Лекция 3.

Энтропия, информация, семантика

В физическом пространстве любое взаимодействие характеризуется инвариантом - энергией, для которой сформулированы законы сохранения.

В информационном пространстве модели взаимодействия объектов имеют другую природу, которая связана с понятием информации как меры разнообразия рассматриваемых объектов. В этом смысле информация является объективной характеристикой реальности.

1. Что такое мера? Это положительное число.

Формально понятие меры в математике связано с обобщением «длины», площади или объема.

Мера вводится на множестве объектов и является положительным числом, которое сопоставляется этому множеству по некоторым правилам. Если правило такое, что число существует, то множество называется измеримым по отношению к этой мере.

Мера м.б. НОЛЬ! Множество меры ноль – это точка, поле рациональных числе, а мера поля вещественных чисел это не ноль, а континуум. = С.

Вернемся к информации. Как сформировать меру «разнообразия»?. Для этого есть несколько возможностей.

Первая – это сформировать количественную меру информации, которая важна для систем связи и передачи информации

Вторая – сформировать смысловую меру информации, которая связана с «моделью мира».

Итак, знания об окружающем мире составляются из

Тип знания Область

1. физический мир
2. мыслимое
3. исчислимое

e\*\*(-i\*pi)=1.

Информационная **двоичная энтропия** для независимых случайных событий *x* с *n* возможными состояниями (от 1 до *n*, *p* - функция вероятности) рассчитывается по формуле:

H(x)=-\sum_{i=1}^np(i)\log_2 p(i).

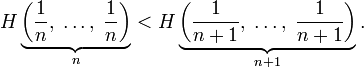
Эта величина также называется *средней энтропией сообщения*. Величина \log_2\frac{1}{p(i)}называется *частной энтропией*, характеризующей только *i*-e состояние.

Таким образом, энтропия события *x* является суммой с противоположным знаком всех произведений относительных частот появления события *i*, умноженных на их же двоичные логарифм. Это определение для дискретных случайных событий можно расширить для функции распределения [вероятностей](http://ru.wikipedia.org/wiki/%D0%92%D0%B5%D1%80%D0%BE%D1%8F%D1%82%D0%BD%D0%BE%D1%81%D1%82%D1%8C).

Итак, прирост информации равен утраченной неопределённости.

Мера информации непрерывна; то есть изменение значения величины вероятности на малую величину должно вызывать малое результирующее изменение функции;

 Для целых положительных *n*, выполняется следующее неравенство:



энтропия является разницей между информацией, содержащейся в сообщении, и той частью информации, которая точно известна (хорошо предсказуема) в сообщении. Примером этого является [избыточность языка](http://ru.wikipedia.org/wiki/%D0%98%D0%B7%D0%B1%D1%8B%D1%82%D0%BE%D1%87%D0%BD%D0%BE%D1%81%D1%82%D1%8C_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0) — имеются явные статистические закономерности в появлении букв, пар последовательных букв, троек и т.

Все это справедливо в контексте вероятностной модели для [источника данных](http://ru.wikipedia.org/w/index.php?title=%D0%98%D1%81%D1%82%D0%BE%D1%87%D0%BD%D0%B8%D0%BA_%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D1%85&action=edit&redlink=1).

 Если X,\;Yнезависимы, то H(X\cdot Y)=H(X)+H(Y).

 Энтропия — выпуклая вверх функция распределения вероятностей элементов.

 Если X,\;Yимеют одинаковое распределение вероятностей элементов, то *H*(*X*) = *H*(*Y*).

Учет корреляции данных происходит с использованием условной энтропии.

Единица измерения — бит/два символа, это объясняется тем, что взаимная энтропия описывает неопределённость на пару символов: отправленного и полученного. Путём несложных преобразований также получаем

H(AB)=H(A)+H(B\mid A)=H(B)+H(A\mid B).

Взаимная энтропия обладает свойством *информационной полноты* — из неё можно получить все рассматриваемые величины.

СЕМАНТИЧЕСКИЙ АСПЕКТ ИНФОРМАЦИИ [semantic aspect of information] — характеристика информации с точки зрения ее смысла, содержания. Для восприятия информации необходимо, чтобы передаваемые сообщения в определенной мере соответствовали тезаурусу знаний получателя: если они не имеют точек соприкосновения с ним, сообщение понято не будет (так, человек, не знающий математики, не поймет математическую формулу).

**Сема́нтика** (от [др.-греч.](http://ru.wikipedia.org/wiki/%D0%94%D1%80%D0%B5%D0%B2%D0%BD%D0%B5%D0%B3%D1%80%D0%B5%D1%87%D0%B5%D1%81%D0%BA%D0%B8%D0%B9_%D1%8F%D0%B7%D1%8B%D0%BA) σημαντικός — обозначающий) — раздел [языкознания](http://ru.wikipedia.org/wiki/%D0%9B%D0%B8%D0%BD%D0%B3%D0%B2%D0%B8%D1%81%D1%82%D0%B8%D0%BA%D0%B0), изучающий [значение](http://ru.wikipedia.org/wiki/%D0%97%D0%BD%D0%B0%D1%87%D0%B5%D0%BD%D0%B8%D0%B5) [единиц языка](http://ru.wikipedia.org/wiki/%D0%95%D0%B4%D0%B8%D0%BD%D0%B8%D1%86%D1%8B_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0).

**СЕМАНТИКА,** в широком смысле слова –отношение между языковыми выражениями и миром, реальным или воображаемым.

каким образом человек, зная слова и грамматические правила какого-либо естественного языка, оказывается способным передать с их помощью самую разнообразную информацию о мире (в том числе и о собственном внутреннем мире),

Как понять, какую информацию о мире заключает в себе любое обращенное к человеку высказывание, даже если человек впервые слышит его.

Та сущность в мире, к которой относится слово, называется его референтом. Так, если я, описывая кому-то произошедшее событие, говорю: *Вчера я посадил под своим окном дерево*, то слово *дерево* отсылает к единичной индивидуальной сущности – тому самому единственному в своем роде дереву, которое я вчера посадил под своим окном.

Существует множество сущностей, которые могут быть обозначены с помощью данного слова

Вещи, достаточно сходные друг с другом, мы называем одним и тем же именем. Деревья отличаются друг от друга размерами, очертаниями, распределением листвы, но они обладают некоторыми сходными чертами, позволяющими называть их все деревьями.

необходимым различать языковое значение слова и связанное с этим словом мыслительное содержание – понятие. И языковое значение, и понятие являются категориями мышления. То и другое суть отражения мира в нашем сознании. Но это разные виды отражения.

Многие слова (возможно, даже большинство слов) используются в более чем одном смысле. Cлово *лук* может применяться как для обозначения огородного растения со съедобной луковицей и съедобными трубчатыми листьями, так и для обозначения старинного оружия для метания стрел.