

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

федеральное государственное автономное образовательное учреждение высшего образования

Санкт-Петербургский национальный исследовательский университет информационных
технологий, механики и оптики

Мегафакультет трансляционных информационных технологий

Факультет информационных технологий и программирования

Лабораторная работа № 4

По дисциплине «Введение в цифровую культуру и программирование»

Исправление ошибок

Выполнил студент группы М3101

Дудко Матвей Владимирович

Проверил:

Хлопотов Максим Валерьевич

САНКТ-ПЕТЕРБУРГ

2019

ЛАБОРАТОРНАЯ РАБОТА 4. ИСПРАВЛЕНИЕ ОШИБОК

В вашу тему на Piazza загружен текст.

На странице Resources загружен словарь (dict1.txt)

Ход работы

1. Предобработка текста

1.1 Текст нужно разделить на слова.

1.2 Удалить следующие знаки препинания: ! ? , . : « () » ‘

Не удаляйте из слов дефисы.

1.3 Перевести все буквы в строчные (маленькие). Например, "Средний" - заменить на "средний".

Приводить слова к нормальной форме не нужно, так как в словаре присутствуют различные словоформы.

Например, которая которого которое которой которому которую которые который которым которыми которых. Это всё разные словоформы. Всего в словаре 4773 разных словоформ, отсортированных по алфавиту.

2. Первичные расчёты

2.1 Посчитайте словоформы в своём тексте

2.2 Посчитайте разные словоформы

2.3 Посчитайте сколько разных словоформ из вашего текста присутствуют в словаре

Обратите внимание, что в словаре после слова через пробел написано число – это частота встречаемости во всём тексте.

3. Поиск и исправление ошибок

3.1 Посчитайте, сколько словоформ не присутствует в словаре ("потенциальные ошибки")

3.2 Найдите для каждого из них редакторское расстояние до ближайшего слова.

Исправление ошибок

Редакторское расстояние – это минимальное количество разрешённых операций, необходимых для превращения одной строки в другую. В настоящем задании разрешены следующие операции: вставка одного символа, удаление одного символа и замена одного символа на другой. Допустимо в строку вставить символ «пробел», превратив строку в две.

3.3 В настоящем задании, если редакторское расстояние равняется 1 или 2, то словоформа в вашем тексте признаётся ошибочной и её нужно заменить словоформу на соответствующую словоформу из словаря.

Если в словаре оказалось несколько словоформ с одинаковым редакторским расстоянием до ошибочной словоформы из текста, то нужно заменить на то, у которого частота выше.

4. После поиска и исправления ошибок повторите расчёты:

4.1 Посчитайте словоформы в своём тексте

4.2 Посчитайте разные словоформы

4.3 Посчитайте сколько разных словоформ из вашего текста присутствуют в словаре

5. Выведите все "потенциальные ошибки" в порядке встречаемости в тексте в следующем виде: словоформа из текста - словоформа из словаря - редакторское расстояние.

Если удалось исправить не все "потенциальные ошибки", то нужно вывести только неисправлённое слово из текста с пометкой "не найдено".

Например,

1) Сридний - средний - 1

2) гепаталамус - гипоталамус - 2

3) гепотоламмус - не найдено - >2

Ответы на вопросы нужно опубликовать в ответном сообщении на Piazza.

Также к ответам нужно прикрепить отчёт в формате pdf. Имя файла должно включать номер группы, ФИО и номер работы.

В отчёт включить.

1. Задание на лабораторную работу
2. Выданный вам текст
3. Код функции (метода) для подсчёта редакторского расстояния
4. Ответы на вопросы задания 2-5 и программный код, с помощью которого эти ответы были получены.
5. Текст с исправленными ошибками

Выданный текст: brain073.txt

Мозг и сенсорные системы

Наш организм довольно богато снабжен различными органами чувств. Еще в античные времена выделили основные пять чувств: зрение, слух, обоняние, осязание и вкус. На самом деле мы снабжены сенсорными системами гораздо богаче. Их назначение понятно: мы собираем информацию из внешней среды и из внутренней среды организма, потому что нашему мозгу важно, в каком состоянии находятся внутренние органы, насколько растянут кишечник или бронхи — все это достаточно значимо.

Большинство сенсорных систем имеют стандартное строение, и все начинается с клеток-рецепторов, то есть таких датчиков, которые реагируют на сигнал — на химический сигнал (молекулы появились в окружающей среде) или на физический, прикосновения, электромагнитные волны, как в случае зрения. Дальше этот датчик, клетка-рецептор, передает электрические импульсы на проводящий нерв. Нерв — это такой провод, который связывает датчик и центральный процессор, головной и спинной мозг. У нас, как известно, 31 пара спинно-мозговых нервов, и все они занимаются передачей сенсорных сигналов от разных этажей тела. Кроме того, из 12 пар черепных нервов большинство тоже занимаются сенсорикой. И наконец, третий, самый сложный этап: сигнал попадает в центральную нервную систему и дальше сначала внутри спинного мозга, а потом и головного мозга последовательно обрабатывается, запускаются те или иные реакции, информация запоминается. Чем выше сигналы двигаются по центральной нервной системе, тем более сложные вычислительные операции реализуются. Самые сложные человеческие моменты обработки информации случаются в коре больших полушарий.

