

Вопрос: Можно ли утверждать, что *Интервальное планирование* \leq_p *Вершинное покрытие*?

(b) Вопрос: Можно ли утверждать, что

Независимое множество \leq_p *Интервальное планирование*?

2. Для анализа поведения покупателей в магазинах часто ведется двумерный массив A , строки которого соответствуют клиентам, а столбцы — продаваемым товарам. Элемент $A[i, j]$ определяет количество единиц товара j , приобретенных покупателем i . Ниже приведен небольшой пример такого массива.

	Моющее средство	Пиво	Пеленки	Наполнитель для кошачьего туалета
Радж	0	6	0	3
Аланис	2	3	0	0
Челси	0	0	0	7

Один из способов обработки полученных данных выглядит так: допустим, подмножество S покупателей называется *разнородным*, если никакие два покупателя из S никогда не покупали один и тот же товар (то есть каждый товар был куплен не более чем одним покупателем из S). Например, разнородное множество покупателей может пригодиться для проведения маркетингового исследования.

Теперь мы можем определить задачу о разнородном множестве следующим образом: для заданного массива A размером $m \times n$, определенного выше, и числа $k \leq m$ существует ли разнородное подмножество, содержащее не менее k покупателей?

Покажите, что задача о разнородном множестве является NP-полной.

3. При организации летнего спортивного лагеря возникает следующая задача. В лагере должен быть как минимум один квалифицированный тренер для каждого из n видов спорта (баскетбол, волейбол и т. д.). Организаторы получают заявки от m потенциальных тренеров. Для каждого из n видов спорта существует подмножество из m претендентов, квалифицированных в этом виде спорта. Вопрос заключается в следующем: для заданного числа $k < m$ возможно ли нанять не более k тренеров, чтобы для каждого из n видов спорта нашелся хотя бы один квалифицированный тренер? Назовем эту задачу *задачей эффективного найма*.

Покажите, что задача эффективного найма является NP-полной.

4. Предположим, вы занимаетесь администрированием высокопроизводительной системы реального времени, в которой асинхронные процессы работают с общими ресурсами. В системе существует множество из n процессов и множество из m ресурсов. В любой заданный момент времени каждый процесс определяет набор ресурсов, которые он намеревается использовать. Каждый ресурс может запрашиваться сразу несколькими процессами, но использоваться в любой момент времени он может только одним процессом. Ваша задача — распределить ресурсы между запрашивающими их процессами. Если процесс получает все

запрошенные им ресурсы, он остается активным; в противном случае его выполнение блокируется. Соответственно *задача резервирования ресурсов* формулируется следующим образом: для заданных множество процессов и ресурсов, множества запрашиваемых ресурсов для каждого процесса и числа k возможно ли распределить ресурсы между процессами так, чтобы не менее k процессов оставались активными?

Рассмотрите задачи из следующего списка и для каждой задачи либо приведите алгоритм с полиномиальным временем, либо докажите, что задача является NP -полной.

- (а) Общая задача резервирования ресурсов, определенная выше.
- (б) Частный случай задачи при $k = 2$.
- (с) Частный случай задачи с двумя типами ресурсов (допустим, люди и оборудование); каждый процесс запрашивает максимум один ресурс каждого типа (другими словами, каждый процесс запрашивает одного конкретного человека и одно конкретное устройство).
- (д) Частный случай задачи, в котором каждый ресурс запрашивается максимум двумя процессами.

5. Рассмотрим множество $A = \{a_1, \dots, a_n\}$ и набор B_1, B_2, \dots, B_m подмножеств A (то есть $B_i \subseteq A$ для всех i).

Множество $H \subseteq A$ называется *множеством представителей* для набора B_1, B_2, \dots, B_m , если H содержит минимум один элемент из каждого B_i — то есть если $H \cap B_i$ не пусто для каждого i (таким образом, H содержит «представителей» из всех множеств B_i).

Задача множества представителей определяется следующим образом: имеется множество $A = \{a_1, \dots, a_n\}$, набор B_1, B_2, \dots, B_m подмножеств A и число k . Существует ли множество представителей $H \subseteq A$ для B_1, B_2, \dots, B_m , размер которого не превышает k ?

Докажите, что задача множества представителей является NP -полной.

