, претендующих на универсальность. Их, казалось бы, можно натравить на один и тот же экзаменационный материал и таким образом сравнить по эффективности. Но здесь есть свои трудности: системы рассчитаны на разные языки, одни лучше понимают, другие лучше синтезируют ответ и т.д. Проблему языка, естественно, решать нам, оставляя родной и английский, как самый распространенный в такого рода исследованиях. Что касается разных стадий обработки текста, то мы остановимся на самой трудной из них - понимание текста.

Итак, мы поставили вопрос, которому посвящена эта работа: КАК СРАВНИВАТЬ ПО ЭФФЕКТИВНОСТИ (и в этом смысле измерять) РАЗНЫЕ ПРОГРАММНЫЕ СИСТЕМЫ, КОТОРЫЕ ЗАНИМАЮТСЯ (может быть, наряду с другими задачами) ПОНИМАНИЕМ ТЕКСТОВ НА РУССКОМ И/ИЛИ АНГЛИЙСКОМ ЯЗЫКАХ?

2. Разные методы оценки понимания

Перечислим наиболее распространенные методики оценки меры понимания текста.

А. Оценка переводом на другой язык

Пусть, скажем, программа машинного перевода (МП) с английского на русский правильно перевела предложение

I saw the table (1)

как

Я видел стол (2)

Хорошо ли поняла текст (1) эта программа? Такой перевод может сделать даже тривиальная программа пословного перевода с учетом частоты слов: она переведет *saw* как *видел*, а не как *пила*, потому что первое значение имеет большую частоту; по этой же причине она переведет *table* как *стол*, а не как *таблица*. Если повезло и *я* мужского рода, то перевод хорош, хотя программа совсем лишена понимания.

Второй недостаток перевода как мерила понимания состоит в том, что качество перевода существенно зависит от умения синтезировать ответ, а это - преодоление трудности другой природы.

Наконец, использование перевода как мерила понимания предусматривает умение оценивать качество перевода, а это весьма трудное дело. Перевод всегда неточен. Это подтверждено блестящим экспериментом [4], который был устроен так. Все начиналось с текста T_0 на языке 0. Переводчик 1 переводил его в текст T_1 на языке 1. Этот текст давали переводчику 2, который переводил его в текст T_2 и так далее 30 раз. Последний раз текст переводился снова на язык 0. Все переводчики были профессионалы. В результате текст (отрывок из Гоголя):

Он хотел взглянуть на прыщик, который вчерашнего вечера вскочил у него на носу; но, к величайшему изумлению, увидел, что у него вместо носа совершенно гладкое место (3)

превратился в

Вглядываясь с маяка в даль, она выронила трубу, узнав остатки его фрегата (4)

Непонятно, как мерить смысловую погрешность перевода и, следовательно, как мерить погрешность понимания.

B. Оценка через вопросно-ответные системы

Простейшая оценка получается введением текста, после чего задаются вопросы по тексту типа «Дед репку что?» Большинство правильных ответов здесь можно получить механической рекомбинацией слов исходного текста.

Использование более сложных диалоговых систем (типа Элизы) не спасает положения. Однажды произошел курьезный случай, когда вице-президент крупной американской компании по производству ЭВМ, сам того не ведая, реализовал тест Тьюринга. Как-то ему понадобилось связаться с известным системным программистом Д. Бобровым, работающим в этой компании. Стандартная связь по телетайпу привела к следующей беседе [5, с. 131-132]:

> *«Вице-президент: Нельзя ли мне сейчас воспользоваться ЭВМ?*
> *Бобров: А почему Вы спрашиваете?*
> *Вице-президент: Может быть, мне удастся заключить сделку.*
> *Бобров: Почему Вы в этом уверены?*
> *Вице-президент: Мои предполагаемые покупатели никогда не видели нашу систему в действии.*
> *Бобров: Это имеет для Вас большое значение?*
> *Вице-президент: Конечно.*
> *Бобров: Вы как будто сомневаетесь в этом?*
> *Вице-президент: Что за нелепое замечание!*
> *Бобров: Что Вы имеете в виду, говоря о нелепом замечании?*
> *Вице-президент: Возражаете ли Вы против того, чтобы я сейчас воспользовался ЭВМ?*
> *Бобров: Расскажите мне побольше о Ваших предполагаемых покупателях.*
> *Вице-президент: Позвоните мне 491-1850!»*

