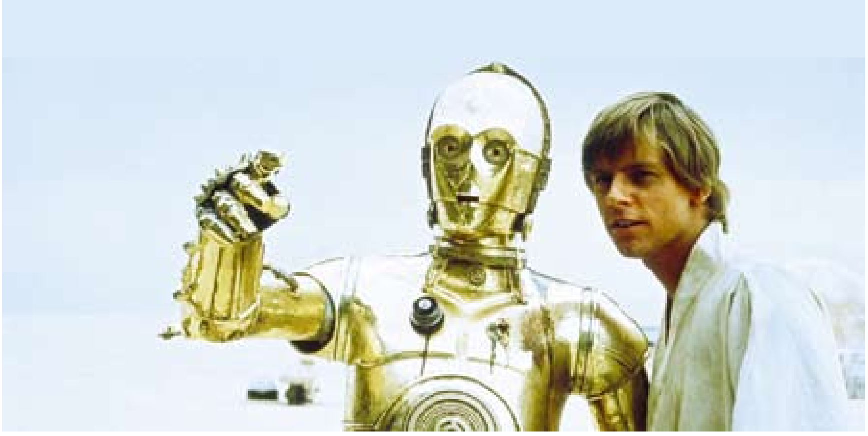

О. МИТРЕНИНА,
Санкт-Петербургский
государственный университет

МАТЕМАТИКА И МОДЕЛЬ ЯЗЫКА

Робот C-3PO из «Звездных войн» может общаться шестью миллионами способов. Но его выдает легкий металлический акцент — мы понимаем, что говорит робот.

Современные электронные устройства — компьютеры, смартфоны, дикторы в метро — умеют произносить предложения без акцента. Но они не смогут поддержать свободную беседу.

Как научить компьютер человеческому языку? Может показаться, что надо поместить в него словарь и правила, а также все те премудрости, которые изучают на уроках русского языка. Но, получив эти знания, компьютер будет говорить плохо и почти ничего не поймет. Он будет вести себя как человек, который попытался выучить незнакомый язык с помощью словаря и грамматики.

Чтобы хорошо выучить чужой язык, надо оказаться в иной языковой среде. Например, можно поехать в другую страну, где никто не владеет вашим родным языком или хотя бы смотреть фильмы на нужном языке и беседовать о чем-то интересном в социальных сетях. Тогда незнакомый язык сможет «пропитать» ваш мозг, стать привычным и понятным.

Почти то же самое можно сделать и с машиной. Правила и словари пригодятся, но главное — компьютер должен получить много настоящих текстов на том языке, который он «изучает». А дальше он «обучится» с помощью математики.

Так происходит и при машинном переводе. Вначале компьютер получает много готовых предложений вместе с их переводами на другой язык. Далее он изучает, какие куски текста повторялись в переводе и в оригинале, — это надо делать на очень большом количестве примеров. После этого компьютер сможет переводить новые предложения, а точнее — правильно собирать перевод из фрагментов, которые он проанализировал ранее. При этом он проследит, чтобы перевод звучал гладко.

Что такое модель?

Модель отражает свойства какого-нибудь предмета или явления.

Всем известны, например, модели самолетов. Это может быть пластмассовая фигурка, которая стоит на книжной полке, — такая модель отражает внешний вид самолета. Более сложную модель собирают в кружке авиамоделирования — она умеет летать по кругу на веревочке и обладает многими свойствами летательного аппарата.

Но можно построить и другую модель: деревянный самолет, который стоит на земле, издает звуки двигателя и трясется. Внутри него расположены пассажирские кресла, мимо которых иногда проходит стюардесса в форме и раздает карамельки. Такая модель никогда не взлетит, но и она отражает некоторые важные свойства самолета. С помощью нее можно лечить тех людей, которые боятся авиаперелетов.