6. Рассмотрим экземпляр задачи выполнимости, заданный условиями C_1, \dots, C_k по множеству булевых переменных x_1, \dots, x_n . Экземпляр называется *монотонным*, если литерал в каждом условии состоит из неинвертированной переменной, то есть каждый литерал равен x_i для некоторого i , а не \bar{x}_i . Монотонные экземпляры задачи выполнимости решаются очень легко: они всегда выполнимы присваиванием каждой переменной 1.

Предположим, имеются три условия

$$(x_1 \vee x_2), (x_1 \vee x_3), (x_2 \vee x_3).$$

Экземпляр является монотонным, и присваивание, задающее все три переменные равными 1, действительно выполняет все условия. Но это не единственное выполняющее присваивание; также можно было присвоить x_1 и x_2 значение 1, а x_3 значение 0. В самом деле, для любого монотонного экземпляра естественно задать вопрос о минимальном количестве переменных, которым необходимо присвоить 1 для его выполнения.

Задача монотонной выполнимости с минимумом истинных переменных формулируется так: для заданного монотонного экземпляра задачи выполнимости и числа k существует ли выполняющее присваивание для экземпляра, в котором не более k переменных задано значение 1? Докажите, что задача является NP-полной.

7. Так как задача о трехмерном сочетании является NP-полной, естественно ожидать, что аналогичная задача о четырехмерном сочетании будет хотя бы не менее сложной. Определим *четырёхмерное сочетание* следующим образом: для заданных множеств W, X, Y и Z , каждое из которых имеет размер n , и набора C упорядоченных четверок в форме (w_p, x_j, y_k, z_l) существуют ли n четверок из C , среди которых никакие два не имеют общих элементов?

Докажите, что задача о четырехмерном сочетании является NP-полной.

8. Малолетняя дочь ваших знакомых Мэдисон учится читать. Чтобы помочь ей, родители купили набор цветных магнитов на холодильник с буквами алфавита (несколько экземпляров буквы А, несколько экземпляров буквы В и т. д.). При последней встрече вы строили из магнитов слова, которые она уже знает. Изначальные условия игры немного изменились. Вскоре вы пытались строить слова так, чтобы использовать все магниты из набора — то есть выбирали слова, которые она уже знает, но так, чтобы после выкладывания всех слов каждый магнит участвовал в записи ровно одного слова. (Разрешается повторение слов; например, если в набор магнитов входят по два экземпляра букв С, А и Т, слово САТ можно построить дважды.)

Задача оказалась довольно сложной, и только позднее вы поняли причину. Допустим, мы рассматриваем обобщенную версию *задачи использования всех магнитов*, в которой английский алфавит заменяется произвольным набором символов, а словарь Мэдисон моделируется произвольным набором строк из символов этого набора. Цель остается той же, что в предыдущем абзаце.

Докажите, что задача использования всех магнитов является NP-полной.

9. Вы управляете сетью передачи данных, моделируемой направленным графом $G = (V, E)$. Имеется s пользователей, которые намерены использовать эту сеть. Пользователь i (для всех $i = 1, 2, \dots, s$) выдает *запрос* на резервирование пути P_i в графе G , по которому должны передаваться данные.

Вы заинтересованы в том, чтобы принять как можно больше таких запросов, со следующим ограничением: если приняты запросы на пути P_i и P_j , то P_i и P_j не могут содержать общих узлов.

Итак, задача о выборе путей формулируется следующим образом: для направленного графа $G = (V, E)$, множества запросов P_1, P_2, \dots, P_s (каждый из которых определяет путь в G) и числа k возможно ли выбрать как минимум k путей, чтобы никакие два из выбранных путей не содержали общих узлов?

Докажите, что задача о выборе путей является NP-полной.

10. Ваши друзья из фирмы WebExodus недавно консультировали компании с большими общедоступными сайтами (условия контракта не позволяют сказать, какими именно), и в ходе работы они столкнулись со следующей *задачей стратегического размещения рекламы*.

Компания предоставляет карту сайта, которая моделируется направленным графом $G = (V, E)$. Компания также предоставляет множество t типичных маршрутов перемещения по сайту; маршруты моделируются направленными путями P_1, P_2, \dots, P_t в графе G (то есть каждый маршрут P_i представляется путем в G). Компания хочет получить от WebExodus ответ на следующий вопрос: для заданного графа G , путей $\{P_i\}$ и числа k возможно ли разместить рекламу не более чем на k узлах G , чтобы каждый путь P_i включал как минимум один узел с рекламой? Назовем эту задачу «задачей стратегического размещения рекламы» со входными данными $G, \{P_i: i = 1, \dots, t\}$ и k .