Если смотреть подробнее на наши рецепторы, с них, собственно, все начинается. Мы видим, что они делятся на два типа: это могут быть нервные клетки или ненервные клетки. Если рецептор — это нейрон или его отросток, такие рецепторы называются первично чувствующими. В каком-то смысле эволюция начинала с них. На нервные клетки приходил сигнал, дальше генерировался электрический импульс, и в этом понятном для мозга виде информация поднималась в спинной мозг, головной мозг. Но сигналов очень много, и они разные. Видимо, ресурсов нейронов не хватает на то, чтобы реагировать на все на свете, а чем больше сенсорных потоков вы считываете, тем полней информация об окружающей среде, тем правильнее ваше поведение, поэтому эволюция искала еще какие-то датчики, кроме нейронов. В конце концов ряд клеток, прежде всего эпителиальных, покровных клеток на поверхности кожи или на поверхности полостей организма, тоже превратились в рецепторы. Но это уже не нервные клетки, а такие рецепторы называются вторично чувствующими. Для того чтобы им передать сигнал в ЦНС, нужна помочь нейронов периферической нервной системы. То есть рецептор реагирует на стимул, потом он должен

передать его на так называемый проводящий нейрон, и уже только отростки проводящего нейрона доберутся до головного и спинного мозга.

К первично чувствующим рецепторам относятся рецепторы нашей обонятельной системы, а еще рецепторы таких систем, как кожная, мышечная, болевая, еще рецепторы системы внутренней чувствительности. Вторично чувствующие рецепторы — это зрение, слух, вестибулярная система и вкус. Получается, что у нас девять больших серьезных сенсорных систем. Хотя на самом деле иногда предлагаю выделить больше. Критерий выделения некой части нашего тела в отдельную сенсорную систему в общем и целом довольно понятен. Мы говорим об особой сенсорной системе, если есть свои рецепторы, свои проводящие пути и свои отдельные центры в головном и спинном мозге, которые обмениваются внутри сенсорной системы информацией. С этой точки зрения кожная чувствительность, болевая чувствительность и мышечная чувствительность — это разные сенсорные системы, хотя когда-то говорилось об общей чувствительности тела. Обоняние — отдельная сенсорная система, но существует так называемая дополнительная обонятельная система — вомероназальный орган. Эта конструкция, хоть она и небольшая, удовлетворяет всем критериям, приложимым к сенсорной системе. Поэтому довольно логично вомероназальный орган и все, что с ним связано, то есть сигналы, которые возникают при появлении феромонов, а потом уходят в гипоталамус, выделять в отдельную сенсорную систему. Но она получается уж больно маленькая, просто она очень сильно редуцирована.

Как рецептор вообще реагирует на сигнал? За счет чего чувствительная клетка или ее отросток отвечает на физическое или химическое воздействие? Логика работы здесь довольно близка к тому, что вообще делают нейроны. Обычная нервная клетка отвечает на появление вещества медиатора. Рецепторы вкуса или рецепторы обоняния, рецепторы внутренней чувствительности примерно так же реагируют на появление химического вещества. На мембране рецептора есть чувствительные белки, с которыми связаны ионные каналы. При появлении определенного запаха они открываются, в клетку входят положительно заряженные ионы, возникает сдвиг заряда вверх, деполяризация, и это может служить причиной генерации электрических импульсов. Дальше эти импульсы убегут опять-таки в головной или спинной мозг. Примерно по такому же принципу работают рецепторы механической чувствительности и даже зрительные рецепторы. Как правило, некое адекватное сенсорное воздействие вызывает на мембране рецептора открывание (правда, иногда закрывание) тех или иных ионных каналов, возникает сдвиг заряда в клетке и генерируется потенциал действия, убегающий в центральную нервную систему. И чем сильнее сенсорное воздействие, тем чаще бегут импульсы (потенциалы действия) сначала по сенсорному нерву, а потом уже внутри сенсорных центров головного и спинного мозга.

