



OPEN.AZ



Уральский
федеральный
университет

ИНФОРМАТИКА

Старший преподаватель департамента информационных технологий и автоматике
Шеклеин Алексей Александрович

- 17 лекций по 2 часа в неделю
- 17 лабораторных 2 часа в неделю
- Контрольные работы + домашние задания
- Экзамен
- При оценке знаний и достижений используется БРС

- Лекция1 Измерение информации
- Лекция2 Коды
- Лекция3 Сжатие
- Лекция4 Системы счисления, Кодирование чисел
- Лекция5 Представление информации в ЭВМ
- Лекция6 MS ExcelГрафики
- Лекция7 MS ExcelФункции
- Лекция8 MS ExcelДанные
- Лекция9 Алгоритм, Блок-схема
- Лекция10 Алгоритмы
- Лекция11 Программное обеспечение
- Лекция12 Операционная система
- Лекция13 Поколения вычислительной техники
- Лекция14 Аппаратное обеспечение, ПК
- Лекция15 Аппаратное обеспечение, ПК
- Лекция16 Компьютерные сети
- Лекция17 Защита данных

ЛИТЕРАТУРА

- 1.** Информатика. Базовый курс / под ред. С.В. Симоновича. 3-е изд. - СПб.: 2011. — 640 с
- 2.** Информатика: учебник. 3-е перераб. изд. / под ред. проф. Н.В. Макаровой. М.: Финансы и статистика, 2001. 768 с.: ил. (и последующих лет издания)
- 3.** Информатика / Луенбергер Дэвид Дж. М.: Техносфера, 2008. – 448 с.

Методические материалы:

study.urfu.ru

<http://learn.urfu.ru> СДО УрФУ Гиперметод

ЛЕКЦИЯ 1. ИНФОРМАТИКА: ПРЕДМЕТ И ЗАДАЧИ.

- Понятие и структура информатики
- Информация и данные
- Свойства информации
- Формы адекватности информации
- Носители информации

ПОЯВЛЕНИЕ И РАЗВИТИЕ ИНФОРМАТИКИ

Термин *информатика* возник в 60-х гг.

Французский термин *informatique* (информатика) образован путем слияния слов *information* (информация) и *automatique* (автоматика) и означает «информационная автоматика» или «автоматизированная переработка информации».

В англоязычных странах этому термину соответствует синоним *computer science* (наука о компьютерной технике).

Другие термины:

information technology,

information science.

Информатика – это область человеческой деятельности, связанная с процессами преобразования информации с помощью компьютеров и их взаимодействием со средой применения.

РОЛЬ ИНФОРМАТИЗАЦИИ В РАЗВИТИИ ОБЩЕСТВА

- **Информационные ресурсы** – это отдельные документы и отдельные массивы документов в информационных системах (библиотеках, архивах, фондах, банках данных, других информационных системах).
- **Информатизация общества** – организованный процесс (социально-экономический и научно-технический) создания оптимальных условий для удовлетворения информационных потребностей граждан, работы органов власти и управления, организаций и общественных объединений на основе формирования и использования информационных ресурсов.
- **Информационная культура** – умение целенаправленно работать с информацией и использовать для её получения, обработки и передачи компьютерную информационную технологию, современные технические средства и методы.

Структура информатики



Отрасль
народного
хозяйства

Производство
технических
средств

Производство
программных
продуктов

Разработка
технологий
переработки
информации

Фундаментальная
наука

Методология
создания
информационного
обеспечения

Теория
информационных
систем и
технологий

Прикладная
дисциплина
(для конкретных
областей)

Изучение
закономерностей в
информационных
процессах

Создание
информационных
моделей
коммуникаций

Разработка
информационных
систем и
технологий

ПЛАТФОРМЫ ПРАКТИЧЕСКОЙ ДЕЯТЕЛЬНОСТИ В ОБЛАСТИ ИНФОРМАТИКИ

- архитектура вычислительных систем;
- программирование;
- преобразование данных;
- защита информации;
- автоматизация;
- стандартизация;
- интерфейсы вычислительных систем:
 - 1) аппаратные интерфейсы
 - 2) программные интерфейсы
 - 3) аппаратно-программные интерфейсы
 - 4) пользовательский интерфейс

ИНФОРМАЦИЯ И ДАННЫЕ

- Термин ***информация*** происходит от латинского ***informatio*** (разъяснение, осведомление, изложение)
- С позиций материалистической философии информация – отражение реального мира с помощью сведений (сообщений).