Самолет — это искусственный предмет, строить его модели достаточно просто. Гораздо сложнее дело обстоит с естественными объектами. Так, долгое время люди мечтали создать махолет — модель птицы, которая летает с помощью взмахов крыльев. Можно измерить мышцы птицы и густоту ее перьев, но нельзя «заглянуть» в ее голову, чтобы понять, как именно она управляет своим организмом, чтобы летать. Однако ученые смогли построить математическую модель, которая заменила вестибулярный аппарат птицы, и первые махолеты поднялись в воздух.

Способность говорить похожа на способность летать. До сих пор неизвестно, как возникают в нашей голове предложения и как мы понимаем то, что нам говорят. Исследователи могут видеть или слышать только сами предложения — это то же самое, что наблюдать за летящей птицей со стороны. Впрочем, это не так уж и мало: замедленная съемка, например, позволяет многое узнать о движении крыльев. А описание предложений с помощью математических методов дает возможность создавать модели, на основе которых компьютер учится работать с текстами и речью.

Модель языка

Модели языка, как и модели самолетов, бывают разные.

В прикладной лингвистике моделью языка (Language Model — LM) часто называют только один вид модели: систему, которая позволяет определять, какие предложения и сочетания более вероятны.

Допустим, мы собрали некоторое множество словоформ — слов в различных формах:

$$v = \{\text{этот, большой, дракон, мальчик, увидел, мальчика, ...}\}.$$

Из этих словоформ можно построить бесчисленное множество предложений. Далеко не все они окажутся правильными, но для машины все они допустимы, ведь она не знает правил русского языка:

этот большой дракон увидел мальчика
этот мальчик большой
дракон этот этот
дракон дракон дракон

Какое из этих предложений более вероятно? Наверное, второе и первое. А третье и четвертое маловероятны, но нельзя их считать совершенно невозможными — ведь вы их уже встретили в этой статье. Да и вообще, если человек выйдет из своего подъезда и увидит дракона, он вполне сможет воскликнуть: «Дракон этот этот!»

Модель языка позволяет компьютеру определять, какие предложения и сочетания более вероятны. Это необходимо в самых разных задачах автоматической обработки текстов. Например, при машинном переводе.

Английское слово *table* означает не только стол, но и таблицу. А если к нему добавить слово *wooden* (деревянный), то какое сочетание мы выберем?

1. Деревянный стол.
2. Деревянная таблица.

Люди понимают, что речь идет о деревянном столе, а не о деревянной таблице. Но как это объяснить компьютеру? Можно заранее задать ему списки материалов: из чего делают столы и из чего не делают таблицы. Для этого потребуется загружать в машину целые базы знаний, которые до сих пор не удается сделать всеобъемлющими. Есть более простой путь: компьютер может просто «вспомнить», что сочетание «деревянный стол» ему уже попадалось, а «деревянная таблица» нет.

Похожая ситуация возникает при автоматическом распознавании речи. Представим, что компьютер услышал песню «Расцветали яблони и груши...». Он пытается построить соответствующую ей цепочку слов, и у него получаются два варианта.

1. Раз цвета ли яблони?
2. Расцветали яблони.

Мы, люди, сразу понимаем, что первый вариант неправильный. При этом мы ориентируемся на свой опыт языка. А компьютер сможет это понять с помощью модели языка: в хорошей модели второй вариант будет более вероятным.

Какое предложение вероятнее?

Мы выбираем нужные предложения, основываясь на своем языковом опыте. Точно так же и компьютеру можно передать «языковой опыт», если показать ему много правильных предложений. С их помощью он построит модель языка и сможет понимать, какие предложения более вероятны.

Наша задача звучит так.

У нас есть большой набор текстов — он называется корпус. Как с его помощью оценить, какие предложения более вероятны?