Исправление ошибок

Это является первым из двух базовых законов работы сенсорных систем. Закон звучит так: интенсивность энергии сенсорного сигнала кодируется частотой потенциала действия в проводящем нерве. То есть чем громче звук, чем ярче свет, чем более концентрированный раствор, например, глюкозы, тем чаще бегут импульсы по тому или иному нерву. В зависимости от этой частоты наш головной мозг и высшие центры узнают об интенсивности сенсорного сигнала. Если говорить уже о реальных цифрах, то сигнал, который субъективно воспринимается как довольно слабый, где-то 20–40 импульсов в секунду. Если импульсы бегут с частотой 55–70 Гц по нерву, то это для нас субъективно сигнал средней силы. Когда ближе к 100 импульсам в секунду, то есть 100 Гц, это сильный сигнал. А когда уходит за 100 Гц, это уже сверхсильный сигнал, и такие сигналы для нас зачастую субъективно неприятны. Слишком яркий свет, слишком громкий звук — мы стараемся уйти от таких воздействий, потому что велик шанс повреждений тех самых рецепторов или, что еще хуже, сенсорных центров головного и спинного мозга.

Для того чтобы рецепторы хорошо и качественно работали, им, как правило, нужны некие вспомогательные структуры, создающие для них все условия. Рецепторы функционируют уже внутри этих структур. Такие структуры мы называем органами чувств. Не нужно путать понятия «орган чувств» и «сенсорная система». Орган чувств — это место, где рецептору хорошо. Скажем, глаз — это орган зрения. Внутреннее ухо или улитка — орган слуха. Кожа — это орган осязания, болевой чувствительности.

Помимо интенсивности, энергии каждый сенсорный сигнал характеризуется еще одним качеством. С точки зрения организации сенсорной системы качественно разными сигналами считаются те, которые действуют на разные рецепторы. Это не очень сходится с нашим бытовым восприятием работы органов чувств и сенсорных систем, но это именно так. Проще всего это понять на примере кожной чувствительности. У нас есть поверхность кожи, по которой рассеяны рецепторы, отростки нервных клеток, и разные рецепторы обслуживают разные участки кожи. Соответственно, есть рецептор и нейрон, работающий с большим пальцем, и есть нейрон, работающий с мизинцем. Качественно разные сигналы — это сигналы, которые считаются от разных участков кожи. Для слуховой системы организация нашей улитки такова, что разные рецепторы реагируют на сигналы разной тональности. Есть рецепторы, настроенные на высокие частоты, на низкие частоты, средние частоты. Для нашей зрительной системы качественно разными сигналами являются сигналы, приходящие от разных точек пространства, потому что разные фоторецепторы у нас на сетчатке как бы сканируют свой кусочек этой 2D-картинки и отчитываются перед центральной нервной системой о неких точках в определенных местах пространства. То есть качественно разные сигналы — это сигналы, действующие на разные рецепторы.

Исправление ошибок

Дело еще в том, что рецепторы, как правило, расположены в определенном месте нашего тела. Эта зона называется рецепторная поверхность. Каждый рецептор передает сигнал своим нервным клеткам, информация от соседних рецепторов передается соседним нервным клеткам. В итоге рецепторная поверхность параллельно отображается на структурах головного и спинного мозга — этот параллельный перенос знаком вам из геометрии. В результате возникает очень интересный эффект: у нас в головном или спинном мозге формируется карта рецепторных поверхностей. Наша кожная поверхность с большим пальцем, ухом, спиной, мизинцем, коленом и так далее отображается в центрах кожной чувствительности, сетчатка отображается в зрительных центрах, а улитка и ее базилярная мембрана — в слуховых центрах. Параллельный перенос позволяет нашему мозгу различать сигналы разного качества. Каким образом мозг узнает, что прикоснулись к носу или к колену? Ведь импульсы, которые бегут по нервным клеткам, абсолютно одинаковые. Узнать можно, только если посмотреть, по какому аксону прибежал сигнал. В кибернетике это называется кодировка номером канала. Принцип кодировки номером канала лежит в основе работы сенсорных систем. Это второй базовый закон работы сенсорных систем. Он звучит так: качество сенсорного сигнала кодируется номером канала. Мы можем закодировать интенсивность сигнала с помощью частоты ПД, закодировать качественные характеристики с помощью номера канала, и этого достаточно головному мозгу, для того чтобы дальше обрабатывать эту сенсорную информацию.

Что происходит в головном и спинном мозге с сенсорными сигналами? Они фильтруются и способны запускать различные реакции. Головной и спинной мозг, особенно головной, способны опознавать так называемые сенсорные образы. Сенсорная область — это совокупность нескольких сенсорных сигналов, информационная сущность более высокого порядка. Спинной мозг в основном работает с чувствительностью тела, 31 сегмент спинного мозга считывает информацию с 31-го этажа нашего тела: это болевая чувствительность, кожная, мышечная чувствительность и сигналы от внутренних органов — это называется интерорецепция, внутренняя чувствительность. Дальше белое вещество спинного мозга, скопление аксонов позволяет провести, передать эту информацию уже в головной мозг. Главными восходящими трактами спинного мозга, теми, которые передают такую сенсорную информацию, являются так называемые дорзальные столбы, идущие на самой задней поверхности спинного мозга. Еще есть спинно-мозжечковые тракты, взаимодействующие с мозжечком. Для передачи болевой чувствительности очень важен спинно-таламический тракт.