Сообщение – это форма представления информации в виде речи, текста, изображения, цифровых данных, графиков, таблиц и т.п.

- **Данные** – это зарегистрированные сигналы. Они несут в себе информацию о событиях, произошедших в материальном мире, т.к. являются регистрацией сигналов, возникших в результате этих событий.

ВЗАИМОСВЯЗЬ ПОНЯТИЙ «ИНФОРМАЦИЯ» И «ДААННЫЕ»

Информация является продуктом взаимодействия данных и соответствующих им методов восприятия.

Данные (data) – это записанные наблюдения, признаки, которые по какой-то причине только хранятся, а не используются.

В том случае, если появляется возможность использовать эти данные для уменьшения неопределенности о чем-либо, данные превращаются в информацию.

Таким образом, информация – это используемые данные.

- **Информация** – сведения об объектах и явлениях окружающей среды, их параметрах, свойствах и состоянии, которые уменьшают имеющуюся степень неопределенности, неполноту знаний о НИХ.

Пусть

α – система (объект реального мира);

$H(\alpha)$ – энтропия системы, мера недостающей информации о системе;

β – сообщение, дополнительные сведения о системе;

$H_{\beta}(\alpha)$ – неопределенность состояния системы после получения сообщения β ;

$I_{\beta}(\alpha)$ – количество информации о системе, полученной в сообщении β .

$$I_{\beta}(\alpha) = H(\alpha) - H_{\beta}(\alpha)$$

- Таким образом, информация – это используемые данные.

ДИАЛЕКТИЧЕСКОЕ ЕДИНСТВО ДАННЫХ И МЕТОДОВ ИХ ОБРАБОТКИ В ИНФОРМАЦИОННОМ ПРОЦЕССЕ

- Динамический характер информации (существует только в момент взаимодействия данных и методов, все прочее время пребывает в состоянии данных).
- Требование адекватности методов. Одни и те же данные могут в момент потребления поставлять разную информацию в зависимости от степени адекватности взаимодействующих с ними методов.
- Диалектический характер данных и методов. Информация возникает и существует в момент диалектического взаимодействия объективных данных и субъективных методов.

СВОЙСТВА ИНФОРМАЦИИ

- Объективность
- Полнота
- Достоверность
- Доступность
- Актуальность
- Адекватность

АДЕКВАТНОСТЬ ИНФОРМАЦИИ

- **Адекватность информации** – это определенный уровень соответствия создаваемого с помощью полученной информации образа реальному объекту, процессу, явлению и т.п.
- При работе с информацией всегда имеется **источник информации** и **потребитель** (получатель).
 - Информационные коммуникации – передача сообщений от источника информации ее получателю.
 - Для потребителя информации очень важной характеристикой является ее адекватность.
- От степени адекватности информации реальному объекту, процессу или явлению зависит правильность принятых решений человеком.
- Адекватность информации можно оценивать в разных формах.

Формы адекватности информации

Синтаксическая
адекватность

Семантическая
(смысловая)
адекватность

Прагматическая
адекватность

ИЗМЕРЕНИЕ ИНФОРМАЦИИ

Меры информации

Синтаксическая
мера

Семантическая
мера

Прагматическая
мера

Объем данных
 V_D

Количество
Информации
 $I_C = C * V_D$

Количество
информации
 $I_B(\alpha)$

СИНТАКСИЧЕСКАЯ МЕРА ИНФОРМАЦИИ

Не затрагивает смыслового содержания

Объем данных V_d

Объем данных в сообщении измеряется количеством символов (разрядов) в этом сообщении.

