

ТЕОРИЯ ИНФОРМАЦИИ

Кодирование информации в
информатике и в биологии

План занятия:

1. Систематизация и обобщение знаний по теории информации:
 - Основные понятия теории информации;
 - Кодирование текстовой информации;
 - Кодирование информации в живой природе.
2. Решение задач на кодирование информации.
3. Информационные процессы в живой природе.

Домашнее задание:

- Биология § 8;
- Информатика стр.107-121, Н.Д. Угринович;
стр. 104 – 108, Ю.А. Шафрин;
- Задача № 5;
- Подготовка к защите
лабораторной работы.

Информация

```
graph TD; A[Информация] --> B[Свойства информации]; A --> C[Виды информации]; A --> D[Единицы измерения информации]; A --> E[Информационные процессы];
```

Свойства
информации

Виды
информации

Единицы
измерения
информации

Информационные
процессы

КОДИРОВАНИЕ ТЕКСТОВОЙ ИНФОРМАЦИИ

Традиционно для кодирования одного символа используется количество информации, равное 1 байту.

$$N = 2^l = 2^8 = 256$$

Такое количество символов вполне достаточно для представления текстовой информации, включая прописные и строчные буквы русского и латинского алфавита, цифры, знаки, графические символы и пр.

Сущность кодирования

Каждому символу ставится в соответствие уникальный десятичный код от 0 до 255 или соответствующий ему двоичный код от 00000000 до 11111111.

Таким образом, человек различает символы по их начертаниям, а компьютер — по их кодам.

соглашение

Кодовая таблица

- Первые 33 кода (с 0 по 32) соответствуют не символам, а операциям (перевод строки, ввод пробела и так далее).
- Коды с 33 по 127 являются интернациональными и соответствуют символам латинского алфавита, цифрам, знакам арифметических операций и знакам препинания.
- Коды с 128 по 255 являются национальными .

Сегодня в России наиболее активно используются кодовые таблицы CP-1251, CP-866, CP KOI-8R и CP-ISO-8859-5.

Кодовые таблицы	использование
CP-1251	102589
KOI-8r	7335
ISO-8859-5	810
CP-866	575
CP-Mac	107

Сравнительная диаграмма

кодовые таблицы

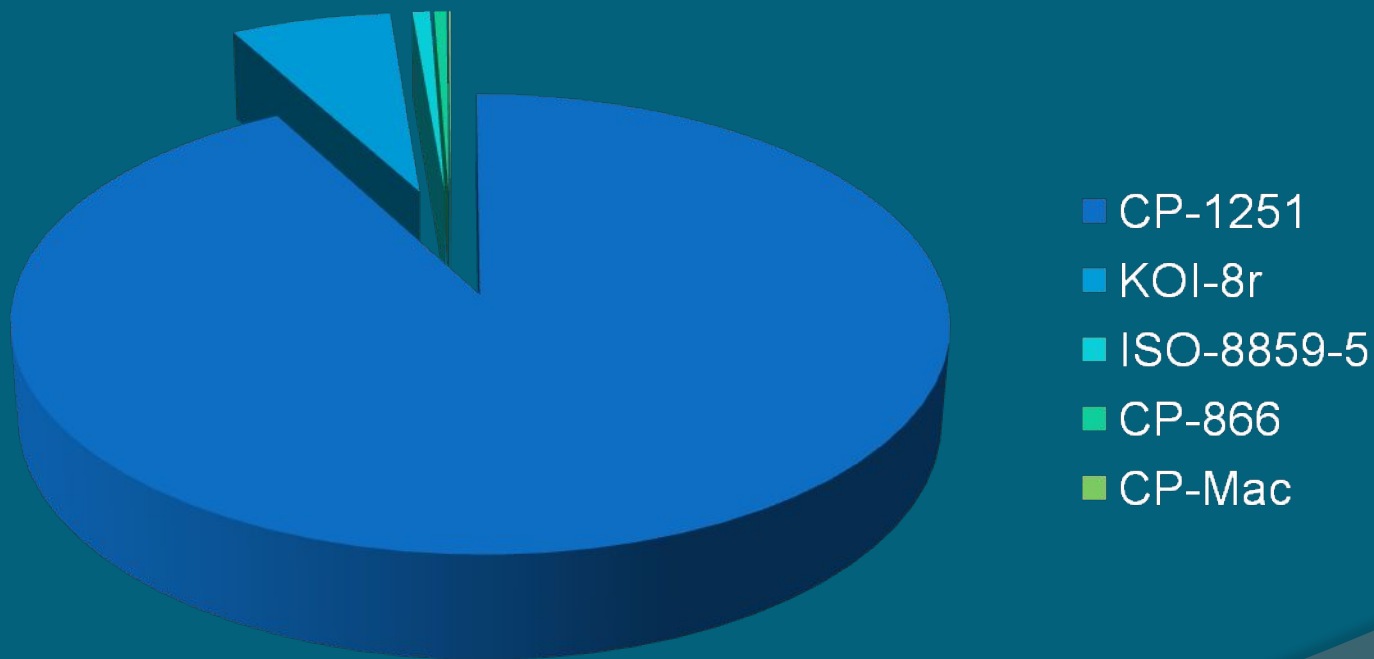
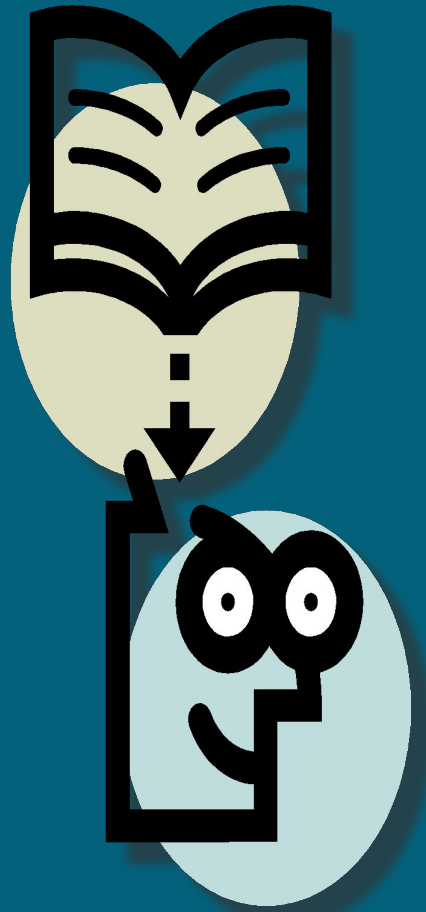