Недоразумение выяснилось лишь позднее, когда вице-президент стал отчитывать своего сотрудника. Оказалось, что телетайп соединил его не с самим Бобровым, а с ЭВМ, на которой была реализована экспериментальная диалоговая программа, осуществляющая, подобно Элизе, фатический диалог, т.е. диалог, когда следующая ремарка строится вероятностно-механическим преобразованием предыдущей.

Возможность фатического диалога делает вопросно-ответные системы негодным мерилом понимания смысла.

C. Оценка по действиям

Правильность выполнения действий в манипуляторных системах так же мало годится для оценки понимания смысла, поскольку такие программы используют неестественно узкое подмножество ЕЯ. Кроме того, ошибки непонимания бывают разной величины, которую непонятно, как измерять.

Нужен какой-то другой подход к сравнению эффективности систем рассматриваемого вида. Мы предлагаем такой подход.

3. Сравнение понимания тест-текстов (ТТ)

Предлагаемая методика называется компараторным методом [6] и состоит в том, что оцениваемой системе предлагаются пары текстов на ЕЯ, одни из которых схожи по смыслу, а другие – нет. Эффективность системы оценивается по числу ошибок. Очевидно, ошибки бывают двух родов: первый – сходство утверждается на паре несхожих текстов и второй – сходство отрицается на паре схожих текстов. Ошибки не равносильны – сделать ошибку первого рода труднее, чем второго. Поэтому оценка эффективности системы на p парах:

$$E = \sum_{k=1}^{p} c_1 e_k^{(1)} + c_2 e_k^{(2)}, \qquad (5)$$

где c_1, c_2 – веса ошибок первого и второго рода, а $e_k^{(1)} = (0, 1)$ и $e_k^{(2)} = (0, 1)$ – число ошибок первого и второго рода в k-м сравнении.

Сразу встает вопрос, какого вида должны быть ТТ, и кто и как будет определять их схожесть по значению. Нам представляется, что по виду ТТ должны быть краткими (установление схожести кратких текстов труднее, чем длинных). Далее, ТТ должны быть приблизительно одинаковыми по размеру, чтобы тривиальный признак размера не помогал в сравнении. Желательно также, чтобы ТТ были самозамкнутыми: тогда не возникают дополнительные трудности связи с объемлющим контекстом. Что касается схожести текстов по смыслу, то это должны быть группы (удобнее – пары) разных переводов одного и того же оригинала или пары оригинал-перевод.

Приведем пример набора из шести ТТ, образующих три пары схожих текстов. Это первые три (из 44) афоризма Ницше из книги «Сумерки идолов» в переводах Н.Полилова (номера без штрихов) и Г.Снежинской (номера со штрихами).

Праздность есть мать всей психологии. Как? Разве психология – порок? (6)

Праздность – мать всей психологии. Да ну? Неужто психология – порок? (6’)

И самый мужественный из нас лишь редко обладает мужеством на то, что он собственно знает… (7)

Порой и у самых отважных не хватает отваги на то, что им известно... (7’)

Чтобы жить в одиночестве, надо быть животным или Богом, говорит Аристотель. Не хватает третьего случая: надо быть и тем, и другим – философом. (8)

Чтобы пребывать в одиночестве, надо быть животным или богом, утверждает Аристотель. Отсутствует третий случай: быть и тем, и другим – философом. (8')

Приведенные тексты – очень легкие, их можно успешно решить простым формальным алгоритмом, который, даже не вникая в структуру предложения, выполняет подсчет общих слов. Мы имеем в виду следующий

Алгоритм 1

1. Ввести очередную пару текстов T_i и T_k; удалить все союзы, частицы, междометия, предлоги как неинформативные слова, часто встречающиеся во всех текстах; привести все изменяемые слова к канонической форме.