Как мы оцениваем вероятность того, что монета упадет вверх орлом? Мы рассматриваем, сколько возможно вариантов. Их два: орел и решка. Они одинаково вероятны. Поэтому вероятность выпадения орла — $\frac{1}{2}$. А если мы кидаем шестигранную игральную кость, стороны у которой пронумерованы от 1 до 6, то вероятность выпадения единицы равна $\frac{1}{6}$.

Похожим образом мы можем оценивать и вероятность предложений — с помощью корпуса.

Допустим, Вася решил создать корпус высказываний своей бабушки. Два дня он записывал все, что она говорит. Получилось 500 предложений, многие из которых повторялись, но он записал и повторы. Как с помощью этого корпуса сравнить вероятности предложений?

Способ I. Можно оценить вероятность каждого предложения простым способом: посчитать, сколько раз оно встретилось в корпусе, и разделить это на количество всех предложений корпуса.

Допустим, пока Вася создавал корпус, бабушка 10 раз сказала фразу «Ешь скорее». При этом объем корпуса получился 500 предложений. Тогда вероятность фразы «Ешь скорее» равна

$$\frac{10}{500}=0{,}02.$$

Это очень простой способ вычисления вероятности, но у него есть важный недостаток. У всех предложений, которые не встретились в корпусе, вероятность будет равна нулю. Например, у предложения «Ешь медленнее». Ранее бабушка говорила только «Иди медленнее» или «Ешь скорее». Теоретически она вполне может сказать «Ешь медленнее», но если это предложение не встречалось в корпусе, его вероятность будет нулевой.

Желательно придумать какой-то другой подход, который позволял бы оценивать вероятности таких предложений, которые в корпусе не встречались.

Способ II. Можно учитывать вероятность появления отдельных слов. Например, вероятности слов «Ешь» и «медленнее». А далее посчитать их совместную вероятность.

Как считают совместную вероятность двух событий? Например, если мы кидаем одновременно монетку и игральную кость, какова вероятность того, выпадут одновременно орел и единица? Поскольку эти события независимы, их вероятности просто перемножаются:

$$\frac{1}{2}\cdot\frac{1}{6}=\frac{1}{12}.$$

Можно попытаться применить этот подход и к текстам. Всего в корпусе бабушки получилось 1500 слов (500 предложений, в среднем по 3 слова в каждом). При этом слово «Ешь» встретилось 30 раз. Тогда его вероятность можно оценить как

$$\frac{30}{1500}=0{,}02.$$

Запишем это с помощью формулы:

$$p(\text{Ешь})=\frac{c(\text{Ешь})}{c(\;)}=\frac{30}{1500}=0{,}02.$$

Здесь вероятность обозначается буквой p, а количество появлений слова — буквой c. Расположенная в знаменателе запись $c\;(\;)$ означает общее число слов в корпусе.

Точно так же вычисляется вероятность слова «медленнее», которое, допустим, встретилось в корпусе 6 раз.

$$p(\text{медленнее})=\frac{c(\text{медленнее})}{c(\;)}=\frac{6}{1500}=0{,}004.$$

Тогда вероятность предложения «Ешь медленнее» можно посчитать как произведение вероятностей входящих в него слов «Ешь» и «медленнее»:

$$p\,(\text{Ешь медленнее}) =$$
$$= p\,(\text{Ешь})\cdot p\,(\text{медленнее}) =$$
$$= 0{,}02\cdot 0{,}004 = 0{,}00008.$$

Этот способ лучше, чем предыдущий, но и его можно усовершенствовать.

Способ III (основной). Вероятность слова зависит от тех слов, которые стоят перед ним. После слова «манная», скорее всего, будет идти либо «каша», либо «крупа», либо «запеканка». А сочетание «съешь манную» почти наверняка потянет за собой слово «кашу». Никто не предлагает съесть манную крупу, а запеканка встречается безнадежно редко.

Вероятность слова «кашу» можно оценить точнее, если учесть, что перед ним появились слова «съешь манную». Такую вероятность называют условной и записывают ее так:

$$p\,(\text{кашу}\,|\,\text{съешь манную}).$$

Современные модели языка используют условную вероятность слов. Во второй половине статьи мы рассмотрим подробнее, как это происходит.

