

# Измерение информации



В настоящее время не существует внятного определения Информации, и само это понятие окружено мистикой. Цель данных заметок дать техническое определение этому понятию и предложить способы его количественной оценки.

В 1948 году Клодом Шенноном была предложена расчетная формула для вычисления того, что он (крайне неудачно) назвал "количество информации".

Один из устойчивых в нетехнических кругах мифов заключается в том, что шенноновское количество информации якобы как-то связано с информацией (в ее обычном, обыденном смысле). В самых запущенных случаях, оценку Шеннона пытаются отождествить с термодинамической энтропией (т.н. негэнтропия Бриллюэна) и использовать в энергетических расчетах.

Прежде, чем перейти к дальнейшему изложению, сразу заметим, что шенноновское "количество информации" не имеет ни малейшего отношения к информации (ни как к понятию - качеству, ни как к количеству этого качества) и является исключительно внутренней характеристикой самого сообщения. И уж тем более, не имеет никакого отношения к термодинамической энтропии.

## Терминология

*Система* - нечто, имеющее различные состояния.

*Измерение* - процесс синхронизации тестовой и эталонной систем до полного совпадения состояний

*Эталонная система* - неизменяемая в процессе измерения, принимая за образец.

*Тестовая система* - изменяемая в процессе измерения до совпадения с эталонной.

*Информация о тестовой системе* - разность состояний эталонной и тестовой системы

## Коммуникация по Шеннону

В своей статье 1948 года Шеннон пишет:

*"Основная задача связи состоит в точном или приближенном воспроизведении в некотором месте сообщения, выбранного для передачи в другом месте. Часто сообщения имеют значение т. е. относятся к некоторой системе, имеющей определенную физическую или умозрительную сущность, или находятся в соответствии с некоторой системой. Эти семантические аспекты связи не имеют отношения к технической стороне вопроса."*

С технической точки зрения, речь идет о копире - устройстве, для точного воспроизведения образца. Примером такого копирования является взвешивание на рычажных весах. Весовщик (активный элемент) выполняет роль третьих сил, компенсирующих потери в канале связи. Взвешивание происходит путем точного уравнивания весов на обеих чашах. Процедуру добавления гирь будем называть сообщением.

### Три необходимых условия измерения:

До начала измерения обе системы синхронизированы (чаши весов уравновешены)

Обе сравниваемые системы функционально идентичны (в данном случае, весы равноплечие)

Взвешивание происходит до полной синхронизации состояний систем (уравнивания весов)

Применительно к передаче сообщений, предварительная синхронизация и функциональная идентичность означают наличие общего языка, алфавита, шифра, "секрета" итп. Коммуникация считается завершенной после полного получения сообщения.

Из рассмотрения взвешивания немедленно выводится важное наблюдение ("инвариант взвешивания"): набор гирь может быть произвольным, при этом сумма произведений числа гирь на их веса равна взвешиваемому грузу ("количество информации" по Шеннону).

Рассмотрение процесса взвешивания позволяет дать общее определение информации.

Будем рассматривать две синхронизируемые системы как два фрактала, единственными измеряемыми характеристиками которых являются их геометрический размер и плотность (два объема газа, два текста итп).

Число несовпадающих состояний обеих систем (геометрический размер) назовем семантической компонентой информации, и плотность фрактала (информационную энтропию по Шеннону) синтаксической компонентой. Их произведение (мера) есть полная информация об измеряемой системе.

Может показаться разочаровывающим, что интегральная оценка (полная информация) ничего не говорит о структуре объекта. На самом деле, это следует из самого свойства меры - отсутствия формы и бесконечной делимости. Измерение - это эквивалентирование фракталом. При этом масса фрактала - единственная доступная характеристика.

Например, если имеется 0.5 жидкости, то неизвестна ни форма сосуда, в которую она налита, ни даже один это сосуд или разлито на троих.