Ваши друзья считают, что хороший алгоритм решения задачи их обогатит; к сожалению, все не так просто.

(а) Докажите, что задача стратегического размещения рекламы является NP -полной.

(б) Ваши друзья из WebExodus написали довольно быстрый алгоритм S , который выдает ответы «да/нет» для произвольного экземпляра задачи стратегического размещения рекламы. Будем считать, что алгоритм S всегда работает правильно.

Используя алгоритм S как «черный ящик», разработайте алгоритм, который получает входные данные $G, \{P_i\}$ и k из части (а) и делает одно из двух:

- выводит множество, содержащее не более k узлов из G , для которого каждый путь P_i содержит как минимум один такой узел, или
- выводит (обоснованно) сообщение о том, что такое множество из максимум k узлов не существует.

Ваш алгоритм должен использовать не более полиномиального количества вычислительных шагов, в сочетании с не более чем полиномиальным количеством вызовов алгоритма S .

11. Как еще помнят некоторые (а другие слышали), идея гипертекста появилась за десятилетия до Всемирной паутины. Даже гипертекстовая литература появилась относительно давно: вместо линейной печатной страницы автор пишет историю, состоящую из набора связанных виртуальных мест, соединенных виртуальными «проходами»¹. Таким образом, гипертекстовое произведение в действительности строится на основе направленного графа; для конкретности (хотя и с сужением полного диапазона возможностей) мы воспользуемся следующей моделью.

Будем рассматривать структуру гипертекстового произведения как направленный граф $G = (V, E)$. Каждый узел $u \in V$ содержит некоторый текст; когда читатель достигает u , он может выбрать любое ребро, ведущее из u ; и если было выбрано ребро $e = (u, v)$, читатель переходит к узлу v . Чтение начинается с узла $s \in V$ и заканчивается на узле $t \in V$; когда читатель впервые достигает t , история завершается. Следовательно, любой путь от s к t является действительным сюжетом произведения. Учтите, что в отличие от просмотра страниц

¹ Например, см. <http://www.eastgate.com>.

в браузере возможность возврата может отсутствовать; после перехода от u к v вы уже можете никогда не вернуться к u .

В такой модели гипертекстовая структура определяет огромное количество разных сюжетов с одним базовым контентом; и отношения между многочисленными вариантами порой становятся весьма нетривиальными. Пример задачи, встречающейся при таком подходе к структуре: рассмотрим фрагмент гипертекстового произведения на базе графа $G = (V, E)$, в котором содержатся некоторые ключевые тематические элементы: любовь, смерть, война, стремление получить ученую степень по информатике и т. д. Каждый тематический элемент i представляется множеством $T_i \subseteq V$, состоящим из всех узлов G , в которых встречается данная тема. Теперь для заданного множества тематических элементов можно задать вопрос: существует ли действительное развитие сюжета, в котором встречается каждый из таких элементов? Конкретнее, для направленного графа G с начальным узлом s и конечным узлом t и тематическими элементами, представленными множествами T_1, T_2, \dots, T_k , *задача воплощения сюжета* формулируется так: существует ли путь из s в t , содержащий по крайней мере один узел каждого из множеств T_i ?

Докажите, что задача воплощения сюжета является NP-полной.

12. Ваши друзья открыли популярный сайт с новостями и форумом. Трафик достиг уровня, при котором они хотят различать платных и бесплатных посетителей. Обычно в такой схеме весь контент сайта доступен для посетителей, оплачивающих ежемесячную подписку; посетители, которые не платят за пользование сайтом, могут просматривать некоторое подмножество страниц (и надоедливую рекламу, которая предлагает им перейти на платную подписку).

Существуют два простых способа управления доступом для бесплатных посетителей: (1) фиксированное подмножество страниц помечается как разрешенное для просмотра бесплатными посетителями или (2) просматриваться могут все страницы, но ограничивается максимальное количество страниц, которые могут просматриваться бесплатным посетителем за один сеанс. (Будем считать, что сайт может отслеживать действия посетителя на сайте.)