Если мы говорим о головном мозге, то ему достается львиная доля сенсорных входов. Существует обонятельный нерв, зрительный нерв, вестибуло-слуховой нерв — три нерва, занимающиеся исключительно сенсорикой. Кроме того, такие нервы, как лицевой, языкоглоточный, тройничный, тоже передают различные сенсорные сигналы.

Исправление ошибок

Очень важным уровнем обработки сенсорных сигналов является таламус — структура, через которую все сенсорные потоки, кроме обоняния, поднимаются в кору больших полушарий. Таламус — это важнейший информационный фильтр, работающий по заказу коры больших полушарий и пропускающий то, что здесь и сейчас является значимым. Кроме того, таламус очень охотно пропускает новые сильные сигналы. В выполнении этой функции ему помогает четверохолмие среднего мозга, где расположены наши древние зрительные и слуховые центры. В конце концов сенсорная информация поднимается в кору больших полушарий, где есть зриетльные центры, слуховые центры, вкусовые центры. Затылочная доля — это зрительная кора, височная доля — слуховая, область в районе центральной борозды — это наша чувствительность. Внутри этих сенсорных зон выделяют первичную, вторичную, а также третичную кору, которая занимается узнаванием все более сложных образов. Первичная зрительная кора — это узнавание линий, вторичная зрительная кора — узнавание геометрических фигур, а третичная — это уже лица конкретных людей. После обработки в конкретных сенсорных центрах сенсорная информация передается в ассоциативную теменную кору, где находятся нейроны, способные работать одновременно с разными сенсорными потоками.

Код функций для подсчёта редакторского расстояния

```
def distance(a, b):
    """Calculates the Levenshtein distance between a and b."""
    """Wagner-Fischer algorithm"""

    n, m = len(a), len(b)
    if n > m:
        # Make sure n <= m, to use O(min(n, m)) space
        a, b = b, a
        n, m = m, n

    current_row = range(n + 1) # Keep current and previous row, not entire matrix
    for i in range(1, m + 1):
        previous_row, current_row = current_row, [i] + [0] * n
        for j in range(1, n + 1):
            add, delete, change = previous_row[j] + 1, current_row[j - 1] + 1, previous_row[j - 1]
            if a[j - 1] != b[i - 1]:
                change += 1
            current_row[j] = min(add, delete, change)

    return current_row[n]
```

Исправление ошибок

```
def edit_mistaken_word(mistaken_word):
    minimum_distance = len(mistaken_word)
    probable_word = mistaken_word
    probable_word_freq = -1
    for word in dict_words:
        calc_distance = distance(word, mistaken_word)
        if calc_distance < minimum_distance:
            probable_word = word
            probable_word_freq = dictionary[dict_words.index(word)][1]
            minimum_distance = calc_distance
        elif calc_distance == minimum_distance:
            if probable_word_freq < dictionary[dict_words.index(word)][1]:
                probable_word = word
                probable_word_freq = dictionary[dict_words.index(word)][1]

    if minimum_distance <= 1:
        return probable_word, minimum_distance

    # Try to split into two words
    # Note: могут быть опечатки внутри двух слов помимо пробелов
    for place in range(len(mistaken_word) + 1):
        word1, word2 = mistaken_word[:place], mistaken_word[place:]
        if word1 in dict_words and word2 in dict_words:
            return word1 + ' ' + word2, 1

        elif word1 in dict_words:
            minimum_distance2 = len(word2)
            probable_word2 = word2
            probable_word_freq2 = -1
            for word in dict_words:
                calc_distance = distance(word, mistaken_word)
                if calc_distance < minimum_distance2:
                    probable_word2 = word
                    probable_word_freq2 = dictionary[dict_words.index(word)][1]
                    minimum_distance2 = calc_distance
                elif calc_distance == minimum_distance2:
                    if probable_word_freq2 < dictionary[dict_words.index(word)][1]:
                        probable_word2 = word
                        probable_word_freq2 = dictionary[dict_words.index(word)][1]
            if minimum_distance2 <= 1:
                return word1 + ' ' + probable_word2, 2
        elif word2 in dict_words:
            minimum_distance1 = len(word1)
            probable_word1 = word1
            probable_word_freq1 = -1
            for word in dict_words:
                calc_distance = distance(word, mistaken_word)
                if calc_distance < minimum_distance1:
                    probable_word1 = word
                    probable_word_freq1 = dictionary[dict_words.index(word)][1]
                    minimum_distance1 = calc_distance
                elif calc_distance == minimum_distance1:
                    if probable_word_freq1 < dictionary[dict_words.index(word)][1]:
                        probable_word1 = word
                        probable_word_freq1 = dictionary[dict_words.index(word)][1]
            if minimum_distance1 <= 1:
                return probable_word1 + ' ' + word2, 2

    return probable_word, minimum_distance
```

Ответы на вопросы:

