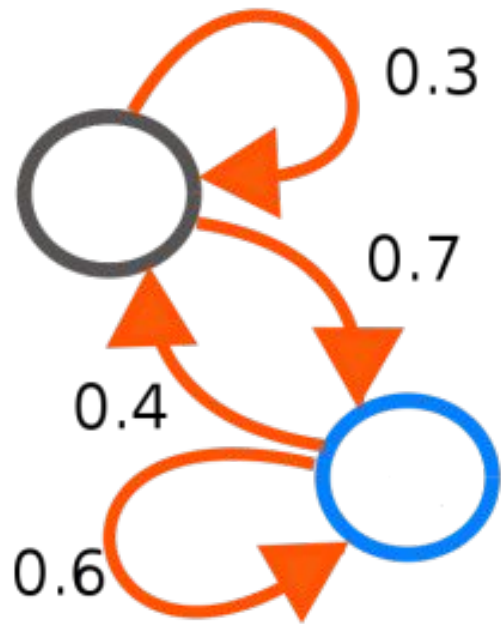


Цепи Маркова

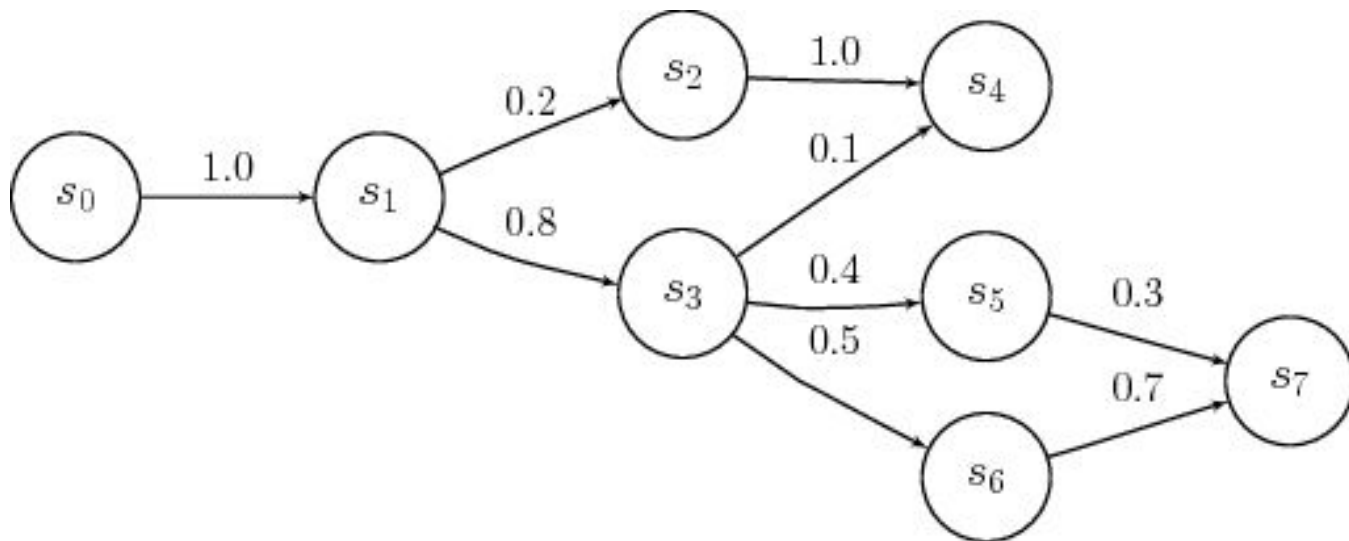
Марковский процесс

Марковский процесс - случайный процесс, поведение которого зависит только от текущего состояния, т.е. не зависит от прошлого



Примеры Марковских процессов

- 1) Случайные блуждания
- 2) Предсказания погоды
- 3) Оценка продаж



Цепи Маркова в генераторе текстов

One fish two fish red fish blue fish

One fish two fish red fish blue fish

One : 1

fish : 4

two : 1

red : 1

blue : 1

(*Start*, One)

(One, fish)

(fish, two) (fish, red) (fish, blue) (fish, *END*)

