Generative Flow Networks, построение модели

Имя входного файла:	`стандартный ввод`
Имя выходного файла:	`стандартный вывод`
Ограничение по времени:	2 секунды
Ограничение по памяти:	512 мегабайт

Введение

Эта задача посвящена моделям **Generative Flow Networks (GFlowNets)**, которые являются одним из методов генерации дискретных объектов.

Формально задача генерации дискретных объектов формулируется так: у нас есть конечное множество объектов $\mathcal{X}$ и неотрицательная reward функция $\mathcal{R}(x)$ определенная для всех $x \in \mathcal{X}$. Задача состоит в том, чтобы построить и обучить модель, которая генерирует объекты из $\mathcal{X}$ из вероятностного распределения $\mathcal{P}(x) = \frac{1}{Z}\mathcal{R}(x)$ для $Z = \sum\limits_{x \in \mathcal{X}} \mathcal{R}(x)$.

Генеративный процесс с помощью GFlowNet может быть рассмотрен как последовательность действий, при которой мы стартуем из некоторого пустого объекта и на каждом шаге добавляем некоторый компонент в него.

Определим процесс формально.

- Рассмотрим ориентированный ациклический граф (DAG) $G = (\mathcal{S}, \mathcal{E})$, где S это множество состояний и $\mathcal{E} \subseteq \mathcal{S} \times \mathcal{S}$ это множество возможных переходов от одного состояния к следующему.

- В множестве состояний $\mathcal{S}$ есть ровно одна вершина исток s_0 (стартовое состояние), в которую не входит ни одного ребра. Гарантируется, что все вершины графа G достижимы из s_0 по ребрам.

- В множестве состояний $\mathcal{S}$ содержится множество финальных объектов $\mathcal{X} \subseteq \mathcal{S}$, причем это множество в точности совпадает с множеством стоков графа (вершин, из которых не выходят ребра).

- Для каждого состояния $s \in \mathcal{S}$ рассмотрим множество возможных действий $\mathcal{A}_s$, а именно множество возможных состояний, в которые можно попасть из s. Формально $\mathcal{A}_s = \{s' \in \mathcal{S} | (s, s') \in \mathcal{E}\}$

- Генеративная модель GFlowNet $\mathcal{F}$ задается как множество распределений для действий из каждого состояния. Формально для всех состояний $s \in \mathcal{S}$ и следующих состояний $s' \in \mathcal{A}_s$ мы знаем неотрицательные числа $\mathcal{F}(s \to s')$, в соответствии с которыми будут генерироваться действия.

Определим процесс генерации финальных объектов из множества $\mathcal{X}$ с помощью заданной модели GFlowNet $\mathcal{F}$:

- Стартуем из стартового состояния $s := s_0$.

- До тех пор пока $s \notin \mathcal{X}$ генерируем следующее состояние $s' \in \mathcal{A}_s$ из вероятностного распределения $\mathcal{P}(s'|s) = \frac{1}{Z_s}\mathcal{F}(s \to s')$, где $Z_s = \sum\limits_{s' \in \mathcal{A}_s} \mathcal{F}(s \to s')$. После этого делаем переход в сгенерированную вершину $s := s'$.

- Полученное финальное состояние $s \in \mathcal{X}$ в конце процесса генерации будет являться сгенерированным объектом.

На практике GFlowNet-ы могут быть применены следующим образом. Обычно нам известно некоторое множество объектов $\mathcal{X}$ и reward функция $\mathcal{R}$ из реальной жизни (например, это могут быть геномные строки или молекулярные графы и reward функция взятая из их химических или физических свойств). Задача состоит в том, чтобы задать процесс конструирования объектов (граф $\mathcal{G}$) и построить для него модель $\mathcal{F}$ так, что описанный генеративный процесс будет генерировать объекты из желаемого распределения, заданного $\mathcal{R}$. Обычно в реальности множество объектов $\mathcal{X}$ очень большое и распределения $\mathcal{F}$ параметризуются нейронной сетью.

Задача

Вам полностью дан граф генеративных переходов $\mathcal{G}$, а также желаемое вероятностное распределение на финальных объектах $x \in \mathcal{X}$.

