

Деятельность 5

Двадцать предположений—Теория информации

Краткое содержание

Какое количество информации содержится в 1000 страницах книги? Это больше, чем количество информации в 1000 страницах телефонной книги, или в папке незаполненных бланков на 1000 строках, или в книге «Властелин колец» Толкиена? Если мы сможем измерить это, то у нас будет возможность оценить объём необходимого пространства для хранения этой информации. Например, вы сможете прочитать следующее предложение?

Ths sntnc hs th vwls mssng.

Вы, вероятно, прочтёте это сообщение, так как гласные звуки не содержат большого количества информации. В данном виде деятельности мы знакомимся с способом измерения информационного содержания.

Учебные направления

Математика: Уровень 3 и выше «Числа». Исследование числа: больше, чем, меньше чем, диапазоны. Модели и последовательности.
Английский язык

Навыки

Сравнение чисел и работа с диапазоном чисел
Дедукция
Задание вопросов

Возраст

От 10 и старше

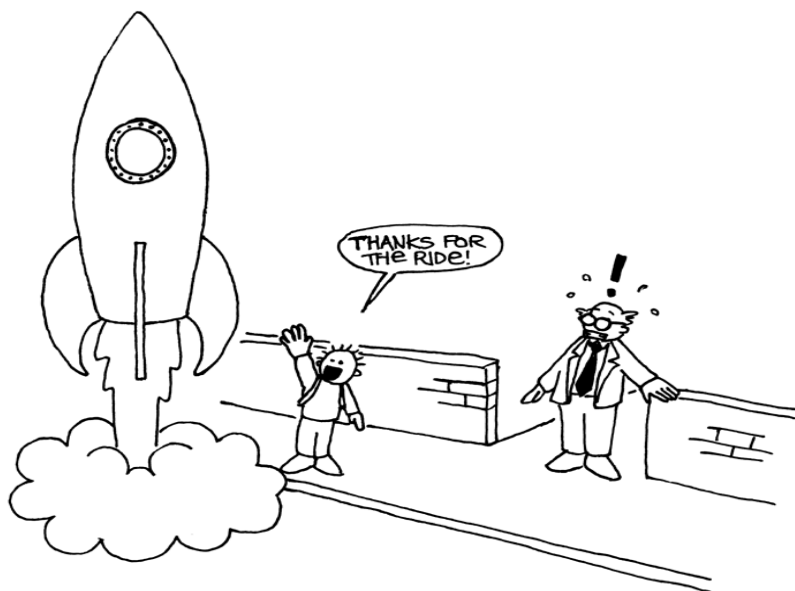
Материалы

Для основной деятельности нет необходимости в материалах
Для расширенной деятельности каждому ребёнку необходимо:
Лист деятельности: Деревья решений (стр. 4)

Двадцать предположений

Обсуждение

1. Обсудите с детьми понятие информации.
2. Как мы можем определить, сколько информации в книге? В книге важнее количество страниц или количество слов? Может ли в одной книге содержаться больше информации, чем в другой? Что если эта книга очень скучная, или очень интересная? Будет ли на 400 страницах книги, содержащей фразу «вздор, вздор, вздор» больше информации, чем скажем, в телефонном справочнике?
3. Объясните, что в специалисты в компьютерных науках, меру информации определяют по тому, как их удивила информация (или книга!). Говоря вам кое-что, что вы уже знаете, например, когда друг, который всегда идет в школе, говорит вам, « что я иду в школу сегодня», не сообщает вам информации, потому что это не удивительно. Если же он вместо этого сказал бы, что «я сегодня прилечу в школу на вертолете», то такое сообщение удивительно, и поэтому будет содержать определенное количество информации.
4. Как можно удивительные сообщения измерить?
5. Один из методов состоит в том, чтобы измерить интенсивность угадывания информации. Если ваш друг скажет: «Угадай, как я добрался до школы сегодня», а так как вы шли вместе, то, вероятно, вы догадаетесь с первого раза. Потребовалось бы большее количество предположений прежде, чем вы бы угадали способ передвижения – вертолет, и еще больше, если бы это был космический корабль.
6. Количество информации, которую содержат сообщения, измеряется тем, насколько легко или трудно они угадываются. Следующая игра даст нам представление об этом.