Количество словоформ: 1725

Количество различных словоформ: 865

Количество различных словоформ присутствующих в словаре: 858

Не найдено словоформ: 7

Ошибки:

- 1) клеток- - клеток - 1
- 2) химический - химический - 1
- 3) эектрические - электрические - 1
- 4) спинно-мозговых - спинномозговых - 1
- 5) сингалов - сигналов - 2
- 6) 55–70 - 50–70 - 1
- 7) зриетльные - зрительные – 2

Расчеты после исправления опечаток:

Количество словоформ: 1725

Количество различных словоформ: 862

Количество различных словоформ присутствующих в словаре: 862

Исправленный текст

Мозг и сенсорные системы

Наш организм довольно богато снабжен различными органами чувств. Еще в античные времена выделили основные пять чувств: зрение, слух, обоняние, осязание и вкус. На самом деле мы снабжены сенсорными системами гораздо богаче. Их назначение понятно: мы собираем информацию из внешней среды и из внутренней среды организма, потому что нашему мозгу важно, в каком состоянии находятся внутренние органы, насколько растянут кишечник или бронхи — все это достаточно значимо.

Большинство сенсорных систем имеют стандартное строение, и все начинается с клеток рецепторов, то есть таких датчиков, которые реагируют на сигнал — на химический сигнал (молекулы появились в окружающей среде) или на физический, прикосновения, электромагнитные волны, как в случае зрения. Дальше этот датчик, клетка-рецептор, передает электрические импульсы на проводящий нерв. Нерв — это такой провод, который связывает датчик и центральный процессор, головной и спинной мозг. У нас, как известно, 31 пара спинномозговых нервов, и все они занимаются передачей сенсорных сигналов от разных этажей тела. Кроме того, из 12 пар черепных нервов большинство тоже занимаются сенсорикой. И наконец, третий, самый сложный этап: сигнал попадает в центральную нервную систему и дальше сначала внутри спинного мозга, а потом и головного мозга последовательно обрабатывается, запускаются те или иные реакции, информация запоминается. Чем выше сигналы двигаются по центральной нервной системе, тем более сложные вычислительные операции реализуются. Самые сложные человеческие моменты обработки информации случаются в коре больших полушарий.

Если смотреть подробнее на наши рецепторы, с них, собственно, все начинается. Мы видим, что они делятся на два типа: это могут быть нервные клетки или ненервные клетки. Если рецептор — это нейрон или его отросток, такие рецепторы называются первично чувствующими. В каком-то смысле эволюция начинала с них. На нервные клетки приходил сигнал, дальше генерировался электрический импульс, и в этом понятном для мозга виде информация поднималась в спинной мозг, головной мозг. Но сигналов очень много, и они разные. Видимо, ресурсов нейронов не хватает на то, чтобы реагировать на все на свете, а чем больше сенсорных потоков вы считываете, тем полней информация об окружающей среде, тем правильнее ваше поведение, поэтому эволюция искала еще какие-то датчики, кроме нейронов. В конце концов ряд клеток, прежде всего эпителиальных, покровных клеток на поверхности кожи или на поверхности полостей организма, тоже превратились в рецепторы. Но это уже не нервные клетки, а такие рецепторы называются вторично чувствующими. Для того чтобы им передать сигнал в ЦНС, нужна помочь нейронов периферической нервной системы. То есть рецептор реагирует на стимул, потом он должен

передать его на так называемый проводящий нейрон, и уже только отростки проводящего нейрона доберутся до головного и спинного мозга.

К первично чувствующим рецепторам относятся рецепторы нашей обонятельной системы, а еще рецепторы таких систем, как кожная, мышечная, болевая, еще рецепторы системы внутренней чувствительности. Вторично чувствующие рецепторы — это зрение, слух, вестибулярная система и вкус. Получается, что у нас девять больших серьезных сенсорных систем. Хотя на самом деле иногда предлагаю выделить больше. Критерий выделения некой части нашего тела в отдельную сенсорную систему в общем и целом довольно понятен. Мы говорим об особой сенсорной системе, если есть свои рецепторы, свои проводящие пути и свои отдельные центры в головном и спинном мозге, которые обмениваются внутри сенсорной системы информацией. С этой точки зрения кожная чувствительность, болевая чувствительность и мышечная чувствительность — это разные сенсорные системы, хотя когда-то говорилось об общей чувствительности тела. Обоняние — отдельная сенсорная система, но существует так называемая дополнительная обонятельная система — вомероназальный орган. Эта конструкция, хоть она и небольшая, удовлетворяет всем критериям, приложимым к сенсорной системе. Поэтому довольно логично вомероназальный орган и все, что с ним связано, то есть сигналы, которые возникают при появлении феромонов, а потом уходят в гипоталамус, выделять в отдельную сенсорную систему. Но она получается уж больно маленькая, просто она очень сильно редуцирована.