Например:

- в двоичной системе счисления единица измерения – **бит** (bit - binary digit - двоичный разряд);
- в десятичной системе счисления единица измерения – **дит** (десятичный разряд).

ИНФОРМАЦИЯ, НЕОПРЕДЕЛЕННОСТЬ, ВОЗМОЖНОСТЬ ВЫБОРА

Количество информации $I_{\beta}(\alpha)$

- Любая неопределенность предполагает – возможность выбора.
- При полной информации выбора нет (нет неопределенности).
- Частичная информация уменьшает число вариантов выбора, сокращая тем самым неопределенность.
- *Пример 1.* Бросание монетки. Обе стороны равноправны. Начальная неопределенность характеризуется **двумя** возможными состояниями. После того как монета упала, достигается полная ясность и неопределенность исчезает.
- *Пример 2.*

Лампа А	0	0	1	1
Лампа В	0	1	0	1

СИНТАКСИЧЕСКАЯ МЕРА ИНФОРМАЦИИ

$$N = m^n$$

где

N – число всевозможных отображаемых состояний системы;

m – основание системы счисления (разнообразие символов, применяемых в алфавите);

n – число разрядов (символов) в сообщении.

Например:

$N = 2^8 = 256$ – в одном байте (8 разрядов или битов) можно закодировать 256 различных символов

$N = 2^{16} = 65536$ – в двух байтах (16 разрядов)

Синтаксическая адекватность отображает формально-структурные характеристики информации и не затрагивает ее смыслового содержания.

Здесь учитываются:

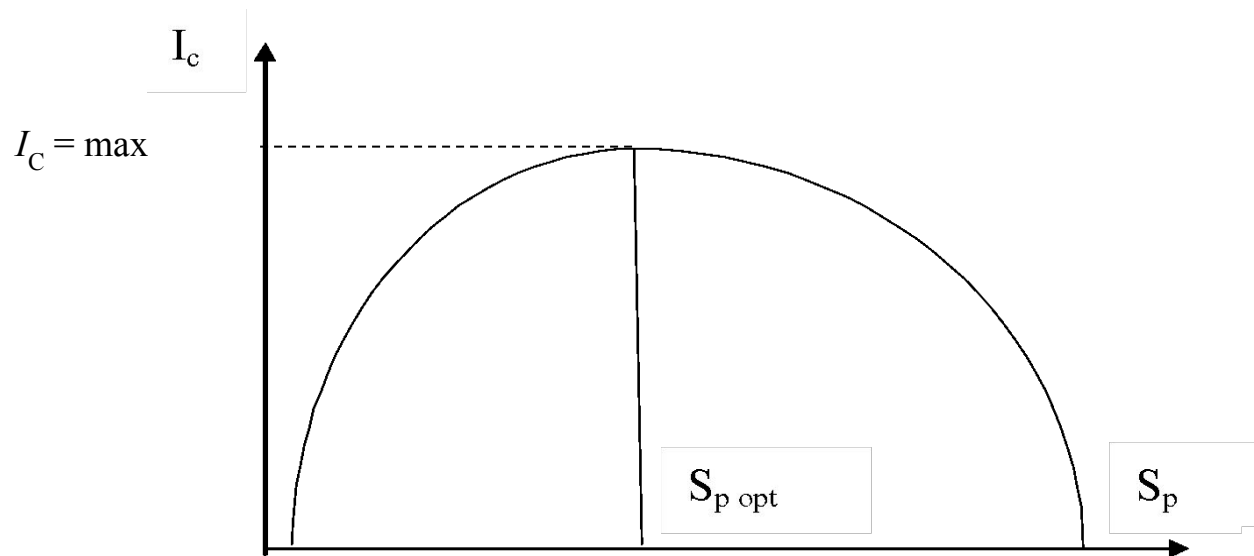
- тип носителя и способ представления информации,
- скорость передачи и обработки,
- размеры кодов её представления,
- надежность и точность преобразования этих кодов

и т. п.

Информацию рассматривают как данные.

