



Из мухи — СЛОНА

В повести известного писателя Юрия Трифонова (1925–1981) «Долгое прощание» рассказывается, как герой проводит время в поезде: «Третьи сутки Ребров, лёжа на верхней полке, мучил себя – делал из мухи слона. На листке бумаги писал: муха – мура – кура – кора – корт – торт – торс...».

Правила игры очень просты: надо построить цепочку слов от начального (МУХА) до конечного (СЛОН), на каждом шаге меняя только одну букву. При этом могут использоваться только русские 4-буквенные нарицательные существительные в начальной форме: например, слова БАЗА, НОЧЬ, САНИ допускаются, а слова ЛИТЬ, ХОТЯ, РУКУ, НОЧИ, САНЯ, ОСЛО, АБВГ, ФЦНМ – нет (первые два – не существительные, следующие два – не в начальной форме, следующие два – собственные, а не нарицательные, а два последних вовсе не существуют в языке).

Эта игра под названием «Дублеты» приобрела известность благодаря Льюису Кэрроллу – не только автору книг про Алису, но ещё и замечательному математику. В марте 1879 года он начал раз в неделю публиковать в журнале «Ярмарка тщеславия» по три задания в форме броских фраз: «Turn POOR into RICH» – «Преврати бедного в богатого», «Evolve MAN from APE» – «Выведи человека из обезьяны», «Make TEA HOT» – «Сделай чай горячим». В том же году он выпустил брошюру «Дублеты», подробно описал в ней правила и предложил читателям попрактиковаться на нескольких десятках примеров.

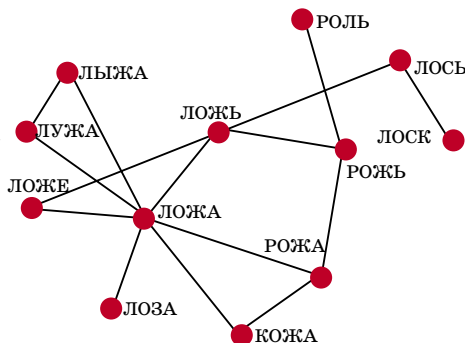
Вот и вам пять пар для тренировки – попробуйте построить для них цепочки. Сразу предупреждаю, что в одном случае, скорее всего, не получится: БОРЩ → ПОСТ; ЛИПА → ЖАРА; КИНО → ВАТА; КЛЕН → ЕЛКА; ПАУК → ЛОСЬ (ответы см. в конце номера).

Вы наверняка нашли четыре из пяти цепочек, а не смогли построить только одну. Но как доказать, что это действительно невозможно? Может быть, мы просто не придумали способ, а вообще-то он есть. Разобраться нам поможет теория графов.

Первым делом договоримся о том, что именно мы считаем словами, а то может получиться, что я дам

вам задание МУХА → СЛОН и вы скажете: «4 хода! МУХА → МУХН → МУОН → МЛОН → СЛОН». Тогда мне придётся со словарями в руках доказывать, что это жульничество, потому что слов МУХН, МУОН и МЛОН не существует. Чтобы избежать этого, мы обратимся к словарю с самого начала и постановим, что будем пользоваться только теми 4-буквенными существительными, которые есть в заранее выбранном источнике. Для этой статьи я взял «Грамматический словарь русского языка» Андрея Зализняка – этот словарь чаще всего применяют в компьютерной лингвистике. Кстати, мысль о том, что очень важно заранее договориться о словаре, пришла в голову ещё Льюису Кэрроллу: 28 из 39 страниц его книги как раз и занимает перечень английских слов, которые можно использовать в игре. Кроме того, условимся, что мы не используем в игре букву Ё и заменяем её на Е.

Всего в «Грамматическом словаре» 1712 четырёхбуквенных существительных. Возьмём, к примеру, существительное ЛОЖЬ и изобразим его точкой. Найдём в словаре все слова, которые отличаются от него на одну букву; их ровно четыре: ЛОЖА, ЛОЖЕ, ЛОСЬ и РОЖЬ. Изобразим их точками, соединёнными со словом ЛОЖЬ; наличие связи означает, что между словами можно перейти за один ход. (Кстати, какую ещё пару слов надо не забыть соединить?) Затем добавим слова, которые за один ход получаются из этих четырёх слов: КОЖА, ЛОЗА, ЛУЖА, ЛЫЖА, РОЖА, РОЛЬ, ЛОСК, – и нарисуем нужные связи.