Как рецептор вообще реагирует на сигнал? За счет чего чувствительная клетка или ее отросток отвечает на физическое или химическое воздействие? Логика работы здесь довольно близка к тому, что вообще делают нейроны. Обычная нервная клетка отвечает на появление вещества медиатора. Рецепторы вкуса или рецепторы обоняния, рецепторы внутренней чувствительности примерно так же реагируют на появление химического вещества. На мембране рецептора есть чувствительные белки, с которыми связаны ионные каналы. При появлении определенного запаха они открываются, в клетку входят положительно заряженные ионы, возникает сдвиг заряда вверх, деполяризация, и это может служить причиной генерации электрических импульсов. Дальше эти импульсы убегут опять-таки в головной или спинной мозг. Примерно по такому же принципу работают рецепторы механической чувствительности и даже зрительные рецепторы. Как правило, некое адекватное сенсорное воздействие вызывает на мембране рецептора открывание (правда, иногда закрывание) тех или иных ионных каналов, возникает сдвиг заряда в клетке и генерируется потенциал действия, убегающий в центральную нервную систему. И чем сильнее сенсорное воздействие, тем чаще бегут импульсы (потенциалы действия) сначала по сенсорному нерву, а потом уже внутри сенсорных центров головного и спинного мозга.

Это является первым из двух базовых законов работы сенсорных систем. Закон звучит так: интенсивность энергии сенсорного сигнала кодируется частотой потенциала действия в проводящем нерве. То есть чем громче звук, чем ярче свет, чем более концентрированный раствор, например, глюкозы, тем чаще бегут импульсы по тому или иному нерву. В зависимости от этой частоты наш головной мозг и высшие центры узнают об интенсивности сенсорного сигнала. Если говорить уже о реальных цифрах, то сигнал, который субъективно воспринимается как довольно слабый, где-то 20–40 импульсов в секунду. Если импульсы бегут с частотой 50–70 Гц по нерву, то это для нас субъективно сигнал средней силы. Когда ближе к 100 импульсам в секунду, то есть 100 Гц, это сильный сигнал. А когда уходит за 100 Гц, это уже сверхсильный сигнал, и такие сигналы для нас зачастую субъективно неприятны. Слишком яркий свет, слишком громкий звук — мы стараемся уйти от таких воздействий, потому что велик шанс повреждений тех самых рецепторов или, что еще хуже, сенсорных центров головного и спинного мозга.

Для того чтобы рецепторы хорошо и качественно работали, им, как правило, нужны некие вспомогательные структуры, создающие для них все условия. Рецепторы функционируют уже внутри этих структур. Такие структуры мы называем органами чувств. Не нужно путать понятия «орган чувств» и «сенсорная система». Орган чувств — это место, где рецептору хорошо. Скажем, глаз — это орган зрения. Внутреннее ухо или улитка — орган слуха. Кожа — это орган осознания, болевой чувствительности.

Помимо интенсивности, энергии каждый сенсорный сигнал характеризуется еще одним качеством. С точки зрения организации сенсорной системы качественно разными сигналами считаются те, которые действуют на разные рецепторы. Это не очень сходится с нашим бытовым восприятием работы органов чувств и сенсорных систем, но это именно так. Проще всего это понять на примере кожной чувствительности. У нас есть поверхность кожи, по которой рассеяны рецепторы, отростки нервных клеток, и разные рецепторы обслуживаю разные участки кожи. Соответственно, есть рецептор и нейрон, работающий с большим пальцем, и есть нейрон, работающий с мизинцем. Качественно разные сигналы — это сигналы, которые считаются от разных участков кожи. Для слуховой системы организация нашей улитки такова, что разные рецепторы реагируют на сигналы разной тональности. Есть рецепторы, настроенные на высокие частоты, на низкие частоты, средние частоты. Для нашей зрительной системы качественно разными сигналами являются сигналы, приходящие от разных точек пространства, потому что разные фоторецепторы у нас на сетчатке как бы сканируют свой кусочек этой 2D-картинки и отчитываются перед центральной нервной системой о неких точках в определенных местах пространства. То есть качественно разные сигналы — это сигналы, действующие на разные рецепторы.

Дело еще в том, что рецепторы, как правило, расположены в определенном месте нашего тела. Эта зона называется рецепторная поверхность. Каждый рецептор передает сигнал своим нервным клеткам, информация от соседних рецепторов передается соседним нервным клеткам. В итоге рецепторная поверхность параллельно отображается на структурах головного и спинного мозга — этот параллельный перенос знаком вам из геометрии. В результате возникает очень интересный эффект: у нас в головном или спинном мозге формируется карта рецепторных поверхностей. Наша кожная поверхность с большим пальцем, ухом, спиной, мизинцем, коленом и так далее отображается в центрах кожной чувствительности, сетчатка отображается в зрительных центрах, а улитка и ее базилярная мембрана — в слуховых центрах. Параллельный перенос позволяет нашему мозгу различать сигналы разного качества. Каким образом мозг узнает, что прикоснулись к носу или к колену? Ведь импульсы, которые бегут по нервным клеткам, абсолютно одинаковые. Узнать можно, только если посмотреть, по какому аксону прибежал сигнал. В кибернетике это называется кодировка номером канала. Принцип кодировки номером канала лежит в основе работы сенсорных систем. Это второй базовый закон работы сенсорных систем. Он звучит так: качество сенсорного сигнала кодируется номером канала. Мы можем закодировать интенсивность сигнала с помощью частоты ПД, закодировать качественные характеристики с помощью номера канала, и этого достаточно головному мозгу, для того чтобы дальше обрабатывать эту сенсорную информацию.

