

Возможно, неправильно

О науке, математике и вычислениях

Среди нас: головоломка с азбукой Морзе

Опубликовано [21 ноября 2020 г.](#)

В онлайн-игре [Among Us](#) игроки, посещающие комнату связи, слышат нечеткую аудиозапись серии высокочастотных сигналов, которые звучат как азбука Морзе. Я впервые услышал запись [здесь](#), но [это более позднее видео](#) также воспроизводит ее около 5:00, а затем следует хорошее объяснение проблемы с попыткой расшифровать код.

На следующем рисунке показана спектрограмма аудиоклипа, где время отложено по оси x , а каждый вертикальный срез показывает преобразование Фурье короткого (примерно 50 мс) скользящего окна сигнала, centered в соответствующее время. Мы можем ясно видеть «точки» и «тире» на частоте около 1 кГц с соответствующим переводом, наложенным желтым цветом.

Если мы проигнорируем цифры на самом нижнем уровне дерева, мы увидим, что буквы кода Морзе не только неоднозначны (т. е. не **свободны от префиксов**), но и почти «**максимально** неоднозначны» в том смысле, что дерево букв почти **полное**. То есть, для почти любого префикса из четырех точек и тире, с которым мы можем столкнуться, пробел, указывающий на конец первой буквы, может быть после *любого* из этих первых четырех символов.

Это было бы хорошим упражнением по программированию для студентов, чтобы показать, что эта конкретная последовательность из 24 символов может быть декодирована в последовательность букв ровно 3 457 592 возможными способами. Конечно, большинство этих декодирований приводят к бессмыслице, например, АЕАВКГЕАЕАЕЕЕ. Но более интересная и сложная проблема заключается в эффективном поиске *разумных* декодирований, то есть сообщений, состоящих из реальных (английских?) слов, возможно, дополнительно ограниченных грамматическими связями *между* словами.

Конечно, также возможно — вероятно? — что этот аудиоклип просто выдуман, случайная последовательность точек и тире, призванная *звучать* как «настоящая» азбука Морзе. И даже если это не так, *мы можем не заметить разницы*. Какой интересный вопрос побудил нас написать этот пост: если мы сгенерируем совершенно случайную и, таким образом, *намеренно* непонятную последовательность из 24 точек и тире, какова вероятность того, что она все равно даст «разумное» возможное декодирование для достаточно больших значений «разумного»?

Эта запись была размещена в **Uncategorized** пользователем **perhapswrong**. Добавьте в закладки **постоянную ссылку** [<https://possiblywrong.wordpress.com/2020/11/21/among-us-morse-code-puzzle/>].

6 МЫСЛЕЙ О « СРЕДИ НАС: ГОЛОВОЛОМКА С АЗБУКОЙ МОРЗЕ »



Крис Уэллонс

на **21 ноября 2020 г. в 14:13** сказал:

эффективно искать разумные расшифровки

Если вы забыли, это был [r/DailyProgrammer Challenge #380](#) в августе 2019 года. 😊

Есть старый трюк с использованием общего компрессора (DEFLATE и т. д.) для определения языка. Начните с корпуса для каждого языка-кандидата, который вы хотите определить. Добавьте неопознанный текст к каждому и сожмите все целиком. Тот, который сжимает меньше всего, является вероятным языком идентифицированного текста, поскольку он наиболее похож. Интересно, можно ли использовать эту технику для автоматического определения разумных кандидатов на декодирование для достаточно больших последовательностей кода Морзе.

**возможнонеправильно**на **21 ноября 2020 г. в 17:21** сказал:

Я помню, что видел эту задачу — по сути, это та же задача, что и сложная (пятничная) [версия](#) .

Несколько хороших и плохих новостей в продолжение: плохая новость в том, что я узнал, что эти звуки типичны для «маркеров каналов» или идентификаторов станций с азбукой Морзе, которые обычно можно услышать на полицейских сканерах... что в точности соответствует контексту окружающего звука, описанного парнем в изначальном видео по ссылке. Смотрите [здесь](#) пример.

(Потенциально) хорошая новость заключается в том, что эти маркеры каналов, похоже, являются «настоящей» азбукой Морзе с требуемыми пробелами между буквами/цифрами. Например, приведенный выше клип — это однозначно TAB269.

Я говорю «хорошие» новости, потому что код в игре, в котором отсутствуют пробелы, необходимые для его фактической интерпретации, вряд ли был взят из реального/стандартного звука сканера. То есть, я предполагаю, что разработчики «выдумали его»... и в процессе испортили его, оставив пробелы. Это предполагает — по крайней мере, для моего чрезмерно оптимистичного мышления решения головоломок :) — что они *намеревались* вставить что-то свое в этот звук.

**Крис Уэллонс**на **22 ноября 2020 г. в 22:10** сказал:

Ваша иллюстрация trie и комментарий о том, что она почти максимально неоднозначна (т. е. trie имеет мало «дырок», когда представлено с помощью массива) дали мне идею для декодирующего автомата. Trie здесь закодирован как байтовый массив:

<https://gist.github.com/skeeto/f9e198b913b228f3fd773a0c4e266579>

Подставив это в небольшую рекурсивную функцию для подсчета всех декодирований, получим точное совпадение с вашим подсчетом (что неудивительно).

**возможнонеправильно**на **23 ноября 2020 г. в 23:20** сказал:

Здорово! Мы могли бы пойти дальше и искать реальные слова, с большим trie, построенным из словаря, где каждый путь от корня соответствует слову, а не просто букве (по крайней мере, те пути, которые заканчиваются конечной вершиной «это завершает полное допустимое слово»; в trie было бы больше внутренних «дыр»). Это было бы намного эффективнее, чем мой ленивый подход Mathematica к рекурсивному сопоставлению шаблонов регулярных выражений. Результаты не предлагают ничего сногсшибательного: RATSTAMEFREE и тому подобное (хотя интерпретация кода *обратно*, кажется, дает более разумные варианты, VIDEOLATER, VIDEOREWIN(d?), IRENEJRITTER (пасхальное яйцо разработчика?) и т. д.).



Филип Моравец

на **31 января 2021 г. в 5:29 утра** сказал:

Вы пробовали смотреть только на те решения, которые состоят из наиболее часто используемых английских слов? Например, «the».

**возможнонеправильно**на **31 января 2021 г. в 13:05** сказал:

В некотором роде — я использовал полный список слов, но отсортировал результаты по частоте встречаемости в наборе данных Google Ngrams, чтобы более «разумные» расшифровки появлялись в начале списка.

Этот сайт использует Akismet для уменьшения спама. [Узнайте, как обрабатываются данные ваших комментариев.](#)