СЕМАНТИЧЕСКАЯ МЕРА ИНФОРМАЦИИ

- Оценивает смысловое содержание.
- Тезаурус – это совокупность сведений, которыми располагает пользователь или система.



Коэффициент содержательности C – есть отношение количества семантической информации к ее объему:

$$C = I_c / V_d$$

ПРАГМАТИЧЕСКАЯ МЕРА ИНФОРМАЦИИ

- определяет *полезность* информации (*ценность*) для достижения пользователем поставленной цели.
- величина относительная, обусловленная особенностями использования этой информации в той или иной системе.
- измеряется в тех же самых единицах (или близких к ним), в которых измеряется целевая функция.

ПРАГМАТИЧЕСКАЯ МЕРА ИНФОРМАЦИИ

Пример

В экономической системе прагматические свойства (ценность) информации можно определить приростом экономического эффекта функционирования, достигнутым благодаря использованию этой информации для управления системой:

$$I_{\Pi} \beta (\gamma) = \Pi(\gamma/\beta) - \Pi(\gamma), \text{ где}$$

$I_{\Pi} \beta (\gamma)$ – ценность информационного сообщения β для системы управления γ ;

$\Pi(\gamma)$ – априорный экономический эффект функционирования системы управления γ ;

$\Pi(\gamma/\beta)$ – ожидаемый эффект функционирования системы γ при условии, что для управления будет использована информация, содержащаяся в сообщении β .

ЕДИНИЦЫ ИЗМЕРЕНИЯ ИНФОРМАЦИИ И ПРИМЕРЫ

Мера информации	Единицы измерения	Примеры (для компьютерной области)
<p>Синтаксическая: шенноновский подход</p> <p>компьютерный подход</p>	<p>Степень уменьшения неопределенности</p> <p>Единицы представления информации</p>	<p>Вероятность события</p> <p>Бит, байт, Кбайт и т.д.</p>
<p>Семантическая</p>	<p>Тезаурус</p> <p>Экономические показатели</p>	<p>Пакет прикладных программ, персональный компьютер, компьютерные сети и т.д.</p> <p>Рентабельность, производительность, коэффициент амортизации и т. д.</p>
<p>Прагматическая</p>	<p>Ценность использования</p>	<p>Емкость памяти, производительность компьютера, скорость передачи данных</p> <p>Денежное выражение</p> <p>Время обработки информации и принятия решений</p>

НОСИТЕЛИ ДАННЫХ

- Данные – зарегистрированные сигналы.
- В соответствии с методом регистрации данные могут храниться и транспортироваться по-разному (на различных носителях)

Носители данных

- Основанные на изменении оптических свойств
 - бумага
 - CD-диски.
- Использующие магнитные свойства материала
 - магнитные ленты и диски.

хим. состава веществ – в фотографии,

на биохимическом уровне – накопление и передача данных в живой природе.

ХАРАКТЕРИСТИКИ НОСИТЕЛЕЙ ДАННЫХ

Разрешающая способность – количество данных, записанных в соответствующей носителю единице измерения.

Динамический диапазон – логарифмическое отношение интенсивности амплитуд максимального и минимального регистрируемого сигналов.

Пример. Визуальный эффект от просмотра слайдов выше, чем от фото, т.к. диапазон яркости сигналов в проходящем света на 2-3 порядка выше, чем в отраженном.

Задача преобразования данных с целью смены носителя относится к одной из важнейших задач информатики.



ОПЕРАЦИИ С ДАННЫМИ

- сбор данных;
- формализация данных;
- фильтрация данных;
- сортировка данных;
- архивация;
- защита данных;
- передача;
- преобразование данных (*например, изменение типа носителя*),