Ваши друзья экспериментируют с моделью ограничения доступа — более хитрой по сравнению с обоими вариантами. Они хотят, чтобы бесплатные посетители могли получить представление о разных разделах сайта, для чего подмножества страниц обозначаются как образующие некую *зону*: например, на сайте одна зона страниц может быть посвящена политике, другая — музыке и т. д. Страница может принадлежать сразу нескольким зонам. Когда бесплатный пользователь перемещается по сайту, политика доступа позволяет ему посетить одну страницу из каждой зоны, но попытка обратиться ко второй странице из той же зоны пресекается. (Вместо этого пользователь направляется на страницу с предложением оформить подписку.)

Формально сайт может моделироваться направленным графом $G = (V, E)$, в котором ребра представляют веб-страницы, а ребра — направленные гиперссылки. Существует *входной узел* $s \in V$ и *зоны* $Z_1, \dots, Z_k \subseteq V$. Путь P , выбираемый бесплатным пользователем, ограничивается включением одного узла из каждой зоны Z .

Один из недостатков усложненной политики доступа заключается в том, что с ней трудно ответить даже на самые простейшие вопросы относительно достижимости узлов, например: может ли бесплатный пользователь посетить заданный узел t ? Более точная формулировка *задачи проблемного пути* выглядит так: для заданного графа $G, Z_1, \dots, Z_k, s \in V$ и конечного узла $t \in V$ существует ли в G путь $s-t$, включающий не более одного узла из каждой зоны Z_i ? Докажите, что задача проблемного пути является NP -полной.

13. Механизм *комбинаторного аукциона* был разработан экономистами для продажи множества объектов множеству потенциальных покупателей. (Федеральное агентство по связи США изучало эту разновидность аукционов для распределения радиочастот среди вещательных компаний.)

Простой пример комбинаторного аукциона: имеются n продаваемых объектов с метками I_1, \dots, I_n . Каждый объект является неделимым и может быть продан только одному покупателю. Далее, m разных покупателей делают ставки: i -я ставка задает подмножество S_i объектов, а x_i — цена, которую покупатель готов заплатить за объекты множества S_i как за единое целое. (Ставка будет представляться парой (S_i, x_i) .)

Теперь аукционист просматривает множество всех m заявок; он выбирает одни заявки и отклоняет другие. Каждый покупатель, ставка которого i будет принята, покупает все объекты соответствующего множества S_i . Следовательно, две принятые ставки не могут содержать множества, содержащие хотя бы один общий элемент, так как один объект не может быть продан двум разным людям.

Аукционист получает сумму предложенных цен по всем принятым ставкам. (Обратите внимание: у каждого покупателя всего одна попытка, сделать новую ставку уже не получится.) Цель аукциониста — получить как можно больше денег.

Задача определения победителя для комбинаторного аукциона формулируется так: для заданных объектов I_1, \dots, I_n , ставок $(S_1, x_1), \dots, (S_m, x_m)$ и границы B существует ли набор ставок, которые может принять аукционист, чтобы заработанная сумма была не меньше B ?

Пример. Предположим, аукционист решает воспользоваться описанным методом для продажи лишнего компьютерного оборудования. Продаются четыре объекта: «PC», «монитор», «принтер» и «сканер»; трое покупателей делают ставки. Определяем:

$$S_1 = \{PC, \text{монитор}\}, S_2 = \{PC, \text{принтер}\}, S_3 = \{\text{монитор, принтер, сканер}\}$$

и

$$x_1 = x_2 = x_3 = 1$$

Сделаны ставки (S_1, x_1) , (S_2, x_2) и (S_3, x_3) , и граница B равна 2.

В данном экземпляре ответ будет отрицательным: аукционист сможет принять не более одной ставки (так как у любых двух ставок имеются общие объекты), поэтому общая выручка не может превысить 1.

Докажите, что задача определения победителя для комбинаторного аукциона является NP-полной.

14. Задача интервального планирования упоминалась в главах 1 и 4. Рассмотрим намного более сложную с вычислительной точки зрения ее разновидность, которую мы назовем *задачей множественного интервального планирования*. Как и прежде, имеется процессор, способный выполнять задания в течение некоторого периода времени (например, с 9:00 до 17:00).

Пользователи поставляют задания, которые должны выполняться на процессоре; в любой момент времени на процессоре может выполняться только одно задание. Однако в этой модели задания имеют более сложную структуру: каждое задание определяет множество интервалов, в течение которых ему требуется использовать процессоры. Итак, например, одно задание может занимать процессор с 10:00 до 11:00, а потом с 14:00 до 15:00. Если принять это задание, оно занимает процессор в течение двух часов, но вы можете принимать задания на другие периоды (включая промежуток с 11:00 до 14:00).