Деятельность: двадцать вопросов

Эта адаптированная игра - т20 вопросов. Дети могут задавать вопросы, одному из ребят, на которые можно ответить: «да» или «пока ответ не угадан». Задать можно любой вопрос при условии, что ответом будет строго: «да» или «нет».

Предложения:

Я думаю, что:

число между 1 и 100

число между 1 и 1000

число между 1 и 1,000,000.

любое целое число

последовательность из 6 чисел в шаблоне (для соответствующей группы).

Угадайте, в порядке от первого до последнего (например, 2, 4, 6, 8, 10)

Подсчитайте число вопросов, которые были заданы. Это и будет количеством «информации».

Последующее обсуждение

Какие стратегии вы использовали? Какие из них были лучшими? Покажите, что требуется только 7 предположений, для определения числа в диапазоне между 1 и 100, если вы будете делить каждый оставшийся диапазон пополам. Например:

Это число - меньше чем 50?	Да.
Это число - меньше чем 25?	Нет.
Это число - меньше чем 37?	Нет.
Это число - меньше чем 43?	Да.
Это число - меньше чем 40?	Нет.
Это число - меньше чем 41?	Нет.
Это должно быть 42!!	Да!

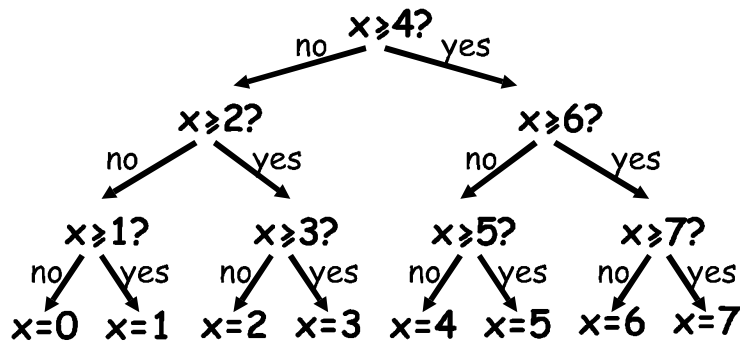
Интересно, если диапазон увеличить до 1000, то для определения задуманного числа, необходимо увеличить число вопросов не в 10 раз, а только на три вопроса. Каждый раз, при удвоении диапазона, необходим просто еще один вопрос для нахождения ответа. Полезно, в дальнейшем, будет игра для детей Mastermind.

Для любознательных: Какое количество информации в сообщении?

IT-специалисты не только угадывают числа, но и предполагают: какой символ, более вероятно, будет следующим в слове или в предложении. Попробуйте угадывание в форме игры с короткими предложениями из 4-6 слов. Буквы должны будут угадываться в определенном порядке: с первой по последнюю. Пусть кто-нибудь записывает буквы и ведет учет всех предположений, необходимых для определения каждой буквы. Могут быть заданы любые вопросы с ответами да/нет. Например, «Это буква - т?»; «Это буква - гласная?»; «Она находится в алфавите перед буквой м?». Пробел между словами, также считается символом, и должен быть угадан. Если вы сможете обнаружить наиболее узнаваемые части сообщений, то угадывание начните именно с них.

Лист деятельности: Деревья решений

Если вы уже знаете стратегию, как задать вопросы, то сможете передать сообщение без предварительного опроса. Вот диаграмма, называемая «деревом решений» для того чтобы определить число между 0 и 7:



Как указать путь «да/нет» в дереве решений при угадывании числа 5?

Сколько ходов «да/нет» в дереве решений вы должны сделать, чтобы определить некоторое число?

Сейчас увидим что-то очень увлекательное. В нижнем ряду дерева решений под числами 0, 1, 2, 3... запишите двоичные числа (см. деятельность 1).

