



# Введение в теорию информации. Методы и средства определения количества информации

Быстренина Ирина Евгеньевна,  
кандидат педагогических наук, доцент  
кафедры прикладной информатики  
РГАУ-МСХА имени К. А. Тимирязева





Под **информацией** понимается совокупность сведений об объектах, явлениях окружающей среды, их свойствах, состояниях и взаимосвязи.

Информация используется во всех областях человеческой деятельности. Установление любой коммуникации между людьми или координации их действий в группе возможны только благодаря обмену информацией.



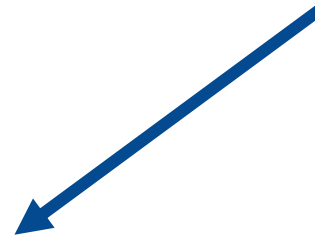
# Измерение информации

---

Важнейшим результатом теории информации является следующий вывод: в определенных условиях можно пренебречь качественными особенностями информации, выразить её количество числом, а также сравнить количество информации, содержащейся в различных группах данных.



## Подходы к измерению информации



Вероятностный

Алфавитный



# Вероятностный подход к измерению информации

---

При этом подходе информацию, содержащуюся в сообщении, можно нестрого трактовать в смысле её новизны или, иначе, уменьшения неопределённости наших знаний об объекте.

# Вероятностный подход к измерению информации

---



Для характеристики неопределенности используется понятие **энтропии**.  
Энтропия и информация - полярные категории.



# Вероятностный подход к измерению информации

---

Американский инженер Ральф Хартли в 1928 г. рассматривал процесс получения информации как выбор одного сообщения из конечного заданного множества из  $N$  равновероятных сообщений, а количество информации  $I$ , содержащееся в выбранном сообщении, определял как двоичный логарифм  $N$ .

Формула Хартли:  $I = \log_2 N$



## Вероятностный подход к измерению информации

---

Допустим, нужно угадать одно число из набора чисел от единицы до ста. По формуле Хартли можно вычислить, какое количество информации для этого требуется:  $I = \log_2 100 = 6,644$ . Таким образом, сообщение о верно угаданном числе содержит количество информации, приблизительно равное 6,644 единицы информации.





# Вероятностный подход к измерению информации

---

К равновероятным событиям также можно отнести:

- Выпадение «орла» или «решки» при подбрасывании монеты;
- Наличие в книге четного или нечетного числа букв.



## Вероятностный подход к измерению информации

---

Для задач, в которых вероятность исходов не одинакова, американский учёный Клод Шеннон предложил в 1948г. другую формулу определения количества информации, учитывающую возможную неравную вероятность сообщений в наборе.



# Вероятностный подход к измерению информации

---

Формула Шеннона:

$$I = - (p_1 * \log_2 p_1 + p_2 * \log_2 p_2 + \dots + p_N * \log_2 p_N),$$

где  $p_i$  — вероятность того, что именно  $i$ -е сообщение выделено в наборе из  $N$  сообщений.



# Вероятностный подход к измерению информации

---

Если вероятности  $p_1, \dots, p_N$  равны, то каждая из них равна  $1/N$ , и формула Шеннона превращается в формулу Хартли.



# Единицы измерения информации

---

В качестве единицы измерения количества информации Клод Шеннон предложил принять один бит (англ. bit — binary digit — двоичная цифра).



# Единицы измерения информации

---

**Бит** в теории информации — количество информации, необходимое для указания одного из двух равновероятных сообщений (типа "орел" — "решка", "чет" — "нечет" и т.п.).



# Единицы измерения информации

---

В вычислительной технике битом называют наименьшую "порцию" памяти компьютера, необходимую для хранения одного из двух знаков "0" и "1", используемых для внутри машинного представления данных и команд.



## Единицы измерения информации

---

На практике чаще применяется более крупная единица измерения количества информации — байт, равная восьми битам. Именно восьмибитовая конструкция требуется для того, чтобы закодировать любой из 256 символов алфавита клавиатуры компьютера ( $256=2^8$ ).





# Единицы измерения информации

---

Широко используются также ещё более крупные производные единицы информации:

- 1 Килобайт (Кбайт) = 1024 байт =  $2^{10}$  байт,
- 1 Мегабайт (Мбайт) = 1024 Кбайт =  $2^{20}$  байт,
- 1 Гигабайт (Гбайт) = 1024 Мбайт =  $2^{30}$  байт.



# Единицы измерения информации

---

- 1 Терабайт (Тбайт) = 1024 Гбайт =  $2^{40}$  байт,
- 1 Петабайт (Пбайт) = 1024 Тбайт =  $2^{50}$  байт.



## Единицы измерения информации

---

За единицу информации можно было бы выбрать количество информации, необходимое для оценки, например, десяти равновероятных сообщений. Это будет не двоичная (бит), а десятичная (дит) единица информации.

Формула нахождения количества информации в этом случае имеет вид:

$$I = \lg N$$



# Единицы измерения информации

---

В некоторых случаях в качестве основания логарифма принимается число  $e$ , такая единица измерения количества информации называется **нат**.

$$I = \ln N$$



# Адекватность информации

---

Адекватность информации может выражаться в трех формах:

- Семантической;
- Синтаксической;
- Прагматической.



## Синтаксическая адекватность

---

**Синтаксическая адекватность** отображает формально-структурные характеристики информации, не затрагивая ее смыслового содержания. На синтаксическом уровне учитываются тип носителя и способ представления информации, скорость передачи и обработки, размеры кодов представления информации, надежность и точность преобразования этих кодов и т.п.



## Синтаксическая адекватность

---

Информацию, рассматриваемую с синтаксических позиций, обычно называют данными, так как при этом не имеет значения смысловая сторона. Эта форма способствует восприятию внешних структурных характеристик информации.



## Семантическая адекватность

---

**Семантическая (смысловая) адекватность** определяет степень соответствия образа объекта и самого объекта. Семантический аспект предполагает учет смыслового содержания информации. На этом уровне анализируются те сведения, которые отражает информация, рассматриваются смысловые связи. В информатике устанавливаются смысловые связи между кодами представления информации.





## Семантическая адекватность

---

Семантическая форма служит для формирования понятий и представлений, выявления смысла, содержания информации и ее обобщения.



## Прагматическая адекватность

---

**Прагматическая (потребительская) адекватность** отражает отношение информации и ее потребителя, соответствие информации цели управления, которая на ее основе реализуется. Проявляются прагматические свойства информации только при наличии единства информации (объекта), пользователя и цели управления.



## Прагматическая адекватность

---

Прагматический аспект рассмотрения связан с ценностью, полезностью использования информации при выработке потребителем решения для достижения своей цели. С этой точки зрения анализируются потребительские свойства информации. Эта форма адекватности непосредственно связана с практическим использованием информации.



# Адекватность информации

---

При работе с информацией всегда имеется ее источник и потребитель (получатель). Ресурсы и процессы, обеспечивающие передачу сообщений от источника информации к ее потребителю, называются информационными коммуникациями.

Для потребителя информации важной характеристикой является ее адекватность.



# Адекватность информации

---

**Адекватность информации** - это определенный уровень соответствия образа, создаваемого с помощью полученной информации, реальному объекту, процессу, явлению и т.п.



---

**Спасибо за внимание!**