Таблица кодов ASCII по России

А	10000000	Р	10010000	а	10100000	р	11100000
Б	10000001	С	10010001	б	10100001	с	11100001
В	10000010	Т	10010010	в	10100010	т	11100010
Г	10000011	У	10010011	г	10100011	у	11100011
Д	10000100	Ф	10010100	д	10100100	ф	11100100
Е	10000101	Х	10010101	е	10100101	х	11100101
Ж	10000110	Ц	10010110	ж	10100110	ц	11100110
З	10000111	Ч	10010111	з	10100111	ч	11100111
И	10001000	Ш	10011000	и	10101000	ш	11101000
Й	10001001	Щ	10011001	й	10101001	щ	11101001
К	10001010	Ъ	10011010	к	10101010	ъ	11101010
Л	10001011	Ы	10011011	л	10101011	ы	11101011
М	10001100	Ь	10011100	м	10101100	ь	11101100
Н	10001101	Э	10011101	н	10101101	э	11101101
О	10001110	Ю	10011110	о	10101110	ю	11101110
П	10001111	Я	10011111	п	10101111	я	11101111

НАСЛЕДСТВЕННАЯ ИНФОРМАЦИЯ



- О ЧЕМ ?
- ГДЕ ХРАНИТСЯ ?
- КАК ЗАКОДИРОВАНА ?