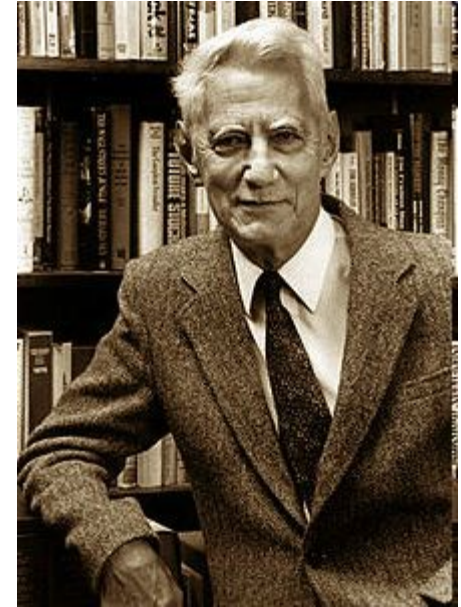
ЛЕКЦИЯ 2. ОПРЕДЕЛЕНИЕ И ИЗМЕРЕНИЕ ИНФОРМАЦИИ

- О Клоде Э. Шенноне
- Вероятность сообщения и количество информации
- Понятие энтропии
- Свойства энтропии
- Источники информации
- Биты в качестве меры информации

О Клоде Э. Шенноне

Клод Элвуд Шеннон родился в 1915 году в г. Петоски, штат Мичиган. Он учился в Университете штата Мичиган, где получил степени бакалавра по электротехнике и бакалавра по математике. Затем учился в Массачусетском технологическом институте и получил степень магистра по электротехнике и степень доктора философии по математике в 1940 году (обе степени в одно и то же время). За диссертацию на соискание степени магистра был награжден премией Альфреда Нобеля, инженерной премией, присуждаемой молодым авторам.

После окончания МТИ Шеннон был принят на работу в компанию «Белл Телефон Лабораториз», сотрудником которой он оставался в течение 15 лет. Именно здесь он разработал свою теорию коммуникации, выполнил свое исследование структуры английского языка и развил теорию кодирования.



Клод Элвуд Шеннон
(1915 – 2001)

Немного о теории информации Шеннона

Теория информации Клода Е. Шеннона была опубликована в середине прошлого века.

Проблема (или задача), которая решалась – быстрая передача информации без ошибок по каналам связи.

Основная идея теории – повышение надёжности передачи данных не повторением (повторение – самый простой способ), а кодированием каждого символа группой, в которую должна быть заложена избыточность.

ЧТО ТАКОЕ ИЗБЫТОЧНОСТЬ

Можно пояснить на примере:

связисты во всех странах, передавая буквы голосом, проговаривают не то как буква звучит, а слово, либо начинающееся на эту букву, либо кодирующее этот знак:

«У» – Ульяна;

«П» – Пётр;

«С»(англ.) – Charlie.

МЕРА ИНФОРМАЦИИ

Шеннон в своей теории использовал **вероятностный подход**.

Идея меры информации следующая:

Чем реже появляются сообщения (т.е. вероятность их появления маленькая), тем больше информации они содержат. И наоборот, сообщения, появляющиеся часто, содержат меньше информации.

Если **вероятность p** ($0 \leq p \leq 1$) **сообщения** (т.е. частота его появления) известна, то можно определить соответствующую информацию.

Определение информации. Информация, соответствующая сообщению с вероятностью p , равняется:

$$I = \log(1 / p) \equiv -\log p$$

Для логарифма может быть использовано любое основание (в частности, **основание e** натуральных логарифмов, **основание 10** или **основание 2**).

Разные основания просто изменяют единицы меры информации.

Данная формула была первоначально предложена Ральфом В.Л. Хартли(1888 – 1970).

Он пользовался **логарифмами с основанием 10** (десятичными логарифмами). При этом единицы информации принято называть **Хартли** или **диты**.

Логарифмы с основанием 2 наиболее часто используются в теории информации, и в этом случае единицами информации являются **биты**.

Например, если $p = 1/2$, тогда

$$I = -\log_2(1/2) = \log_2 2 = 1 \text{ бит.}$$

Для преобразования одного основания в другое
можно воспользоваться следующим
отношением:

$$\log_b x = \log_a x / \log_a b$$

Например, преобразование натурального
логарифма в логарифм по основанию 2:

$$\log_2 x = \ln x / \ln 2,$$

где \ln – логарифм по основанию e .