Условная вероятность

Если мы создаем модель языка, все условные вероятности надо посчитать заранее. Это делается с помощью корпуса, на основе которого строится модель.

Рассмотрим пример.

Допустим, вы едете в поезде. В одном купе с вами оказался шумный маленький мальчик. Не замолкая ни на секунду, он без конца что-то просит у своих родителей. Другой путешественник на вашем месте мог бы огорчиться из-за такого соседства, но настоящий ученый умеет извлекать пользу из неудобств. Вы решаетесь провести лингвистический эксперимент: создать корпус высказываний мальчика и построить языковую модель.

Вот фрагмент этого корпуса — все предложения, начинающиеся со слова «хочу»:

– «хочу собаку» — 21 раз;
– «хочу гулять» — 14 раз;
– «хочу сок» — 4 раза;
– «хочу домой» — 3 раза.

Всего: 42 предложения, начинающихся со слова «хочу».

Когда в очередной раз мальчик начинает медленно произносить слово «хочу», вы можете легко оценить вероятность появления слова «собаку». После слова «хочу» слово «собаку» следовало в 21 предложении из 42. Значит, вероятность оценивается как $\frac{21}{42}=\frac{1}{2}$.

Это и есть условная вероятность — вероятность слова «собаку», при условии, что начало предложения состоит из слова «хочу».

Запишем все условные вероятности, посчитанные с помощью фрагмента нашего корпуса:

$$p(\text{собаку} \mid \text{хочу})=\frac{21}{42}=\frac{1}{2}=0,5;$$

$$p(\text{гулять} \mid \text{хочу})=\frac{14}{42}=\frac{1}{3};$$

$$p(\text{сок} \mid \text{хочу})=\frac{4}{42}=\frac{2}{21};$$

$$p(\text{домой} \mid \text{хочу})=\frac{3}{42}=\frac{1}{14}.$$

В этом примере в качестве условия мы учитывали только одно предшествующее слово, потому что больше слов не было — предложения слишком короткие. Но как быть, если предложение длинное?

«Прошу тебя, мой юный друг,
с аппетитом съешь манную кашу».

В этой изысканной фразе перед словом «кашу» расположено еще 9 слов и два знака препинания. Нужно ли учитывать их при оценке условной вероятности? Интуиция подсказывает, что можно обойтись и без них. И это правильно.

В начале XX в. математик Андрей Марков предположил, что не слишком важно, учитываем мы длинную цепочку предыдущих событий или берем только два или три ближайших события, — в большинстве случаев условная вероятность будет отличаться не очень сильно. Так возникли марковские цепи — модели, в которых вероятность события не зависит от далекого прошлого. В нашем случае это дает вот такое приблизительное равенство:

$$p\,(\text{кашу} \mid \text{Прошу тебя, мой юный друг, с аппетитом съешь манную}) \approx$$
$$\approx p\,(\text{кашу} \mid \text{съешь манную}).$$

В этом примере мы учитываем только два предыдущих слова, и это самый распространенный подход в тех случаях, когда язык исследуется статистическими методами.

Триграммная модель языка

Сочетание из трех идущих подряд слов называют триграммой. Например, «съешь манную кашу» — это триграмма, именно она используется в нашем примере:

$$p\,(\text{кашу} \mid \text{съешь манную}).$$

Основанные на них модели называются триграммными языковыми моделями.

Как построить триграммную модель? Она создается на основе корпуса текстов.

Вначале нужно собрать множество всех словоформ корпуса — слов в различных формах, но мы их тоже будем называть словами. Напомним, как выглядит такое множество:

$$\nu = \{\text{хочу, собаку, собака, где, гулять, ...}\}.$$

Слова «собаку» и «собака» — это два разных элемента множества. Ведь для компьютера это просто две разные цепочки символов.

