**а) Расчет энтропии для алфавитов на латинице и кириллице**

1. **Энтропия как мера неопределенности** в информационных системах определяется по формуле Шеннона:

H(X)=−∑i=1np(xi)log⁡2p(xi)H(X) = - \sum\_{i=1}^{n} p(x\_i) \log\_2 p(x\_i)H(X)=−i=1∑n​p(xi​)log2​p(xi​)

где p(xi)p(x\_i)p(xi​) — вероятность появления символа xix\_ixi​, а nnn — количество различных символов в алфавите.

1. **Преобразование частоты символов в вероятности**: Для каждого алфавита (латиница и кириллица) вы должны посчитать частоту появления каждого символа в тексте. Например, если буква «а» встречается 50 раз из 1000 символов, её вероятность будет p(a)=501000=0.05p(a) = \frac{50}{1000} = 0.05p(a)=100050​=0.05.
2. **Вычисление энтропии для каждого алфавита**: После того как вы рассчитаете вероятности появления каждого символа, подставьте их в формулу Шеннона для вычисления энтропии.

**б) Определение энтропии бинарного алфавита**

Для бинарного алфавита (0 и 1) аналогичным образом считается энтропия. Нужно анализировать входной документ в виде бинарных данных и посчитать частоты появления нулей и единиц. После этого, как и в случае с текстами, переводим эти частоты в вероятности и вычисляем энтропию по формуле Шеннона.

**в) Подсчет количества информации в сообщении**

1. **Информация в сообщении**: Количество информации в сообщении можно вычислить через энтропию. Например, если ваше имя и фамилия содержат несколько символов, то можно подсчитать общее количество информации, используя полученную ранее энтропию алфавита.
2. **Использование кодов ASCII**: В случае с кодами ASCII, для каждого символа (например, буквы) в вашем имени и фамилии можно вычислить количество информации, зная, что в ASCII кодировании каждый символ занимает 1 байт (8 бит), и его вероятность может быть рассчитана аналогично.
3. **Получение результата**: Используйте энтропию для латиницы и кириллицы для текста, и для бинарного алфавита — для ASCII кодов, чтобы подсчитать информацию в сообщении (например, ваше имя). Для этого применяйте формулу:

I=H(X)×LI = H(X) \times LI=H(X)×Lгде H(X)— энтропия алфавита, а L — длина сообщения в символах.

Значение **242.2254 бит** объясняется тем, что имя конвертируется в строку числовых кодов, которая имеет гораздо большую длину, и к ней применяется расчет энтропии. Это не совсем типичный расчет информации для ASCII — скорее, он иллюстрирует объем данных при использовании битового представления кодов символов.

 **Количество информации для фамилии, имени и отчества:**

* Для **польского алфавита**: Рассчитано количество информации для имени "SasnowiecMariaIgorewna" с учетом энтропии польского алфавита (H=4.1763H = 4.1763H=4.1763). Полученный результат: I=H×NI = H \times NI=H×N, где NNN — длина имени (22 символа). Итог: **91.8786 бит**.
* Для **белорусского алфавита**: Имя "СаснавецМарыяІгараўна" использует белорусский алфавит (H=4.4958H = 4.4958H=4.4958). Аналогичный расчет дает: **94.4123 бит**.
* Для **ASCII**: Кодировка ASCII предполагает более высокий уровень детализации (8 бит на символ). Энтропия для ASCII (H=8H = 8H=8) при длине строки даёт результат: **242.2254 бит**.

**г) Подсчет информации с учетом ошибок передачи**

Если вероятность ошибки передачи бита сообщения составляет 0,1; 0,5; 1,0, то энтропия и количество информации в сообщении нужно будет скорректировать с учетом этих вероятностей ошибок. Для этого используется **передача с ошибками** и **коррекция ошибок**.

1. **Коррекция ошибки и расчет новой энтропии**: Для учета ошибок можно использовать теорию вероятности, включая вероятность ошибки при передаче. Этому может быть посвящен отдельный блок по теории кодирования, в котором учитываются параметры канала связи и способы восстановления информации.
   * Для каждого значения вероятности ошибки (0,1; 0,5; 1,0) пересчитывается информация, которая уже не будет полностью передана, из-за ошибок, и на основе этого делают выводы.

**1. Что такое алфавит источника сообщения?**

Алфавит источника сообщения – это конечная или бесконечная совокупность символов и знаков, формирующих сообщение.

**2. Что такое мощность алфавита источника сообщения?**

Мощность алфавита – это количество символов, составляющих алфавит.

**3. Какова мощность алфавита белорусского языка?**

32

**4. Какова мощность алфавита русского языка?**

33

**5. Какова мощность алфавита «компьютерного» языка?**

2

**6. Что такое энтропия алфавита?**

Энтропия алфавита – это информационная характеристика алфавита. Энтропия характеризует количество информации, приходящееся в среднем на один символ алфавита.

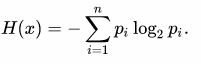
**7. Что такое энтропия сообщения?**

Энтропия алфавита – это информационная характеристика алфавита сообщения. Энтропия характеризует количество информации, приходящееся в среднем на один символ алфавита в сообщении.

**8. От чего зависит энтропия алфавита?**

Зависит только от распределения вероятностей появления символов, а не от алфавита

**9. Напишите формулу вычисления энтропии алфавита.**

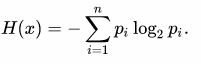


**10. Что нужно знать для вычисления энтропии алфавита?**

Вероятности появления каждого символа алфавита.

**11. Как рассчитываются энтропия Шеннона и энтропия Хартли? В чем принципиальное различие между этими характеристиками? Дайте толкование физического смысла энтропии?**

**По Шеннонну**



где pi – вероятность появления i-ого символа алфавита.

**по Хартли**

*,*

где N – мощность алфавита.

Энтропия по Хартли рассчитывается с учетом того, что алфавит состоит из символов с равными вероятностями появления. Для расчета энтропии по Шеннону необходимо вычислить отдельно вероятность появления каждого символа алфавита.

Физический смысл: показывает сколько в среднем информации приходится на один символ алфавита.

**12. Пояснить назначение знака «минус» в формулах (2.1) и (2.4).**

Энтропия не может быть отрицательной, а логарифм может быть отрицательным. Поэтому чтобы избежать отрицательную энтропию ставят –

**13. Что такое избыточность сообщений?**

Избыточность – термин, означающий превышение количества информации, используемой для передачи или хранения сообщения, над его информационной энтропией. Возникает, когда энтропия Шеннона и энтропия Хартли не равны. Неполно используется алфавит.

**14. Расположить в порядке возрастания энтропии известные Вам алфавиты.**

Английский

Белоруский

Русский