2. Из оставшихся слов в T_i и T_k образовать соответственно два множества M_i и M_k (учитывая, однако, разные словоупотребления одного слова как разные элеметы). Пусть $m_i = |M_i|$ и $m_k = |M_k|$. Пусть m_{ik} – число элементов в пересечении $M_i \cap M_k$, а m_{ik} – средний размер множества: $m_{ik} = (m_i + m_k)/2$. Мера пословного совпадения r_{ik} задается формулой:

$$r_{ik} = n_{ik} / m_{ik} . \quad (9)$$

{Комментарий: r_{ik} меняется от 0, когда нет ни единого общего слова в M_i и M_k, до 1, когда $M_i = M_k$}.

3. Если $r_{ik} \geq h$ (h – некий порог, численное значение которого будет оценено позднее), то $T_i = T_k$ (T_i сходно по значению с T_k), иначе $T_i \neq T_k$.

4. Если не все пары исчерпаны, перейти к п.1.

Пример 1. Пропустим через Алгоритм 1 пару (6), (6'). Получим:

M(6) = {праздность, есть, мать, весь, психология, как, психология', порок},

M(6') = {праздность, мать, весь, психология, как, психология', порок};

m (6, 6') = 7; n (6, 6') = 6; r (6, 6') = 6/7 ≈ 0,86.

Конец примера.

Считая r_{ik} для всех шести пар, получаем следующую (симметрическую) матрицу:

	6	6'	7	7'	8	8'
6	1	0,86	0	0	0	0
6'	0,86	1	0	0	0	0
7	0	0	1	0,32	0,07	0,08
7'	0	0	0,32	1	0,08	0,08
8	0	0	0,07	0,08	1	0,65
8'	0	0	0,08	0,08	0,65	1

Элемент r (7, 7'), выражающий пословное совпадение схожих текстов, достаточно мал. Однако если усилить Алгоритм 1, определяя пересечение не как множество одинаковых слов, а как множество либо одинаковых слов, либо синонимов (что устанавливается формально – введением словаря синонимов), то r (7, 7') возрастет до 0,53. Алгоритм 1, усиленный узнаванием синонимов, назовем **Алгоритмом 2**.

Приведенные и многие другие вычисления показывают, что для Алгоритма 1 годится значение порога h = 0,2, а для Алгоритма 2 – h = 0,4; при этих значениях порога надежно распознается смысловое сходство отрывков прозы.

Тексты для сравниваемых пар можно также подбирать из разных языков. Только здесь пересечение $M_i \cap M_k$ строится не из равных слов и не их синонимов, а иначе. Например, так: пусть имеется пара текстов – оригинал (T_i) и перевод (T_k). Если слово $v \in M_i$, а $w \in M_k$ и в фиксированном двуязычном словаре (из языка текста T_i на язык текста T_k) в статье слова v найдется перевод слова w, то пара (v, w) входит в пересечение. Проиллюстрируем такой **Алгоритм 3** на примере пословиц с большим пословным совпадением.

All that glitters is not gold (10)

Не все то золото, что блестит (10')

As sow, so shall you reap (11)

Что посеешь, то пожнешь (11')

Better late than never (12)

Лучше поздно, чем никогда (12')

Blood is thicker than water (13)

Кровь людская – не водица (13')

Catch the bear before you sell its skin (14)

Не дели шкуру неубитого медведя (14')

The devil is not so black as he is painted (15)

Не так страшен черт, как его малюют (15')

Пример 2. Применим Алгоритм 3 к сравнению текстов (10) и (10'):

M (10) = {all, glitter, be, gold};

M (10') = {весь, тот, золото, блестеть}.

В известном англо-русском словаре Мюллера статья *all* содержит *весь*, статья *glitter* содержит *блестеть*, статья *gold* содержит *золото*. Итак, m (10, 10') = 4, n (10, 10') = 3, r (10, 10') = 0,75.

Конец примера.

