

# Лекция 1

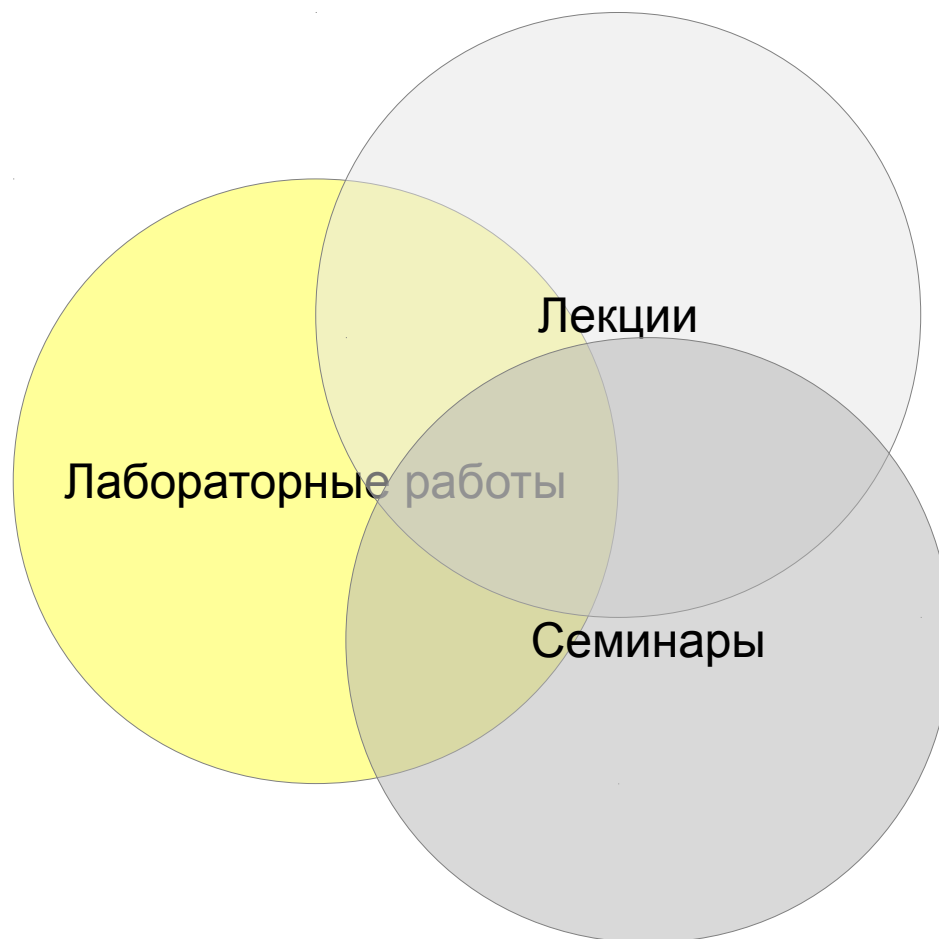
Технологии обработки информации

# Преподаватели

Арбатский Евгений Викторович  
A517

Федоров Вячеслав Викторович  
A501

# Состав курса



# Примерная структура курса

1. Введение. Виды информации.
2. Способы представления информации. Физические носители. Способы кодирования информации.
3. Анализ задачи, постановка задачи, создание отчетов с результатами. Презентации.
4. Поиск информации. Поиск на локальном компьютере.
5. Поиск в интернет. Поисковые системы, базы знаний.
6. Анализ информации. Методы статистического, визуального анализа данных.
7. Релевантность, онтологии.
8. Обработка информации. Электронная почта, базы данных, базы знаний.
9. Технические средства обработки информации. Средства для первичной обработки данных. Средства для подготовки данных для поиска.

# Адрес сайта

Персональная страница Арбатского Е.В.

<http://trpo.is-isea.ru/>

Персональная страница Фёдорова В.В.

<http://web-edu.iriit/sites/index.php?s=1>

<http://www.irgups.ru/web-edu/sites/index.php?s=1>

# Информация

**Информация** (от лат. informatio — разъяснение, изложение), первоначально — сведения, передаваемые одними людьми другим людям устным, письменным или каким-либо другим способом (например, с помощью условных сигналов, с использованием технических средств и т. д.), а также сам процесс передачи или получения этих сведений.

# Информация

Кодирование

Декодирование



Источник

Носитель

Приемник

# Развитие понятия

1. Обмен между человеком и автоматом, автоматом и автоматом;
2. Мера информации.



# Исследование информации

1. Математический аппарат;
2. Свойства информации;
3. Лингвистический анализ.