Имеется множество из n заданий, каждое из которых определяется множеством интервалов. Требуется найти ответ на следующий вопрос: для заданного числа k возможно ли принять не менее k заданий так, чтобы никакие два задания не перекрывались по времени?

Покажите, что задача множественного интервального планирования является NP-полной.

15. Однажды во время работы вам приносят пакет FedEx. В пакете лежит сотовый телефон, который начинает звонить, — оказывается, это ваш друг Нео, от которого уже давно не было никаких известий. Все разговоры с Нео всегда проходят одинаково: сначала он долго и пафосно объясняет, почему он обратился к вам, но в итоге все сводится к просьбам потратить ваше время на то, чтобы помочь ему с решением некоторой проблемы.

На этот раз по причинам, о которых он вынужден умолчать (что-то типа защиты подземного города от роботов-убийц), ему и его знакомым требуется отслеживать радиосигналы в разных точках электромагнитного спектра. Всего есть n разных частот, для наблюдения за которыми существует массив датчиков.

Задача отслеживания состоит из двух компонентов:

- множество L из m мест, в которых могут быть размещены датчики;
- множество S из b источников помех, каждый из которых блокирует некоторые частоты в некоторых местах. А конкретно каждый источник помех i задается парой (F_i, L_i) , где F_i — подмножество частот, а L_i — подмножество мест; это означает, что датчик, помещенный в любое из мест множества L_i , не сможет получать сигналы на любых частотах множества F_i (из-за воздействия помех).

Подмножество $L' \subseteq L$ мест называется *достаточным*, если для каждой из n частот j существует некоторое место в L' , в котором частота j не блокируется ни одним источником помех. Тогда, размещая датчик в каждом месте достаточного множества, вы сможете успешно отслеживать все n частот.

У соратников Нео есть k датчиков, и они хотят узнать, существует ли достаточное множество мест с размером не более k . Назовем следующую формулировку *задачей отслеживания частот*: для заданных частот, мест, источников помех и параметра k существует ли достаточное множество с размером не более k ?

Пример. Допустим, заданы четыре частоты $\{f_1, f_2, f_3, f_4\}$ и четыре места $\{\ell_1, \ell_2, \ell_3, \ell_4\}$. Существуют три источника помех:

$$(F_1, L_1) = (\{f_1, f_2\}, \{\ell_1, \ell_2, \ell_3\})$$

$$(F_2, L_2) = (\{f_3, f_4\}, \{\ell_3, \ell_4\})$$

$$(F_3, L_3) = (\{f_2, f_3\}, \{\ell_1\})$$

В этой ситуации существует достаточное множество с размером 2: можно выбрать места ℓ_2 и ℓ_4 . (Так как f_1 и f_2 не блокируются в ℓ_4 , а f_3 и f_4 не блокируются в ℓ_2).

Докажите, что задача отслеживания частот является NP -полной.

16. Рассмотрим задачу характеристики множества по размерам его пересечений с другими множествами. Имеется конечное множество U размера n , а также набор A_1, \dots, A_m подмножеств U . Также заданы числа c_1, \dots, c_m . Вопрос звучит так: существует ли такое множество $X \subset U$, что для всех $i = 1, 2, \dots, m$ мощность $X \cap A_i$ равна c_i ? Назовем его *задачей выведения пересечений* с входными данными $U, \{A_i\}$ и $\{c_i\}$.

Докажите, что задача выведения пересечений является NP -полной.

17. Задан направленный граф $G = (V, E)$ с весами w_e ребер $e \in E$. Веса могут быть отрицательными или положительными. В задаче *цикла с нулевым весом* требуется решить, существует ли в G простой цикл, сумма весов ребер которого равна 0. Докажите, что эта задача является NP -полной.
18. Вас попросили помочь теоретикам в анализе данных, связанных с принятием решений. В частности, аналитики рассматривают набор данных по решениям, принятых некоторым правительственным комитетом, и пытаются выявить малое подмножество влиятельных членов комитета.