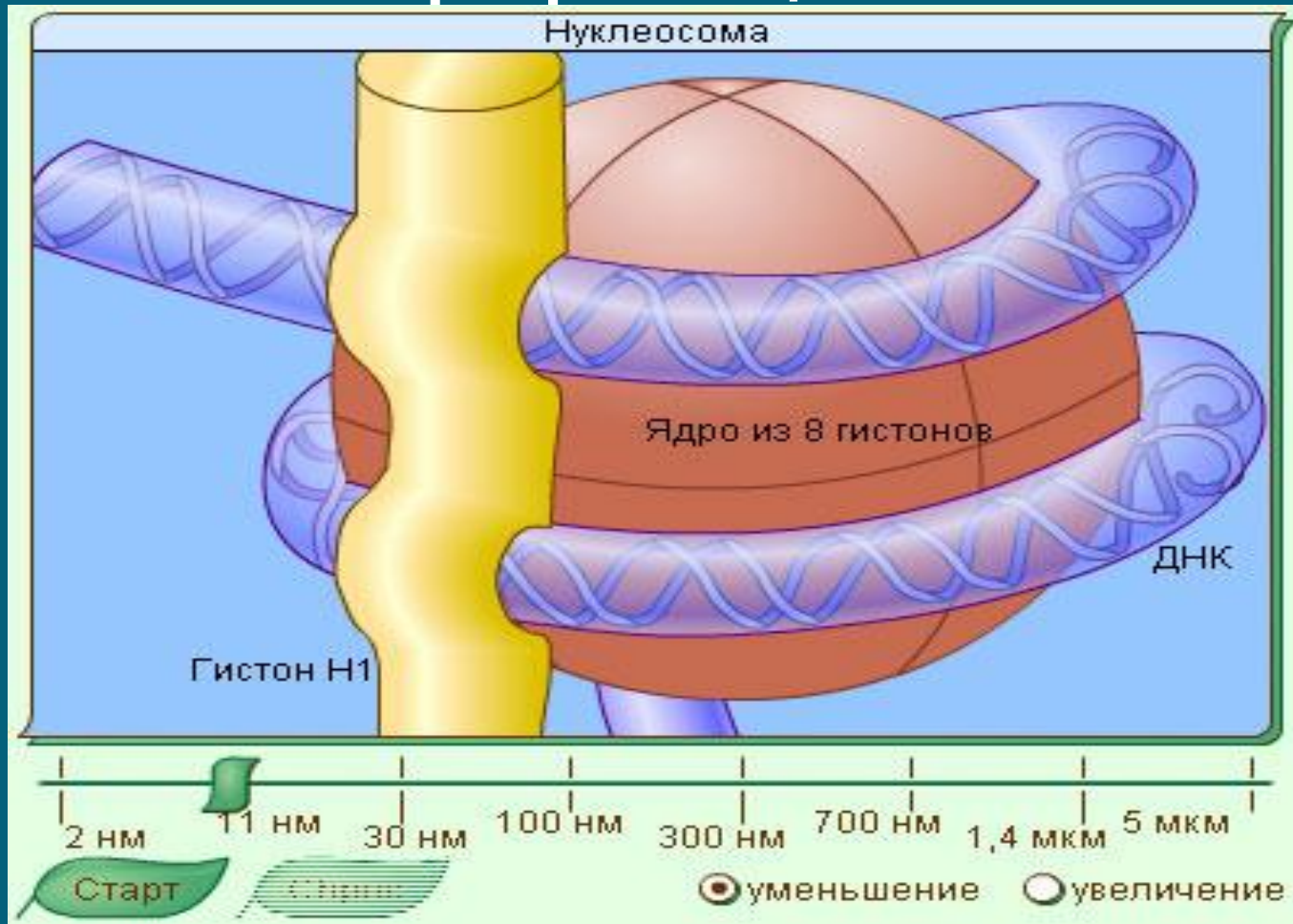
О ЧЕМ?