Сравнивая разноязычные тексты из (10) – (15'), получаем следующую матрицу:

	10'	11'	12'	13'	14'	15'
10	0,75	0	0	0	0	0
11	0	0,4	0	0	0,15	0,27
12	0	0	1	0,22	0	0
13	0	0	0	0,5	0	0,15
14	0	0,15	0	0	0,36	0,14
15	0	0,27	0	0,15	0,14	0,43

Как видим, на диагонали стоят числа, достаточно удаленные от 1, т.е. пословная близость, задаваемая формулой (9), невелика при переводе пословиц. Это замечание относится к переводу текстов любого жанра. Как показала Н.В.Шаронова [7], сравнив переводы одного фрагмента прозы на семь языков, пословная близость колеблется возле r = 0,5 и не уменьшается при переводе на близкородственные языки. Сказанное означает, что системы МП, ориентированные на перевод, близкий к пословному, не могут обеспечить удовлетворительного качества.

Вернемся к оценке компараторным методом. Обсуждение примеров (6) – (8') и (10) – (15') показало, что они слишком просты для тестирования сложных систем. Алгоритмы 1 – 3 описаны, главным образом, для того, чтобы отсевать тривиальные ТТ.

4. Более трудные ТТ

Приведенные выше рассуждения отнюдь не означают, что трудно найти трудные ТТ. Например, такой жанр, как поэтические переводы одного оригинала разными переводчиками, является богатым источником трудных ТТ. Приведем для примера два перевода (О.Румером и Г.Плисецким) одного и того же рубаи Омара Хайяма:

Где высился чертог в далекие года,
И проводила дни султанов череда,
Там нынче горлица кричит среди развалин
И плачет бедная: «Куда? Куда? Куда?» (16)

Здесь владыки блистали в парче и шелку,
К ним гонцы подлетали на полном скаку.
Где все это? В зубчатых развалинах башни
Сиротливо кукушка кукует: «Ку-ку!» (16')

Несмотря на явное сходство смыслов, Алгоритм 2 (полагая с натяжкой, что *чертог* и *башня* являются синонимами, и обнаружив общее слово *развалина*) дает пословную близость r (16, 16') = 0,1. Еще одним богатым источником трудных ТТ являются пословицы. Некоторые, расходясь по словесному выражению, имеют общность структуры, например:

Out of sight, out of mind (17)

С глаз долой – из сердца вон (17')

В профессиональном фольклоре бытует байка, что (17) перевели на русский, а потом обратно на английский и получили *«Invisible idiot»*.

Самый трудный класс ТТ – это пословицы схожего смысла, где разнятся и слова, и структура; например:

Jack of all trades (18)

И швец, и жнец, и в дуду игрец (18')

Наличие тестов для оценки эффективности программ – событие, часто встречающееся в более строгих науках. Так, в теории дискретной оптимизации новые подходы часто пробуют на задаче коммивояжера, которая принята как своеобразный пробный камень. Чтобы результаты можно было сравнивать, опубликовано несколько конкретных задач [8], на которых испытывается эффективность новых процедур.

Приведем десять ТТ типа (16) – (18).

5. Библиотека ТТ повышенной трудности

А. Переводы рубаи Омара Хайяма разными авторами.

О, если б каждый день иметь краюху хлеба,
Над головою кров, и скромный угол, где бы
Ничьим владыкою, ничьим рабом не быть!
Тогда благословить за счастье можно б небо. (19)

Если есть у тебя для жилья закуток –
В наше подлое время – и хлеба кусок,
Если ты никому ни слуга, ни хозяин –
Счастлив ты и воистину духом высок. (19')

Мужи, чьей мудростью был этот мир пленен,
В ком светочей познанья видел он,
Дороги не нашли из этой ночи темной,
Посуесловили и погрузились в сон. (20)

Даже самые светлые в мире умы
Не смогли разогнать окружающей тьмы,
Рассказали нам несколько сказочек на ночь
И отправились, мудрые, спать, как и мы. (20')

Когда бываю трезв, не мил мне белый свет,
Когда бываю пьян, впадает разум в бред,
Лишь состояние меж трезвостью и хмелем
Ценю я, – вне его для нас блаженства нет. (21)

Трезвый я замыкаюсь, как в панцире краб.
Напиваясь, я делаюсь разумом слаб.
Есть мгновенье меж трезвостью и опьяненьем,
Это высшая правда, и я – ее раб! (21')