Состав комитета представлен множеством $M = \{m_1, \dots, m_n\}$ из n членов. За последний год в комитете проводилось голосование по t разным вопросам. По каждому вопросу каждый член комитета мог проголосовать «За», «Против» или «Воздержаться». В результате комитет принимает положительное решение по каждому вопросу, если количество голосов «За» строго превышает количество голосов «Против» (воздержавшиеся не считаются ни за одну из сторон); в противном случае принимается отрицательное решение.

Имеется большая таблица, содержащая данные голосования каждого участника комитета по каждому вопросу. Рассматривается следующее определение: подмножество членов $M' \subseteq M$ называется *решающим*, если при проверке только голосов, поданных членами M' , решение комитета по каждому вопросу будет тем же. (Иначе говоря, результат голосования членов M' по каждому вопросу

определяет общий результат голосования всего комитета.) Такое подмножество можно рассматривать как своего рода «ближнее окружение», которое отражает мнение комитета в целом.

Вопрос: для данных о голосовании каждым членом комитета по каждому вопросу и заданного параметра k требуется узнать, существует ли решающее подмножество, содержащее не более k членов. Назовем эту формулировку *задачей решающего подмножества*.

Пример. Допустим, имеются четыре члена комитета и три вопроса. Мы ищем решающее подмножество с размером не более $k = 2$, а голосование проходило так:

Вопрос	m_1	m_2	m_3	m_4
Вопрос 1	За	За	Воздержаться	Против
Вопрос 2	Воздержаться	Против	Против	Воздержаться
Вопрос 3	За	Воздержаться	За	За

В данном экземпляре результат будет положительным, поскольку члены m_1 и m_3 образуют решающее подмножество.

Докажите, что задача решающего подмножества является NP-полной.

19. Предположим, вы консультируете управление порта в маленькой стране на побережье Тихого океана. Ежегодный объем сделок исчисляется миллиардами, и доход ограничивается исключительно скоростью разгрузки кораблей, прибывающих в порт.

Работа с опасными грузами усложняет и без того непростую задачу. Допустим, утром в порт прибывает морской конвой, который привозит n контейнеров с разными видами опасных материалов. В доке стоят m машин, каждая из которых вмещает до k контейнеров.

Существуют две взаимосвязанные задачи, которые обусловлены разными типами ограничений, действующих при обработке опасных материалов. Для каждой из двух задач выберите один из двух ответов:

- Алгоритм с полиномиальным временем для решения задачи.
- Доказательство NP-полноты задачи.

(а) Для каждого контейнера существует заданное подмножество машин, на которых они могут перевозиться. Возможно ли погрузить все n контейнеров на m машин, чтобы ни одна машина не была перегружена, а каждый контейнер был погружен на машину, для которой разрешена его перевозка?

(б) В другой версии задачи любой контейнер может быть погружен на любую машину, однако некоторые пары контейнеров не могут быть погружены на одну машину. (При контакте содержащихся в них химикатов может произойти взрыв.) Возможно ли погрузить все n контейнеров на m машин так, чтобы ни одна машина не была перегружена и никакие два контейнера не грузились на одну машину, если это запрещено правилами?

20. Существует много разных способов формального представления задачи кластеризации, целью которой является разбиение набора объектов на «похожие» группы.

Естественным способом представления входных данных для задачи кластеризации является множество объектов p_1, p_2, \dots, p_n , в котором для каждой пары определяется числовое расстояние $d(p_i, p_j)$. (Для расстояний устанавливаются требования: $d(p_i, p_i) = 0$; $d(p_i, p_j) > 0$ для разных p_i и p_j и симметричность $d(p_i, p_j) = d(p_j, p_i)$.)

Ранее в разделе 4.7 была рассмотрена одна из возможных формулировок задачи кластеризации: разбить объекты на k множеств так, чтобы максимизировать минимальное расстояние между парами объектов из разных кластеров. Как выясняется, задача решается удачным применением задачи о минимальном остовном дереве.

Другой, но на первый взгляд похожий способ формализации может выглядеть следующим образом: разбить объекты на k множеств так, чтобы минимизировать максимальное расстояние между любой парой объектов в пределах одного кластера. Обратите внимание на отличия: если в формулировке из предыдущего абзаца нужно было найти кластеры, среди которых не было бы двух «слишком близких», в новой формулировке кластеры не должны быть слишком «широкими», то есть ни один кластер не должен содержать двух точек, находящихся на большом расстоянии друг от друга.