Наследственная информация

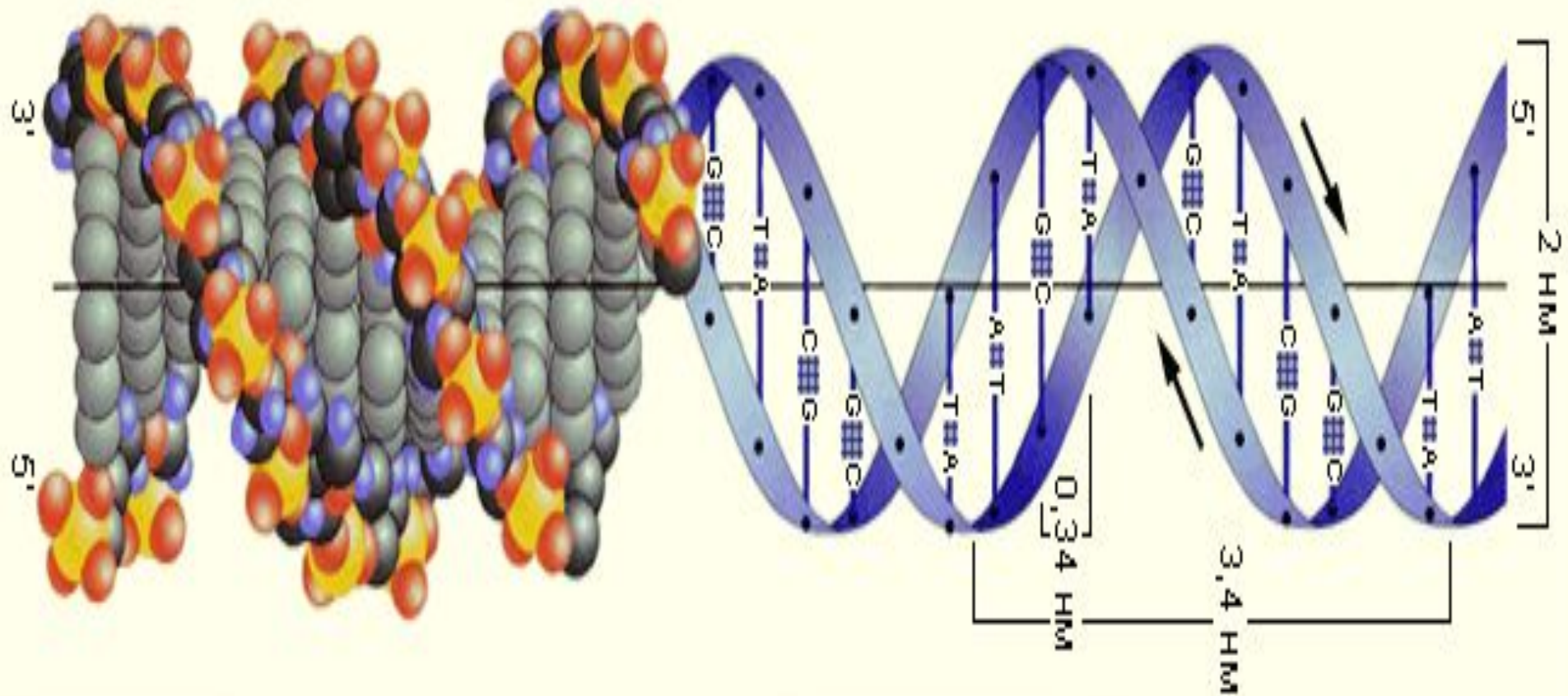
информация о признаках организмов

информация о том, сколько, каких аминокислот и в каком порядке соединятся при образовании молекулы белка

Хранение наследственной информации



Структура ДНК



Авторы пространственной модели ДНК



- **Уотсон (Watson)** Джеймс Дьюи (06.04.1928, Чикаго), американский биохимик, специалист в области молекулярной биологии



- **Крик (Crick) Фрэнсис Харри Комптон** (08.06.1916, Нортгемптон), английский биофизик

Ген

часть хромосомы, определяющая какой-то наследственный признак

ген – фрагмент генетического кода, который может синтезировать только один определённый полипептид

Ген → полипептид(белок) → признак

Генетический код

система записи
информации о
последовательности
аминокислот в
полипептиде
последовательностью
нуклеотидов ДНК или
РНК

Свойства генетического кода

Аминокислота	Кодирующие триплеты РНК (кодоны)
Аланин	ГЦУ ГЦЦ ГЦА ГЦГ
Аргинин	ЦГУ ЦГЦ ЦГА ЦГГ АГА
Аспарагин	АГГ
Аспарагиновая кислота	ААУ ААЦ
Валин	ГАУ ГАЦ
Гистидин	ГУУ ГУЦ ГУА ГУГ
Глицин	ЦАУ ЦАЦ
Глутамик	ГГУ ГГЦ ГГА ГГГ
Глутаминовая кислота	ЦАА ЦАГ
Изолейцин	ГАА ГАГ
Лейцин	АУУ АУЦ АУА
Лизин	ЦУУ ЦУЦ ЦУА ЦУГ
Метионин	УУА УУГ
Пролин	ААА ААГ
Сери	АУГ
Тирозин	ЦЦУ ЦЦЦ ЦЦА ЦЦГ
Треонин	УЦУ УЦЦ УЦА УЦГ
Триптофан	АГУ АГЦ
Фенилаланин	УАУ УАЦ
Цистеин	АЦУ АЦЦ АЦА АЦГ
Знаки препинания	УГГ УУУ УУЦ УГУ УГЦ УАА УАГ УГА

- Триплетность
- Однозначность
- Вырожденность
- Универсальность
- Неперекрываемость

Информация в информатике и биологии

информатика

$$N=2^l$$

$$256=2^8$$

256 символов

8 бит информации несет
1 символ (10100011)

Биология

триплет - 3 нуклеотида
кодирует 1
аминокислоту

$$N=2^l$$

$$64 = 4^3 = (2^2)^3 = 2^6$$

триплет несет 6 бит
информации, а
нуклеотид - 2

нуклеотид кодируется
сочетанием двух знаков
(1 и 0)

Какие информационные процессы осуществляются в живой природе?

- ◎ репликация (копирование родительской ДНК с образованием дочерних ДНК);
- ◎ транскрипция (переписывание генетической информации в форме РНК);
- ◎ трансляция (перевод информации с РНК на белковую форму).