Что происходит в головном и спинном мозге с сенсорными сигналами? Они фильтруются и способны запускать различные реакции. Головной и спинной мозг, особенно головной, способны опознавать так называемые сенсорные образы. Сенсорная область — это совокупность нескольких сенсорных сигналов, информационная сущность более высокого порядка. Спинной мозг в основном работает с чувствительностью тела, 31 сегмент спинного мозга считывает информацию с 31-го этажа нашего тела: это болевая чувствительность, кожная, мышечная чувствительность и сигналы от внутренних органов — это называется интерорецепция, внутренняя чувствительность. Дальше белое вещество спинного мозга, скопление аксонов позволяет провести, передать эту информацию уже в головной мозг. Главными восходящими трактами спинного мозга, теми, которые передают такую сенсорную информацию, являются так называемые дорзальные столбы, идущие на самой задней поверхности спинного мозга. Еще есть спинно-мозжечковые тракты, взаимодействующие с мозжечком. Для передачи болевой чувствительности очень важен спинно-таламический тракт.

Если мы говорим о головном мозге, то ему достается львиная доля сенсорных входов. Существует обонятельный нерв, зрительный нерв, вестибуло-слуховой нерв — три нерва, занимающиеся исключительно сенсорикой. Кроме того, такие нервы, как лицевой, языкоглоточный, тройничный, тоже передают различные сенсорные сигналы.

Исправление ошибок

Очень важным уровнем обработки сенсорных сигналов является таламус — структура, через которую все сенсорные потоки, кроме обоняния, поднимаются в кору больших полушарий. Таламус — это важнейший информационный фильтр, работающий по заказу коры больших полушарий и пропускающий то, что здесь и сейчас является значимым. Кроме того, таламус очень охотно пропускает новые сильные сигналы. В выполнении этой функции ему помогает четверохолмие среднего мозга, где расположены наши древние зрительные и слуховые центры. В конце концов сенсорная информация поднимается в кору больших полушарий, где есть зрительные центры, слуховые центры, вкусовые центры. Затылочная доля — это зрительная кора, височная доля — слуховая, область в районе центральной борозды — это наша чувствительность. Внутри этих сенсорных зон выделяют первичную, вторичную, а также третичную кору, которая занимается узнаванием все более сложных образов. Первичная зрительная кора — это узнавание линий, вторичная зрительная кора — узнавание геометрических фигур, а третичная — это уже лица конкретных людей. После обработки в конкретных сенсорных центрах сенсорная информация передается в ассоциативную теменную кору, где находятся нейроны, способные работать одновременно с разными сенсорными потоками.

Программа, при помощи которой производились вычисления и исправление опечаток

```
import re

# -----
# ----- ONE: Pre-processing --
# -----


dictionary = open("dict1.txt", 'r')
dictionary = [(i[0], int(i[1])) for i in list(map(lambda x: x.split(),
dictionary.readlines()))]
dict_words = [i for i, j in dictionary]

# -----
# Step 1: read whole file in a string
f = open("brain073.txt", "r")
s = ''.join(f.readlines())
source = s

# -----
# Step 2: replace specific symbols
for i in """!?,;.:«()»'":
    s = s.replace(i, ' ')

s = s.replace('\n', ' ')

# -----
# Step 2.1: split by whitespaces
# Note: Sequence of whitespaces doesn't
#       split into empty strings

all_words = s.split()
unique_words = set(all_words)

# -----
# Step 3: set all words to lowercase

all_words = list(map(lambda x: x.lower(), all_words))
unique_words = list(map(lambda x: x.lower(), unique_words))

# -----
# ----- TWO: First counting --
# -----


print('Количество словоформ:', len(all_words))
print('Количество различных словоформ:', len(unique_words))

count = 0
for i in unique_words:
    if i in dict_words:
        count += 1

print('Количество различных словоформ присутствующих в словаре:', count)

print('-' * (len('Количество различных словоформ присутствующих в словаре: ' + str(count))))
```