# Виды информации

1. В быту информацией называют любые данные, сведения, знания, которые кого-либо интересуют. Например, сообщение о каких-либо событиях, о чьей-либо деятельности и т.п.;
2. В технике под информацией понимают сообщения, передаваемые в форме знаков или сигналов (в этом случае есть источник сообщений, получатель (приемник) сообщений, канал связи);

# Виды информации

3. В кибернетике под информацией понимают ту часть знаний, которая используется для ориентирования, активного действия, управления, то есть в целях сохранения, совершенствования, развития системы;

4. В теории информации под информацией понимают сведения об объектах и явлениях окружающей среды, их параметрах, свойствах и состоянии, которые уменьшают имеющуюся о них степень неопределенности, неполноты знаний.

# Виды информации

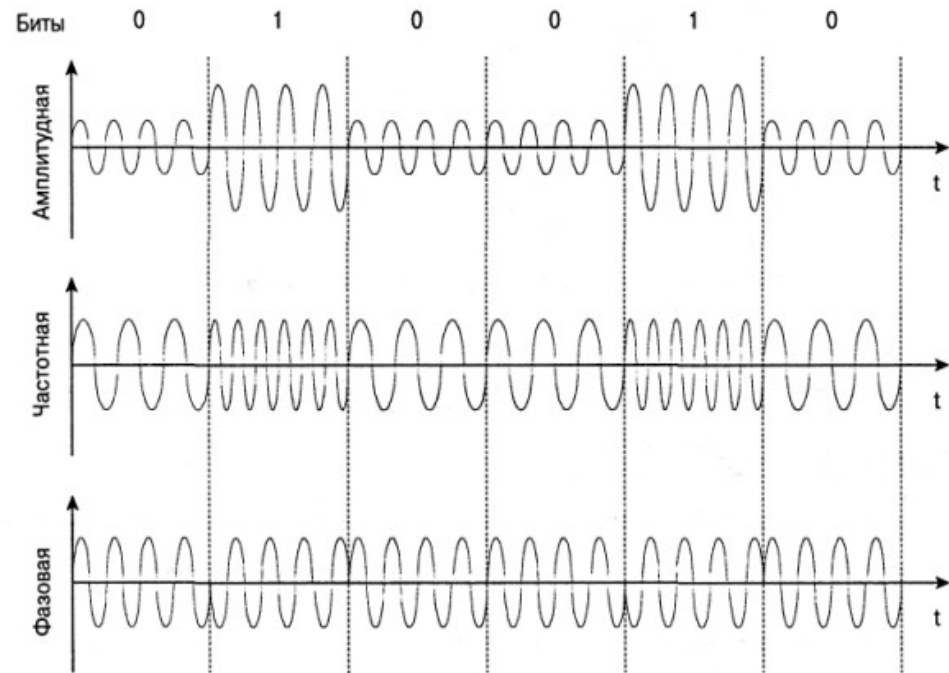
Дискретная

1, 2, 30, 400, 5124

Текстовая информация

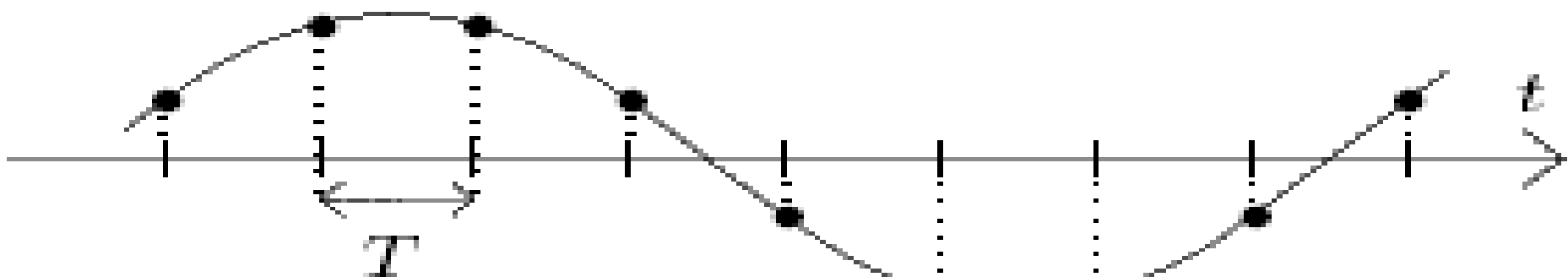
01010101000010101010

Аналоговая

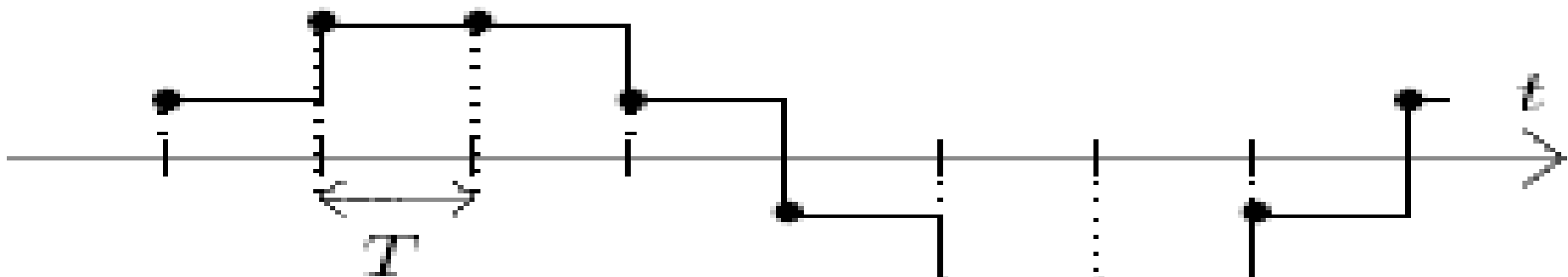


# Обработка аналоговой информации

Исходный сигнал



Дискретизированный сигнал



# Виды информации

- по способам восприятия (визуальная, тактильная, аудио и т.д.);
- по форме представления (текстовая, числовая, графическая, звуковая.);
- по общественному значению (массовая, специальная, личная).

# Свойства информации

1. Полнота;
2. Актуальность;
3. Достоверность;
4. Доступность;
5. Релевантность;
6. Эргономичность.

# Свойства информации

1. Запоминаемость;
2. Передаваемость;
3. Воспроизводимость;
4. Преобразуемость;
5. Стираемость.



# Носители информации

- любой материальный предмет (бумага, камень и т.д.);
- волны различной природы: акустическая (звук), электромагнитная (свет, радиоволна) и т.д.;
- вещество в различном состоянии: концентрация молекул в жидком растворе, температура и т. д.
- машинные носители информации: перфоленты, перфокарты, магнитные ленты, и т. д.

# Информатика

**Информатика** – область человеческой деятельности, связанная с процессами преобразования информации с помощью компьютеров и других средств вычислительной техники.

Информатика как совокупность средств преобразования информации включает технические средства (hardware), программные продукты (software), математические методы, модели и типовые алгоритмы (brainware).