Постройте какую-нибудь GFlowNet модель $\mathcal{F}$, которая будет генерировать финальные объекты из заданного распределения.

Для полного балла также требуется, чтобы **по каждому ребру существовала ненулевая вероятность пройти**, то есть $\mathcal{F}(s \to s') > 0$ для всех $(s, s') \in \mathcal{E}$.

Формат входных данных

В первой строке находится единственное целое число n ($2 \leqslant n \leqslant 50$) — количество состояний, то есть $n = |\mathcal{S}|$.

Для простоты пусть множество состояний будет $\mathcal{S} = \{1, 2, \ldots, n\}$, а стартовое состояние $s_0 = 1$.

В каждой из следующих n строк задаются возможные переходы из каждого состояния.

В строке с номером s сначала находится целое число k_s ($0 \leqslant k_s \leqslant n - s$) — количество возможных переходов из s, то есть $k_s = |\mathcal{A}_s|$. Затем следует k_s целых чисел $s'_1, s'_2, \ldots, s'_{k_s}$ ($s'_j > s$). Множество переходов задается как $\mathcal{A}_s = \{s'_1, s'_2, \ldots, s'_{k_s}\}$. Гарантируется, что заданные s'_i различны.

Заметим, что заданный ориентированный граф $\mathcal{G}$ является ациклическим (поскольку все ребра ведут из вершин с меньшим номером в вершины с большим номером). Гарантируется, что все вершины достижимы из вершины 1 по ребрам.

Множество финальных объектов $\mathcal{X}$ задается как множество стоков графа $\mathcal{G}$ (множество состояний s, для которых $|\mathcal{A}_s| = 0$).

В последней строке задано n целых чисел $w_1, w_2, \ldots, w_n$ ($0 \leqslant w_i \leqslant 10^6$). Гарантируется, что $w_i = 0$, если $i \notin \mathcal{X}$ и $w_i > 0$, иначе.

Желаемое вероятностное распределение на финальных объектах $x \in \mathcal{X}$ задается как $\mathcal{P}(x) = \frac{w_x}{\sum\limits_{i=1}^{n} w_i}$.

Формат выходных данных

Выведите любую подходящую GFlowNet модель $\mathcal{F}$ в следующем формате:

Несколько строк, каждая строка содержит три **целых** числа s, s', $\mathcal{F}(s \to s')$ ($1 \leqslant s, s' \leqslant n$, $0 \leqslant \mathcal{F}(s \to s') < 10^{300}$) — начало и конец ребра из графа $\mathcal{G}$ и значение GFlownet модели $\mathcal{F}$ для него. Для всех строк (s, s') должно быть ребром графа $\mathcal{G}$, для каждого ребра должно быть не более одной строки с этим ребром. Для всех не выведенных ребер (s, s') будет считаться, что $\mathcal{F}(s \to s') = 0$. Ребра можно выводить в любом порядке.

Если рассмотреть подграф, построенный на ребрах (s, s'), таких что $\mathcal{F}(s \to s') > 0$, то все вершины из $\mathcal{X}$ должны быть достижимы из 1, а также не должно существовать других стоков в нем. Для вершин, не достижимых в этом подграфе, можно не задавать значения $\mathcal{F}$ (они могут быть все нулевыми).

- Выведенная GFlowNet модель $\mathcal{F}$ должна генерировать в точности из заданного распределения. Если это не будет выполнено или ваш ответ не будет соответствовать формату в хотя бы одном из тестов, ваше решение получит 0 **баллов**.
- Если все ответы корректные, но в каком-то из тестов существует хотя бы одно ребро (s, s'), такое что $\mathcal{F}(s \to s') = 0$, то ваше решение получит 4 **балла**.
- Если во всех тестах для всех ребер $\mathcal{F}(s \to s') > 0$, ваше решение получит полные 10 **баллов**.

Можно показать, что в заданных ограничениях всегда существует ответ, получающий полный балл. Обратите внимание, что в данной задаче вам может потребоваться длинная арифметика.

Пример

стандартный ввод	стандартный вывод
5 4 3 4 5 2 0 2 5 4 0 0 0 3 0 4 5	3 5 5 1 3 9 1 4 4 3 4 4 1 5 5 1 2 6