Множество слов корпуса может оказаться очень большим, но оно всегда будет конечным.

Далее для любой тройки слов АБВ из множества ν нужно оценить параметр q (В | АБ). Он соответствует вероятности того, что после слов АБ идет слово В. Например, параметр q (кашу | съешь манную).

Параметров должно получиться очень много — по одному для каждой возможной тройки слов. Кроме того, нужны отдельные параметры для тех случаев, когда слово стоит в начале или в конце предложения, об этом мы скажем чуть позже.

Множество ν вместе с набором всех параметров и образуют триграммную модель языка.

Но как с помощью триграммной модели оценить вероятность предложения? Очень просто.

Рассмотрим следующий пример:

сегодня выпал снег STOP.

Искусственное слово STOP появилось для обозначения конца предложения, оно еще пригодится. Кроме того, предложение написано с маленькой буквы, потому что перед обработкой текста обычно производится «понижение регистра» — все прописные буквы заменяются на строчные.

Вероятность этого предложения можно посчитать как произведение условных вероятностей каждого его слова, точнее, как произведение параметров для каждой триграммы:

$$p\,(\text{сегодня выпал снег}) = q\,(\text{сегодня} \mid *\,*) \times q\,(\text{выпал} \mid *\ \text{сегодня}) \times q\,(\text{снег} \mid \text{сегодня выпал}) \times q\,(\text{STOP} \mid \text{выпал снег})$$

Множитель $q\,(\text{сегодня} \mid *\,*)$ соответствует вероятности того, что слово «сегодня» используется в абсолютном начале предложения. Это начало обозначено звездочками, они показывают, что перед словом «сегодня» в предложении нет других слов.

Множитель $q\,(\text{выпал} \mid *\ \text{сегодня})$ соответствует вероятности того, что слово «выпал» находится после слова «сегодня», а перед ними в предложении ничего больше нет (отсутствие других слов перед словом «сегодня» обозначается звездочкой).

Множитель $q\,(\text{снег} \mid \text{сегодня выпал})$ соответствует вероятности того, что после слов «сегодня выпал» идет слово «снег».

Множитель $q\,(\text{STOP} \mid \text{выпал снег})$ обозначает вероятность конца предложения после слов «выпал снег».

Задача 1. Даны следующие параметры:

$q\,(\text{купи} \mid *\,*) = 1$;
$q\,(\text{хлеба} \mid *\ \text{купи}) = 0{,}5$;
$q\,(\text{пожалуйста} \mid *\ \text{купи}) = 0{,}5$;
$q\,(\text{STOP} \mid \text{купи хлеба}) = 1$;
$q\,(\text{STOP} \mid \text{купи пожалуйста}) = 0{,}25$;
$q\,(\text{вертолет} \mid \text{купи пожалуйста}) = 0{,}25$;
$q\,(\text{STOP} \mid \text{пожалуйста поесть}) = 0{,}25$;
$q\,(\text{STOP} \mid \text{пожалуйста вертолет}) = 0{,}5$.

У каких предложений на основании этих параметров есть ненулевая вероятность? Посчитать эту вероятность

Ответ:

купи хлеба STOP: $1 \cdot 0{,}5 \cdot 1 = 0{,}5$;
купи пожалуйста STOP: $1 \cdot 0{,}5 \cdot 0{,}25 = 0{,}125$;
купи пожалуйста вертолет STOP: $1 \cdot 0{,}5 \cdot 0{,}25 \cdot 0{,}5 = 0{,}0625$.

Пояснения. Вероятность первого предложения вычисляется так:

$$p\,(\text{купи хлеба STOP}) = q\,(\text{купи} \mid *\,*) \times q\,(\text{хлеба} \mid *\ \text{купи}) \times q\,(\text{STOP} \mid \text{купи хлеба}) = 1 \cdot 0{,}5 \cdot 1 = 0{,}5.$$

Предложение «купи пожалуйста поесть» невозможно, так как у нас нет параметра $q\,(\text{поесть} \mid \text{купи пожалуйста})$.