В итоге у нас получился граф, в котором некоторые из 12 точек (вершин) соединены отрезками (рёбрами). Если нам нужно превратить одно слово в другое, на математическом языке это формулируется так: найти путь между соответствующими вершинами, желательно кратчайший. Так, если нам надо пройти от ЛЫЖА до РОЛЬ, это займёт четыре шага,

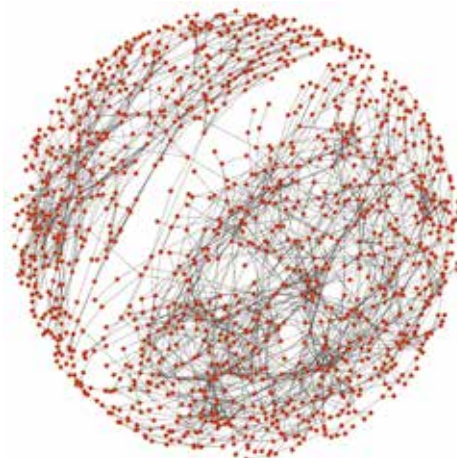


ИГРЫ И ГОЛОВОЛОМКИ

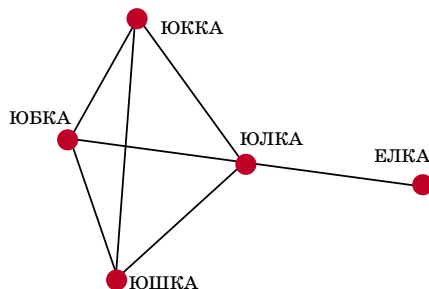


и пути могут быть разными: ЛЫЖА → ЛОЖА → → РОЖА → РОЖЬ → РОЛЬ или ЛЫЖА → ЛОЖА → → ЛОЖЬ → РОЖЬ → РОЛЬ. (Докажите, что более короткого пути между словами ЛЫЖА и РОЛЬ нет.)

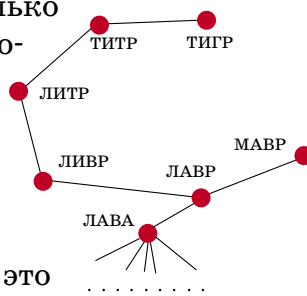
А теперь изобразим так не 12 слов, а все 1712, и посмотрим, как устроен этот граф. Вручную это сделать едва ли возможно, так что понадобится компьютер. Видно, что на графе выделяется большой кусок, где от любого слова можно прийти до любого другого; в теории графов такой подграф называют *компонентой связности*. Есть ещё несколько таких кусков поменьше и много точек, которые не связаны вообще ни с чем (такие отдельно стоящие точки тоже можно считать компонентами связности). Ясно, что от одного слова можно прийти до другого тогда и только тогда, когда они входят в одну и ту же компоненту связности.



Самая большая компонента связности включает в себя 1361 слово (то есть 79,5% всех слов). Кроме неё есть компонента размером 11 слов, ещё одна – размером 6 слов, 3 – размером 5 слов, 4 – размером 4 слова, 14 – размером 3 слова, 25 – размером 2 слова и ещё 211 отдельно стоящих слов (АЛОЭ, ВДОХ, ДЖАЗ, НЕБО, ОПЫТ, СОЮЗ, ТАЙМ и другие). Слово КЛЕН входит в большую компоненту связности, а слово ЕЛКА – в маленькую, 5-словную; этим и объясняется тот факт, что из слова КЛЕН не получится ЕЛКА.



Всего в нашем графе 3172 ребра, а значит, у слова в среднем $\frac{3172}{1712} \cdot 2 = 3,7$ соседей. Возвращаясь к самой большой компоненте связности, обратим внимание на то, что из неё торчат «хвосты». Дело в том, что даже в ней у некоторых слов совсем мало соседей. Скажем, от слова ТИГР можно перейти только к слову ТИТР, от него – только к слову ЛИТР, дальше – только к слову ЛИВР (это старинная французская монета, вспомните «Трёх мушкетёров»), дальше – только к слову ЛАВР, а уже от него – к словам МАВР и ЛАВА, после чего возможностей становится резко больше: от слова ЛАВА отходит ещё 5 слов, и мы попадаем в основную гущу вершин.



Таким образом, слово ТИГР – это конец хвоста, и если понадобится пройти из одного такого хвоста в другой, путь может получиться очень длинным, даже если это кратчайший путь между этими двумя вершинами. Самые длинные пути имеют длину 23 – например, от слова ДЖИП до слова ТУЕС (берестяная коробочка) всего 22 шага: ДЖИП → ДЖИН → УЖИН → УДИН → ОДИН → ОВИН → ОВЕН → ОВЕС → СВЕС → СВЕТ → СВАТ → СВАН → СТАН → СТЕН → СТЕК → САЕК → РАЕК → РОЕК → БОЕК → БУЕК → БУЕР → ТУЕР → ТУЕС. Глядя на эту цепочку, вам наверняка хочется пожаловаться: «Я же не знаю половины этих слов!» (признаюсь честно: я тоже не знаю). Но раз мы договорились использовать «Грамматический словарь», то и будем на него опираться, а к борьбе с незнакомыми словами вернёмся чуть позже.

Для 5-буквенных слов английского языка такой граф впервые построил знаменитый американский учёный и автор классических пособий по программированию Дональд Кнут. А почему, кстати, у него 5 букв, а у нас – 4? Есть ли какое-то объяснение тому, что в игре «Из мухи – слона» по-русски обычно играют 4-буквенными словами? Интуитивно кажется, что так интереснее всего. Но попробуем оценить этот интерес и количественно.

Окончание следует



Художник Мария Усеинова