# Измерение информации

Основоположник теории информации Клод Шеннон определил информацию, как снятую неопределенность. Точнее сказать, получение информации - необходимое условие для снятия неопределенности. Неопределенность возникает в ситуации выбора.

*Количество информации - мера уменьшения неопределенности знаний.*

# Измерение информации

$$I(x) = -\log p(x)$$

$I$  — количество информации

$P(X)$  — вероятность события  $X$

По основанию **2** — бит

По основанию **e** — нат

По основанию **10** — хартли

# Свойства информации

1. Информация не отрицательна ( $0 \leq P(X) \leq 1$ );
2. Свойство аддитивности.  $I(x,y) = I(x) + I(y)$

# Энтропия

$$H = - \sum_{i=1}^N p_i \log_2 p_i$$

Энтропия — это количество информации, приходящейся на одно элементарное сообщение источника, вырабатывающего статистически независимые сообщения.

$H$  — энтропия

$N$  - количество возможных событий

$P_i$  — вероятность события

# Энтропия

$$H = \log_2 N$$

Если события равновероятны

$H$  — энтропия

$N$  - количество возможных событий

# Измерение информации

За единицу количества информации принят 1 бит — количество информации, содержащееся в сообщении, уменьшающем неопределенность знаний в два раза.



# Единицы измерения

1 байт = 8 бит

1 Кбайт =  $2^{10}$  байт

1 Мбайт =  $2^{10}$  Кбайт =  $2^{20}$  байт

1 Гбайт =  $2^{10}$  Мбайт =  $2^{20}$  Кбайт =  $2^{30}$  байт

# Адрес сайта

Персональная страница Арбатского Е.В.

<http://trpo.is-isea.ru/>

Персональная страница Фёдорова В.В.

<http://web-edu.iriit/sites/index.php?s=1>

<http://www.irgups.ru/web-edu/sites/index.php?s=1>

# Бод

Бод (англ. baud) в связи и электронике — единица измерения символьной скорости, количество изменений информационного параметра несущего периодического сигнала в секунду. Названа по имени Эмиля Бодо, изобретателя кода Бодо — кодировки символов для телетайпов.

Зачастую ошибочно считают, что бод — это количество бит, переданное в секунду. В действительности же это верно лишь для двоичного кодирования, которое используется не всегда. Например, при символьной скорости 2400 бод скорость передачи может составлять 9600 бит/с благодаря тому, что в каждом временном интервале передаётся 4 бита.

Кроме этого, бодами выражают полную ёмкость канала, включая служебные символы (биты), если они есть. Эффективная же скорость канала выражается другими единицами, например битами в секунду (бит/с, bps).