Вычисление параметров модели

Как посчитать параметры модели языка? Например, чему равен параметр $q\,(\text{тарелку} \mid \text{скорее неси})$? Он соответствует вероятности того, что после слов «скорее неси» встретится слово «тарелку».

Параметры можно вычислять точно так же, как вы считали вероятности на примере высказываний разговорчивого мальчика в купе. Допустим, сочетание «скорее неси» встретилось в корпусе 15 раз. А «скорее неси тарелку» — 5 раз. Значит, только в пяти случаях из пятнадцати после слов «скорее неси» следовало слово «тарелку».

Тогда параметр $q\,(\text{тарелку} \mid \text{скорее неси})$ можно вычислить так:

$$q\,(\text{тарелку} \mid \text{скорее неси}) = \frac{c\,(\text{скорее неси тарелку})}{c\,(\text{скорее неси})} = \frac{5}{15} = \frac{1}{3}$$

Такой подход называют оценкой максимального правдоподобия. С помощью него вычисляются параметры для всех возможных троек слов, а также для слов в начале и в конце предложения.

Задача 2. Имеется следующий корпус:

я хочу гулять здесь STOP;
хочу собаку STOP;
я хочу эту собаку STOP;
я хочу вертолет STOP;
хочу вертолет STOP.

С помощью оценки максимального правдоподобия найти значение следующих параметров:

$q\,(\text{гулять} \mid \text{я хочу})$;
$q\,(\text{вертолет} \mid *\ \text{хочу})$.

Напоминание. Звездочка обозначает начало предложения, отсутствие предшествующих слов.

Решение.

$$q\,(\text{гулять} \mid \text{я хочу}) = \frac{c\,(\text{я хочу гулять})}{c\,(\text{я хочу})} = \frac{1}{3};$$

$$q\,(\text{вертолет} \mid *\ \text{хочу}) = \frac{c\,(*\ \text{хочу вертолет})}{c\,(*\ \text{хочу})} = \frac{1}{2}.$$

Ответ: $\frac{1}{3}$ и $\frac{1}{2}$.

Чтобы параметры получились более точными, корпус должен быть большим. И он должен соответствовать тому языку, модель которого вы хотите построить.

Например, если ехать с разговорчивым мальчиком в купе трое суток и записывать все его фразы, то можно собрать неплохой корпус. Но он будет отражать язык только этого отдельно взятого мальчика и только в дороге, а не в цирке, например. Если построить модель языка с помощью такого корпуса, то можно научить робота разговаривать как мальчик в поезде, но не как диктор радио или продавщица в магазине.

Идеальную модель языка можно построить только в сказке. Для этого надо собрать все предложения русского языка, которые когда-то были произнесены или написаны, и добавить к ним все предложения, которые когда-либо произнесут или напишут в будущем. Получится очень большой корпус, суперкорпус, и с помощью него нужно будет посчитать параметры модели.

Но собрать все предложения невозможно, поэтому в жизни приходится довольствоваться корпусами поменьше, хотя и они бывают очень большими. Есть правила их составления, но это уже область прикладной лингвистики, а не математики.

Сглаживание

Как быть, если в тексте появится сочетание, ранее не встречавшееся в корпусе? Например, после того, как мы собрали корпус, мальчик в купе мог неожиданно произнести: «дай мне вертолет». Но такой триграммы в нашем корпусе не было, поэтому параметр q (вертолет | дай мне) равен нулю. А значит, и вероятность всего предложения окажется нулевой, ведь параметры перемножаются.

Получается, что модель предсказывает, будто предложение «дай мне вертолет» никогда не может появиться, а это не так.