Я в мечеть не за праведным словом пришел,
Не стремясь приобщиться к основам пришел.
В прошлый раз утащил я молитвенный коврик,
Он истерся до дыр – я за новым пришел. (22)

Вхожу в мечеть. Час поздний и глухой.
Не в жажде чуда я и не с мольбой.
Когда-то коврик я стянул отсюда,
А он истерся; надо бы другой! (22')

Вино запрещено, но есть четыре «но»:
Смотря, кто, с кем, когда и в меру ль пьет вино.
При соблюдении сих четырех условий
Всем здравомыслящим вино разрешено. (23)

Запрет вина – закон, считающийся с тем,
Кем пьется, и когда, и много ли, и с кем.
Когда соблюдены все эти оговорки,
Пить – признак мудрости, а не порок совсем. (23')

Общаясь с дураком, не оберешься срама,
Поэтому совет ты выслушай Хайяма.
Яд, мудрецом предложенный, прими.
Из рук же дурака не принимай бальзама. (24)

Знайся только с достойными дружбы людьми.
С подлецами не знайся, себя не срами.
Если подлый лекарство нальет тебе – вылей!
Если мудрый подаст тебе яду – прими! (24')

Несовместимых мы полны желаний,
В одной руке бокал, другая на Коране…
Вот так мы и живем под небом голубым,
Полубезбожники и полумусульмане. (25)

Держит чашу рука, а другая – Коран,
То молюсь до упаду, то до смерти пьян.
Как лишь терпит нас мраморный свод бирюзовый –
Не кафиров совсем, не совсем мусульман. (25')

Нет благороднее растенья и милее,
Чем черный кипарис и белая лилея.
Он, сто имея рук, не тычет их вперед,
Она всегда молчит, сто языков имея. (26)

Да, лилия и кипарис – два чуда под луной,
О благородстве их твердит язык любой.
Имея десять языков, она всегда молчит,
А он, имея двести рук, не тычет ни одной. (26')

Я – словно старый дуб, что бурею разбит;
Увял и пожелтел гранат моих ланит,
Все естество мое – колонны, стены, кровля, –
Развалиною став, о смерти говорит. (27)

Старость – дерево, корень которого сгнил.
Возраст алые щеки мои посинил.
Крыша, дверь и четыре стены моей жизни
Обветшали и рухнуть грозят до стропил. (27')

Мы больше в этот мир вовек не попадем,
Вовек не встретимся с друзьями за столом,
Лови же каждое летящее мгновенье –
Его не подстеречь уж никогда потом. (28)

Боюсь, что в этот мир мы вновь не попадем,
И там своих друзей – за гробом – не найдем.
Давайте пировать в сей миг, пока мы живы.
Быть может, миг пройдет – мы все навек уйдем. (28')

B. Пословицы, где слова разные, но структуры схожи

Better an egg today than a hen tomorrow (29)
Лучше синица в руках, чем журавль в небе (29')

Stretch your legs according to your coverlet (30)
По одежке протягивай ножки (30')

Do not halloo till you are out of the wood (31)
Не кричи гоп, пока не перескочишь (31')

Do not make a mountain out of a molehill (32)
Не делай из мухи слона (32')

Dog does not eat dog (33)
Ворон ворону глаз не выклюет (33')

East or West, home is best (34)
В гостях хорошо, а дома лучше (34')

Make hay while the sun shines (35)
Куй железо, пока горячо (35')

If "ifs" and "cans" where pots and pans ...(36)
Если бы да кабы во рту выросли грибы (36')

Little strokes fell great oaks (37)
Капля камень точит (37')

Bought a pig in a poke (38)
Купил кота в мешке (38')

C. Пословицы, где и слова, и структуры – разные

All is fish that comes to the net (39)
На безрыбье и рак рыба (39')

The biter is sometimes bit (40)
Вор у вора дубинку украл (40')

Clothes do not make the men (41)
По одежке встречают – по уму провожают (41')

Diamond cuts diamond (42)
Нашла коса на камень (42')