Поскольку $\ln 2 = 0,693$, можно записать:

$$\log_2 x = \ln x / 0,693.$$

Суммируемость информации

Свойство суммируемости. Если событие A имеет вероятность p_A , а событие B имеет вероятность p_B , и являются независимыми событиями (одно событие не находится под влиянием другого), в этом случае вероятность совместного события A и B равняется $p_A p_B$. Соответствующая информация равняется:

$$I_{AB} = -\log p_A p_B = -\log p_A - \log p_B = I_A + I_B.$$

ЭНТРОПИЯ

Энтропия – это мера информации, которую мы ожидаем получить в будущем (мера недостающей информации).

Энтропия – это средняя информация, принятая в отношении всех возможных исходов (сообщений).

Энтропия для двух событий. Пусть событие A имеет вероятность $p_A = p$, а событие B имеет вероятность $p_B = 1 - p$.

Тогда для двух событий с вероятностями p и $1 - p$ энтропия равна:

$$H(p) = -p \log p - (1 - p) \log(1 - p)$$

Пример 1

Погода. Предположим, что погода в Калифорнии либо солнечная, либо облачная с вероятностью $7/8$ и $1/8$ соответственно. Энтропией этого источника информации является средняя информация о солнечных и облачных днях. Отсюда:

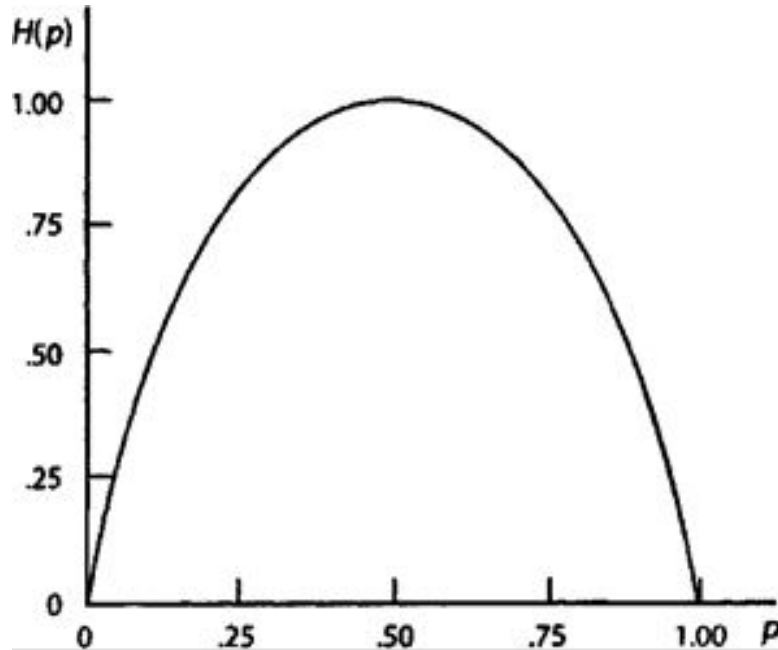
$$\begin{aligned} H &= -(7/8)\log(7/8) - (1/8)\log(1/8) = \\ &= -\frac{1}{8}[7\log 7/8 + \log 1/8] = -\frac{1}{8}[7\log 7 - 7\log 8 - \log 8] = \\ &= -\frac{1}{8}[7 \times 2,81 - 7 \times 3 - 3] = \\ &= -\frac{1}{8}[19,65 - 21 - 3] = \frac{1}{8}[4,349] = 0,54 \text{ бита} \end{aligned}$$

Примечание. В данном расчете:

$\log 8 = 3$, потому что $2^3 = 8$;

$\log 7$ находится из $\log_2 7 = \ln 7 / \ln 2 = 1,946 / 0,693 = 2,81$.

Энтропия $H(p)$ как функция p .



Энтропия симметрична относительно точки $1/2$, в которой она достигает максимального значения, равного 1 биту. Энтропия равна 0, если p равняется нулю или единице.

ИСТОЧНИКИ ИНФОРМАЦИИ

Источник информации или просто источник по определению состоит из всех возможных (взаимно исключающих) событий.