Из похожих предложений мальчик говорил «дай мне машинку» и «хочу вертолет». На основе этих и других его фраз можно оценить вероятность сочетания «дай мне вертолет», а также всех других сочетаний, которые не встретились в корпусе, но вполне могут встретиться в жизни.

Чтобы вероятность не появившихся в корпусе сочетаний была чуть больше нуля, модель надо «пригладить» — большие значения немного уменьшить, а нулевые слегка увеличить. В статистике для этого используются разнообразные методы сглаживания. Например, метод дисконтирования: от числа появлений каждой триграммы отнимают небольшое значение β, так накапливается «запасная» вероятность, которую можно поделить между нулевыми триграммами.

Рассмотрим это на примере. Как сделать ненулевым параметр q (вертолет | дай мне), если мальчик ни разу не сказал «дай мне вертолет».

Установим значение $\beta = 0{,}5$.

Рассмотрим все триграммы нашего корпуса, которые начинались со слов «дай мне»:

– «дай мне машинку» встретилось 15 раз;
– «дай мне это» встретилось 10 раз;
– «дай мне сок» встретилось 5 раз;
– «дай мне STOP» встретилось 5 раз.

В последнем примере мальчик просто сказал «дай мне» и замолчал.

Всего в корпусе встретилось $15 + 10 + 5 + 5 = 35$ триграмм, начинающихся со слов «дай мне».

Запишем эти данные в виде таблицы. Первые две колонки понятны: x — это триграмма, а $c(x)$ — число ее появлений в нашем корпусе. Третья колонка — это вторая колонка минус значение β (в нашем примере $\beta = 0{,}5$). Благодаря данным в третьей колонке можно считать, что триграмма «дай мне машинку» встретилась не 15 раз, а 14,5 раза. Пусть значение не будет целым, ничего страшного.

Последняя колонка — это оценка максимального правдоподобия на основе того, что триграмма встретилась $c(x) - 0{,}5$ раза.

x	$c(x)$	$c(x) - \beta$	$\frac{c(x)-\beta}{35}$
дай мне машинку	15	14,5	$\frac{14{,}5}{35}$
дай мне это	10	9,5	$\frac{9{,}5}{35}$
дай мне сок	5	4,5	$\frac{4{,}5}{35}$
дай мне STOP	5	4,5	$\frac{4{,}5}{35}$

Если триграмма «дай мне машинку» встретилась 15 раз, то оценка ее максимального правдоподобия с учетом дисконтирования вычисляется так:

$$q\,(\text{машинку} \mid \text{дай мне}) = \frac{c\,(\text{дай мне машинку}) - \beta}{c\,(\text{дай мне})} =$$

$$= \frac{15 - 0{,}5}{35} = \frac{14{,}5}{35} = \frac{29}{70}.$$

При таком подходе каждый параметр получится немного заниженным по сравнению с тем, что показывает корпус. Благодаря этому удается сэкономить некоторое число «появлений» триграмм. В нашем случае сэкономлено два появления. Ведь на самом деле триграмм, начинающихся с «дай мне», было 35, но третья колонка показывает, что их было в сумме 33. Два появления триграмм сэкономлено. Осталось только честно поделить эту двойку между всеми триграммами, начинающимися со слов «дай мне», вероятность которых получилась нулевой. Делить можно разными способами. Например, триграмме «дай мне вертолет» можно отдать значение 0,1 — получится, что она появилась в корпусе 0,1 раз.

С точки зрения здравого смысла нецелое появление звучит абсурдно, но для статистики оно сгодится. Ведь теперь мы сможем посчитать параметр q = (вертолет | дай мне):

$$q\,(\text{вертолет}\mid\text{дай мне})=\frac{c\,(\text{дай мне вертолет})}{c\,(\text{дай мне})}=\frac{0{,}1}{35}=\frac{1}{350}.$$

Получилось не очень большое значение, но уже не ноль.