Исправление ошибок

```
# ----- THREE: Search and correct mistakes -----
# -----  
count = 0  
mistakes = []  
for i in all_words:  
    if i not in dict_words and i not in mistakes:  
        count += 1  
        mistakes.append(i)  
print('Не найдено словоформ:', count)  
  
def distance(a, b):  
    """Calculates the Levenshtein distance between a and b."""  
    """Wagner-Fischer algorithm"""  
  
    n, m = len(a), len(b)  
    if n > m:  
        a, b = b, a  
        n, m = m, n  
  
    current_row = range(n + 1)  
    for i in range(1, m + 1):  
        previous_row, current_row = current_row, [i] + [0] * n  
        for j in range(1, n + 1):  
            add, delete, change = previous_row[j] + 1, current_row[j - 1] + 1,  
previous_row[j - 1]  
            if a[j - 1] != b[i - 1]:  
                change += 1  
            current_row[j] = min(add, delete, change)  
  
    return current_row[n]  
  
def edit_mistaken_word(mistaken_word):  
    minimum_distance = len(mistaken_word)  
    probable_word = mistaken_word  
    probable_word_freq = -1  
    for word in dict_words:  
        calc_distance = distance(word, mistaken_word)  
        if calc_distance < minimum_distance:  
            probable_word = word  
            probable_word_freq = dictionary[dict_words.index(word)][1]  
            minimum_distance = calc_distance  
        elif calc_distance == minimum_distance:  
            if probable_word_freq < dictionary[dict_words.index(word)][1]:  
                probable_word = word  
                probable_word_freq = dictionary[dict_words.index(word)][1]  
  
    if minimum_distance <= 1:  
        return probable_word, minimum_distance  
  
    # Try to split into two words  
  
    # Note: могут быть опечатки внутри двух слов помимо пробелов  
    for place in range(len(mistaken_word) + 1):  
        word1, word2 = mistaken_word[:place], mistaken_word[place:]  
        if word1 in dict_words and word2 in dict_words:  
            return word1 + ' ' + word2, 1  
  
    elif word1 in dict_words:
```

Исправление ошибок

```
minimum_distance2 = len(word2)
probable_word2 = word2
probable_word_freq2 = -1
for word in dict_words:
    calc_distance = distance(word, mistaken_word)
    if calc_distance < minimum_distance2:
        probable_word2 = word
        probable_word_freq2 = dictionary[dict_words.index(word)][1]
        minimum_distance2 = calc_distance
    elif calc_distance == minimum_distance2:
        if probable_word_freq2 < dictionary[dict_words.index(word)][1]:
            probable_word2 = word
            probable_word_freq2 = dictionary[dict_words.index(word)][1]
if minimum_distance2 <= 1:
    return word1 + ' ' + probable_word2, 2

elif word2 in dict_words:

    minimum_distance1 = len(word1)
    probable_word1 = word1
    probable_word_freq1 = -1
    for word in dict_words:
        calc_distance = distance(word, mistaken_word)
        if calc_distance < minimum_distance1:
            probable_word1 = word
            probable_word_freq1 = dictionary[dict_words.index(word)][1]
            minimum_distance1 = calc_distance
        elif calc_distance == minimum_distance1:
            if probable_word_freq1 < dictionary[dict_words.index(word)][1]:
                probable_word1 = word
                probable_word_freq1 = dictionary[dict_words.index(word)][1]
    if minimum_distance1 <= 1:
        return probable_word1 + ' ' + word2, 2

return probable_word, minimum_distance

print('\nОшибки:')
correct_mistakes = [[i, i] for i in mistakes]
for i in range(len(mistakes)):
    print("{}:{} - {}".format(i + 1, mistakes[i]), end=' ')
    probable_word, minimum_distance = edit_mistaken_word(mistakes[i])
    if minimum_distance <= 2:
        print("{}:{} - {}".format(probable_word, minimum_distance))
        correct_mistakes[i][1] = probable_word
    else:
        print("не найдено - >2")

print("\n" + '-' * (len('Количество различных словоформ присутствующих в словаре: ') +
str(count)))

# Correcting mistakes

for i in correct_mistakes:
    regex = re.compile(i[0], re.IGNORECASE)
    source = regex.sub(i[1], source)
```

Исправление ошибок

```
# ----- FOUR: Second counting -----
# ----- FIVE: Output -----
# -----



s = source

for i in """!?,;.:«()»""":
    s = s.replace(i, ' ')

s = s.replace('\n', ' ')

all_words = s.split()
unique_words = set(all_words)

all_words = list(map(lambda x: x.lower(), all_words))
unique_words = list(map(lambda x: x.lower(), unique_words))

print('Количество словоформ:', len(all_words))
print('Количество различных словоформ:', len(unique_words))

count = 0
for i in unique_words:
    if i in dict_words:
        count += 1

print('Количество различных словоформ присутствующих в словаре:', count)
print('-' * (len('Количество различных словоформ присутствующих в словаре: ' + str(count))))


# ----- FIVE: Output -----
# -----



print("\n\nИсправленный текст:\n")
print(source)
```