Do as you would be done by (43)
Не рой другому яму (43')

The early bird catches the worm (44)
Кто рано встает, тому Бог дает (44')

Every cloud has its silver lining (45)
Нет худа без добра (45')

The fish will soon be caught that nibbles at every bait (46)
Любопытной Варваре нос оторвали (46')

Great oaks from little acorn grow (47)
Всяк бык теленком был (47')

Haste makes waste (48)
Поспешишь – людей насмешишь (48')

Литература. 1. *Виноград Т.* Программа, понимающая естественный язык. М.: Мир. 1976. 294 с. 2. *Вейценбаум Дж.* Возможности вычислительных машин и человеческий разум: От суждений к вычислениям. М.: Радио и связь. 1982. 370 с. 3. *Командин А.Ф.* Обобщенные пространства и их применение для автоматической обработки текстов естественного языка. Дисс. канд. техн. наук, Х., 1995. 134с. 4. *Рогинский В.Н.* Человек разговаривает с ЭВМ. М.: Знание. 1976. 64 с. 5. *Поспелов Д.А.* Фантазия или наука: На пути к искусственному интеллекту. М.: Наука. 1982. 220 с 6. *.Шабанов-Кушнаренко Ю.П., Шаронова Н.В.* Компараторная идентификация лингвистических объектов. К.: ИСИО, 1993. 116с. 7. *Шаронова Н.В.* Компараторная идентификация лингвистических объектов. Дисс. докт. техн. наук , Х., 1994. 354 с. 8. *Хелд М., Карп Р.М.* Применение динамического программирования к задачам упорядочения / В кн.: Кибернетический сборник. Вып. 9. М.: Мир. 1964. С. 202 - 218.

Поступила в редколлегию 09.02.98

Рецензент: д-р техн. наук Смеляков С.В.

Калиновский Андрей Станиславович, аспирант кафедры ПОЭВМ ХТУРЭ. Научные интересы: интеллектуальные компьютерные технологии, системы обработки естественно-языковой информации, базы данных и знаний. Адрес: Украина, 310726, Харьков, пр. Ленина, 14, тел. 40-94-46.

Рублинецкий Владимир Ильич, старший научный сотрудник кафедры ПОЭВМ ХТУРЭ. Научные интересы: компьютерная лингвистика, машинное понимание естественного языка, машинный перевод. Адрес: Украина, 310726, Харьков, пр. Ленина, 14, тел. 40-94-46.

Рябова Наталия Владимировна, к.т.н., доцент филиал-кафедры ИИИС ХТУРЭ. Научные интересы: интеллектуальные компьютерные системы, обработка естественного языка, извлечение знаний из текстовых баз данных. Адрес: Украина, 310726, Харьков, пр. Ленина, 14, тел. 40-98-90

УДК 519.767.2

Оценка понимания смысла ЕЯ компараторным методом / А.С. Калиновский, В.И. Рублинецкий, Н.В. Рябова // Радиоэлектроника и информатика. – 1998. - № 4. – С. 95 – 99

Многочисленные системы ИИ, которые «понимают» ЕЯ тексты, не сравниваются по эффективности. Предлагается методика измерения эффективности по результатам сравнения тест-текстов (ТТ), приблизительно похожих по форме и либо разных, либо равных по смыслу. Описаны массовые источники легких и трудных ТТ, а также - алгоритмы отсева легких. Приложена библиотека ТТ для систем, обрабатывающих русский и/или английский язык.

Библиогр.: 8 назв.

UDK 519.767.2

The estimation of NL semantic understanding by comparatory identification / A.S. Kalinovsky, V.I. Rublinetsky, N.V. Ryabova // Radioelektronika i informatika.- 1998.- N 4.- P. 95 – 99.

Many AI systems that ‘understand’ NL texts are never compared as to their efficiency. The article suggests a technique for such comparision by the results of understanding test texts (TTs) that are approximately similar in form and equal or different in their meaning. Massive sources of TTs are indicated, both hard and easy. Algorithms for sifting away easy TTs are described. A set of TTs for messuring the efficiency of systems dealing with English or/and Russian is appended.

Refs: 8 items.