Двигаться дальше

30 лет назад компьютер «Урал» в 239-й математической школе занимал целую комнату. Сегодня обычный смартфон во много раз превышает его по своей мощности. Техника продолжает стремительно совершенствоваться, а вместе с ней развиваются и формальные методы анализа действительности.

Математические модели уже применяются в самых разных областях: в радиотехнике, в экономике, в медицине, в генетике и даже в искусстве — с помощью них успешно предсказывают популярность новых книг и киносценариев. Особенно важны они в системах безопасности, ведь современные угрозы невозможно отслеживать вручную.

Наступившее тысячелетие, несомненно, будет все глубже использовать методы автоматической обработки огромных массивов данных. Изучать основы этих методов необходимо уже в школе, чтобы дать нашим учащимся путевку к будущим научным и практическим успехам.

Модели языка — это тот материал, с помощью которого знакомство с современными интеллектуальными технологиями можно сделать простым и интересным.

ТЕОРЕМА О ДЕЛИМОСТИ СУММЫ ЧИСЕЛ

В. БАРАНОВСКИЙ,
Гатчинский р-н, Ленинградская обл.

Мною было замечено, что если номер автомашины представляет собой числа, полученные перестановкой цифр, например, 39 и 93, то их сумма делится на 11. Это свойство я проверил на трехзначных числах: делались все возможные перестановки цифр (их шесть) и складывались полученные числа. Сумма всегда делилась на 111.

Пример. Выполним все возможные перестановки цифр числа 345 и найдем сумму полученных чисел:

$$345 + 435 + 534 + 354 + 453 + 543 = 2664.$$

Полученная сумма делится на 111.

Оказалось, что указанное свойство можно сформулировать в виде теоремы для любого натурального числа.

Теорема. Сумма чисел, полученных от всех возможных перестановок цифр любого натурального числа, делится на число, составленное из единиц того же разряда, что и данное число.

Доказательство. Используем метод математической индукции.

Пусть n — номер разряда числа. Проверим утверждение при различных n.

При $n = 1$ утверждение верно, так как любое однозначное число делится на 1.

При $n = 2$ сумма двузначных чисел, составленных из цифр A и B, может быть записана в виде:

$$A\cdot 10 + B + B\cdot 10 + A = 10(A + B) + (A + B) = 11(A + B) = P_1\cdot 11(A + B),$$

где $P_1 = 1! = 1$.

Очевидно, что выражение $P_1\cdot 11(A + B)$ делится на 11. Таким образом, утверждение верно.

При $n = 3$ сумма трехзначных чисел, составленных из цифр A, B и C, может быть записана в виде:

$$A\cdot 100 + B\cdot 10 + C + A\cdot 100 + C\cdot 10 + B + B\cdot 100 + A\cdot 10 + C + B\cdot 100 + C\cdot 10 + A + C\cdot 100 + A\cdot 10 + B + C\cdot 100 + B\cdot 10 + A = 2\cdot 111(A + B + C) = P_2\cdot 111(A + B + C),$$

где $P_2 = 2! = 1\cdot 2 = 2$.

Выражение $P_2\cdot 111(A + B + C)$ делится на 111. Таким образом, утверждение верно.

При $n = k$ сумма всех перестановок из k цифр будет иметь вид:

$$P_{k-1}\cdot 111...1\cdot(A + B + C + ... + K),$$

где $P_{k-1} = (k - 1)!$

Эта сумма делится на 111...1 — число, составленное из k единиц.

Докажем, что утверждение верно при $n = k + 1$.

$$P_k\cdot(111...11)\cdot(C_1 + C_2 + ... + C_{k+1}),$$

где $P_k = k!$, 111...11 — число, составленное из $k + 1$ единицы. Выражение суммы делится на число, составленное из $k + 1$ единиц.

Следовательно, теорема о делимости суммы чисел, полученных от всех возможных перестановок цифр натурального числа, на число, составленное из единиц того же разряда, что и данное число, доказана.