Предположим, что в источнике имеется n возможных событий, причем i -е событие имеет вероятность p_i для $i = 1, 2, \dots, n$.

$$p_i > 0; \sum_{i=1}^n p_i = 1$$

Информация сообщения о том, что событие i произошло, равняется

$$I_i = \log(1/p_i).$$

Энтропия равняется средней информации этих событий.

Энтропия источника n событий. Энтропия источника n событий с вероятностями p_1, p_2, \dots, p_n равняется:

$$\begin{aligned} H &= p_1 \log(1/p_1) + p_2 \log(1/p_2) + \dots + p_n \log(1/p_n) = \\ &= -[p_1 \log p_1 + p_2 \log p_2 + \dots + p_n \log p_n] = \\ &= -\sum_{i=1}^n p_i \log p_i. \end{aligned}$$

Эта функция иногда обозначается $H(p_1, p_2, \dots, p_n)$.

Пример 2

Источник из трех событий. Предположим, что имеется три события с вероятностями $1/2$, $1/4$, $1/4$. Соответствующая энтропия равняется:

$$\begin{aligned} H(1/2, 1/4, 1/4) &= (1/2)\log_2 + (1/4)\log_4 + (1/4)\log_4 = \\ &= 1/2 + (1/4) \cdot 2 + (1/4) \cdot 2 = 3/2 \text{ бита} \end{aligned}$$

СВОЙСТВА ЭНТРОПИИ

1. **Неотрицательность** $H(p_1, p_2, \dots, p_n) \geq 0$.

Доказательство: поскольку $0 \leq p_i \leq 1$, каждый $\log p_i \leq 0$. Отсюда $-p_i \log p_i \geq 0$ для каждого i , что означает $H \geq 0$.

2. $H(p_1, p_2, \dots, p_n) \leq \log n$.

Максимальное значение H имеет место, когда все вероятности равны, при этом $p_i = 1/n$ для каждого i .

Отсюда

$$H \leq \sum_{i=1}^n (1/n) \log n = \log n$$

3. Суммирующее свойство энтропии. Если источники S и T независимы, в этом случае энтропия $H(S, T)$ источника произведения (S, T) удовлетворяет следующему уравнению:

$$H(S, T) = H(S) + H(T).$$

Доказательство:

Математически два источника S и T являются независимыми, если вероятность каждой пары (s, t) при $s \in S, t \in T$ равняется $p_{st} = p_s p_t$, где p_s и p_t – это вероятности s и t соответственно.

Суммируемость вытекает из **свойства логарифмов**, а именно

$$\log p_s p_t = \log p_s + \log p_t.$$

Формально произведение двух источников S и T обозначается (S, T) и состоит из всех возможных пар (s, t) событий, одно из S и одно из T .

Энтропия S^n

Рассмотрим источник, являющийся произведением любого числа n независимых событий от S , и обозначим этот источник S^n .

Энтропия S^n . Когда независимые образцы берутся из источника S с энтропией $H(S)$, энтропия результирующего источника S^n равняется:

$$H(S^n) = nH(S).$$

Смешанная энтропия

Смешанная энтропия. Энтропия источника, полученного смешиванием независимых источников S_1 и S_2 с учетом вероятностей p и $1-p$ соответственно, равняется:

$$H = pH_1 + (1 - p)H_2 + H(p)$$

где H_1 – это энтропия S_1 , а H_2 – это энтропия S_2 , p – вероятность смешивания.

Пример 3

Пусть S_1 будет монеткой, а S_2 шестисторонней костью. Смешанный источник с вероятностью p будет бросать монетку для того, чтобы получить «орла» или «решку», или же (с вероятностью $1 - p$) будет бросать кость для того, чтобы получить 1, 2, 3, 4, 5 или 6.

Тогда:

$H_1 = \log 2 = 1$; $H_2 = \log 6$ (см. свойство 2 энтропии).

Если $p = 1/2$, то имеем:

$$H = 1/2(1 + \log 6) + H(1/2) = 2 + 1/2 \log 3.$$

(!!!Здесь испол. сво-во логарифмов: $\log 6 = \log 2 + \log 3$).