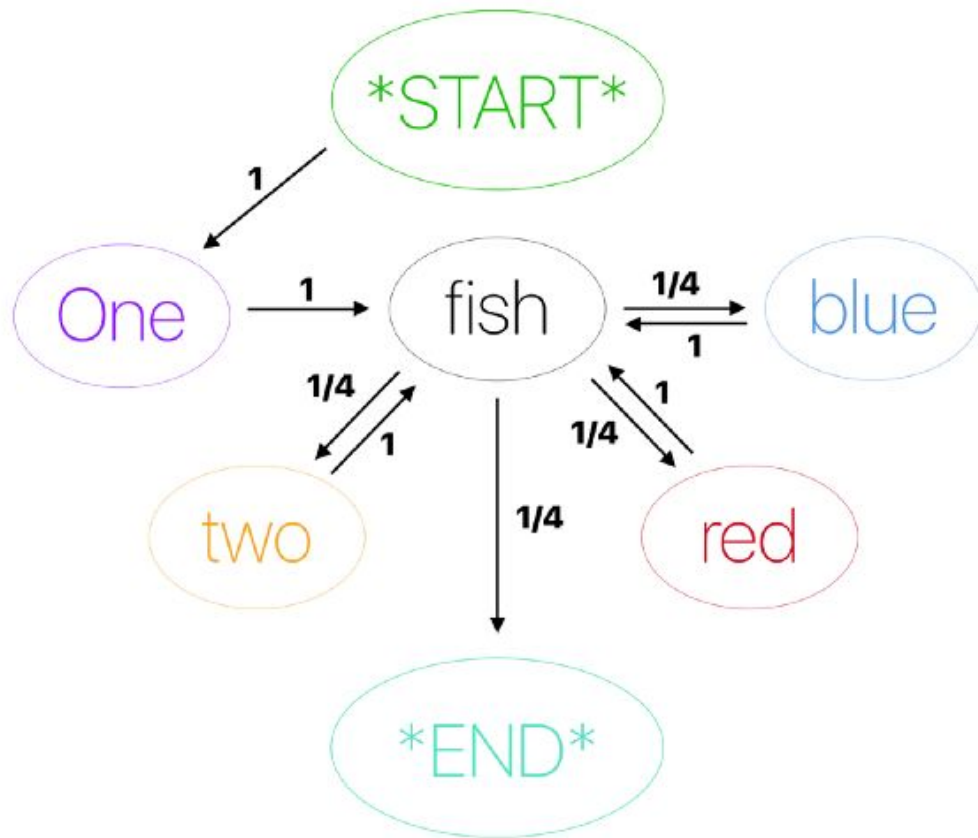
(two, fish)

(red, fish)

(blue, fish)

(*END*, none)

Start : [One]
One : [fish]
fish : [two, red, blue, *END*]
two : [fish]
red : [fish]
blue : [fish]
END : [none]



One fish two fish red fish blue fish

Алгоритм

Считывание данных

- 1) Считать текст и посчитать, сколько раз встречается каждая пара в тексте
- 2) Сохранить базу слов в файл, чтобы потом можно было загрузить готовую и впоследствии дополнять ее новыми словами

Вывод нового текста

- 1) Загрузить базу слов из файла
- 2) Отнормировать частоты по сумме для каждого слова, являющегося первым в паре. Это и будут вероятности
- 3) Выбрать первое слово, длину текста, длину абзаца и т.д.
- 4) Сгенерировать и вывести текст

Как можно улучшить качество текстов?

- 1) Добавить в базу слова, обозначающие начало и конец предложения и разделять текст на предложения и абзацы
- 2) Строить не 2-граммы, а 3-граммы, 4-граммы и т.д.
- 3) Использовать морфологию с помощью специальных библиотек (PyMorphy2, NLTK)

| | | |
|----------------|---|-------------|
| (*Start*, one) | : | [fish : 1] |
| (one, fish) | : | [two : 1] |
| (fish, two) | : | [fish : 1] |
| (two, fish) | : | [red : 1] |
| (fish, red) | : | [fish : 1] |
| (red, fish) | : | [blue : 1] |
| (fish, blue) | : | [fish : 1] |
| (blue, fish) | : | [*END* : 1] |

Полезные ссылки

<https://ru.wikipedia.org/wiki/%D0...>

<https://neerc.ifmo.ru/wiki/index.php?title=%D0...>

<https://tproger.ru/translations/markov-chains/>

<https://ru.wikipedia.org/wiki/N-%D0...>

<https://habr.com/ru/post/88514/>

<https://habr.com/ru/post/334046/>