Посмотрите внимательно на дерево. Если «нет» обозначим 0, а «да» – 1, что вы тогда увидите?

В игре на угадывание чисел, мы пытаемся выбрать, таким образом, вопросы, чтобы последовательность ответов было представлена одинаковым количеством ветвей дерева решений.

Создайте своё дерево решений для угадывания чисел от 0 до 15.

Дополнительно для любознательных: Какое «дерево» вы бы использовали для угадывания возраста кого-нибудь? Как бы вы представили «дерево» для угадывания в предложении последовательности букв?

О чём это всё?

Известный американский математик (фокусник, моноциклист) Клод Шеннон провел много экспериментов с этой игрой. Он измерил объём информации в битах, где каждый ответ: да/нет - эквивалентен 1/0 бит. Он обнаружил, что количество «информации», содержащейся в сообщении зависит от того, что вам уже известно. Иногда вы можете задать один вопрос, который избавит вас от необходимости задавать ещё несколько вопросов. В этом случае сообщение содержит небольшое количество информации. Например, один бит: орёл или решка, как правило, - это информация об одном бросании монеты. Но если монета будет деформированной (неровной), и при рассмотрении переворотов: в девяти случаях из десяти - количество информации уже будет не один бит - верьте, хотите нет, а меньше. Как можно вам связать процесс бросания монеты и задания более одного вопроса, ответом на который будет: да/нет? Просто - только использовать вот такие вопросы: «при следующих двух бросках монеты выпадает орёл?» Около 80% случаев, при последовательном бросании деформированных монет, ответом будет - «да». В то время, на 20%, вам придётся ответить - «нет», при этом ещё задать два дополнительных вопроса. Но в среднем, вы будете уже задавать больше чем один вопрос по сравнению с одним бросанием монеты! Шеннон называл **информационное содержание сообщения «энтропией»**. Энтропия зависит, во-первых, не только от числа возможных исходов бросания одной монеты, но и во-вторых, также от вероятности, что это вообще произойдет. Для угадывания сообщения в случае, невероятных или удивительных событий, необходимо задать больше чем обычно вопросов, поскольку эти события содержат больше информации, по сравнению с обычным сообщением, например, в случае посадки вертолета в школе.



Энтропия сообщения очень важна для компьютерной науки. Вы можете сжать сообщение, только до её энтропии, и уменьшить занимаемое этим сообщением место, и лучшие системы сжатия, подобны игре на угадывания. Поскольку компьютерная программа осуществляет «угадывание» и ответы хранятся в виде битов, то перечень вопросов может быть воспроизведен позже, при восстановлении информации. Лучшие системы сжатия могут уменьшить текстовые файлы примерно в четверть от их первоначального размера - это большая экономия памяти! Метод угадывания может быть использован для создания компьютерного интерфейса, позволяющего предсказывать, что пользователь собирается ввести в забросе! Этот метод может быть полезен людям с ограниченными возможностями, которым трудно набрать текст. Компьютер предлагает или угадывает возможные варианты забросов пользователей, а они соглашаются или нет. Хорошая система нуждается, в среднем, в двух ответах «да/нет» на символ, и может быть большим подспорьем для тех, у кого проблемы в точных движениях, необходимых для управления мышью или клавиатурой. Такие системы используются также в случаях «набора» текста в некоторых мобильных телефонах.

Решения и подсказки

Ответ «да / нет» на вопрос - соответствует ровно одному биту информации, будь то простой вопрос, как: «Это число больше, чем на 50?»; или более сложный, такой, как: «Это число - между 20 и 60?»

При угадывании чисел, если вопросы выбраны определенным образом, последовательность ответов представлена только двоичными числами. Число три - «011» в двоичной системе, и оно представлено ответами «Нет, да, да». Что, то же самое, если мы запишем: «нет» - числом 0 и «да» - числом 1, в дереве решений.

Дерево можно использовать для определения возраста кого-нибудь, как смещение в сторону уменьшения значения возраста.

Решение о букве в предложении, зависит от того, какова была предыдущая буква.