Биты в качестве меры

Бит – это единица измерения, часто используемая в науках об информации.

Обычно, количество бит – это фактическое количество двоичных разрядов, использованных в представлении информации. (Например, выражение 010111 имеет длину, равную 6 битам.)

С комбинаторной точки зрения, все, что имеет n возможностей (или состояний), как обычно говорят, имеет $\log_2 n$ битов.

И наоборот, k битов могут представлять в общей сложности 2^k предметов.

Различие!

В общем виде прямая мера битов, когда они приводятся в виде символов, соответствует мере энтропии, только если все символы встречаются **в равной степени вероятно и являются взаимно независимыми.**

Ни простая комбинаторная мера битов, ни мера энтропии не сообщает ничего о полезности информации, измеряемой в битах.

Термины, определяющие крупные количества битов

байт = 8 битов

килобайт = 2^{10} байтов = 1024 байт

мегабайт = 2^{20} байтов

гигабайт = 2^{30} байтов

терабайт = 2^{40} байтов

петабайт = 2^{50} байтов

эксабайт = 2^{60} байтов

зеттабайт = 2^{70} байтов

йоттабайт = 2^{80} байтов

ВЫВОДЫ

1. Информатика – это широкое понятие, область человеческой деятельности, связанная с обработкой информации на компьютерах.
2. Информация и данные – это различные понятия.
3. На синтаксическом уровне информацию рассматривают как данные и количество информации измеряют с помощью энтропии, битов, байтов ...
4. Количество информации в одном сообщении (событии) зависит от вероятности его получения
5. Энтропия – это среднее количество информации источника с несколькими событиями

Задание 1

1. Рассмотрите источник с четырьмя событиями, имеющими вероятности $1/5$, $1/5$, $1/5$, $2/5$.

(а) Чему равна информация в битах, сообщаемая в репортаже о том, что произошло первое событие?

(б) Какова энтропия источника?

Подсказка: $\ln 2 = 0,693$, $\ln 5 = 1,61$

$$I_{S_1} = -\log(1/5) = \log(5) = \ln 5 / \ln 2 = 1,61 / 0,693 = 2,32 \text{ бит}$$

$$\begin{aligned} H &= (1/5) \log(5) + (1/5) \log(5) + (1/5) \log(5) + (2/5) \log(5/2) = \\ &= (3/5) \log 5 + (2/5) \log(5) + (2/5) \log(1/2) = \\ &= \log 5 - (2/5) = \\ &= 2,32 - 0,4 = 1,92 \text{ бит} \end{aligned}$$

2. (Изменение основания.) Какова общая формула энтропии $H_b(S)$ с использованием логарифмов с основанием b относительно энтропии $H_a(S)$ с использованием логарифмов с основанием a ?

3. (Максимальная энтропия.) Убедительно докажите, что максимально возможная энтропия источника, состоящего из n событий, равняется $\log n$ бит и достигается, когда события имеют равные вероятности.

4. (Смешивание источников.) Профессор Бэббл пишет математическую работу, представляющую собой сочетание английского языка и математики. Энтропия на один символ его английского языка равна H_E , а энтропия его математики (использование математических символов) равняется H_M . Его работа состоит из доли λ английских букв и доли $1-\lambda$ математических символов.

(а) Докажите, что энтропия на один символ его работы равняется:

$$H_p = \lambda H_E + (1 - \lambda) H_M + H(\lambda)$$

(б) Профессор гордится тем, что он сочетает в работе английский язык и математику таким образом, что его работа имеет максимальную энтропию на один символ. Найдите величину λ , которой он пользуется.

5. (Игральные карты.)

(а) Каков объем информации в битах передается в результате объявления названия выбранной карты из колоды, состоящей из 52 игральных карт?

(б) Каково общее количество способов, которыми может быть расположена колода? Подсказка: вначале найдите логарифм этого числа.

(в) Какова энтропия в битах источника, состоящего из произвольной колоды карт?

Спасибо за внимание!