Тот факт, что новая формулировка обладает слишком большой вычислительной сложностью для оптимального решения, может показаться удивительным (при таком сходстве задач). Чтобы задачу можно было проанализировать в контексте NP -полноты, сначала запишем ее в формулировке с принятием решения. Для заданных n объектов p_1, p_2, \dots, p_n с расстояниями, определенными выше, и границы B *задача кластеризации малого диаметра* формулируется следующим образом: возможно ли разбить объекты на k множеств так, чтобы расстояние между любыми двумя точками, принадлежащими одному множеству, не превышало B ?

Докажите, что задача кластеризации малого диаметра является NP -полной.

21. После знакомства с популярной книгой для начинающих предпринимателей вы осознаете, что компьютерная система в вашем офисе нуждается в обновлении. Однако это начинание сталкивается с некоторыми непростыми проблемами.

При планировании новой системы необходимо выбрать k компонентов: операционная система, редактор текстов, почтовый клиент и т. д. Для j -го компонента системы существует множество A_j вариантов; а конфигурация системы состоит из выбора одного элемента из каждого из множеств A_1, A_2, \dots, A_k .

Проблемы возникают из-за того, что некоторые пары вариантов из разных множеств оказываются несовместимыми. Вариант $x_i \in A_i$ и вариант $x_j \in A_j$ образуют несовместимую пару, если они не могут содержаться в одной системе: например, Linux (вариант операционной системы) и Microsoft Word (вариант редактора текстов) образуют несовместимую пару. Конфигурация системы называется *полностью совместимой*, если она состоит из элементов $x_1 \in A_1, x_2 \in A_2, \dots, x_k \in A_k$, из которых ни одна пара (x_i, x_j) не является несовместимой.

Теперь мы можем определить *задачу полностью совместимой конфигурации*. Экземпляр такой задачи состоит из непересекающихся множеств вариантов A_1, A_2, \dots, A_k и множества P несовместимых пар (x, y) , в котором x и y являются элементами разных множеств вариантов. Требуется решить, существует ли полностью совместимая конфигурация: выбор элемента из каждого множества вариантов, при котором ни одна пара выбранных элементов не принадлежит множеству P .

Пример. Допустим, $k = 3$, а множества A_1, A_2, A_3 обозначают варианты выбора операционной системы, редактора текстов и почтового клиента соответственно:

$$A_1 = \{\text{Linux, Windows NT}\},$$

$$A_2 = \{\text{emacs, Word}\},$$

$$A_3 = \{\text{Outlook, Eudora, rmail}\}.$$

Множество несовместимых пар выглядит так:

$$P = \{(\text{Linux, Word}), (\text{Linux, Outlook}), (\text{Word, rmail})\}$$

В этом экземпляре ответ на версию с принятием решения будет положительным: например, выбор $\text{Linux} \in A_1, \text{emacs} \in A_2, \text{rmail} \in A_3$ образует полностью совместимую конфигурацию в соответствии с предшествующими определениями.

Докажите, что задача полностью совместимой конфигурации является NP-полной.

22. Предположим, вы получили алгоритм A — «черный ящик», который получает ненаправленный граф $G = (V, E)$ и число k и ведет себя следующим образом:

- Если граф G не является связным, алгоритм просто возвращает сообщение о несвязности.
- Если граф G является связным и в нем есть независимое множество с размером не менее k , алгоритм возвращает сообщение «Да».
- Если граф G является связным, но в нем нет независимого множества с размером не менее k , алгоритм возвращает сообщение «Нет».

Предположим, алгоритм A выполняется за время, полиномиальное по размеру G и k .

Покажите, как с использованием обращений к A решить за полиномиальное время задачу о независимом множестве: для заданного графа G и числа k содержит ли граф G независимое множество с размером не менее k ?

23. Для множества конечных двоичных строк $S = \{s_1, \dots, s_k\}$ строка u называется *конкатенацией* по S , если она равна $s_{i_1} s_{i_2} \dots s_{i_l}$ для некоторых индексов $i_1, \dots, i_l \in \{1, \dots, k\}$.

Ваш знакомый рассматривает следующую задачу: для заданных двух множеств конечных двоичных строк $A = \{a_1, \dots, a_m\}$ и $B = \{b_1, \dots, b_n\}$ существует ли строка u , которая бы одновременно являлась конкатенацией по A и конкатенацией по B ?

Ваш знакомый заявляет: «По крайней мере задача принадлежит NP, поскольку для доказательства положительного ответа мне достаточно предъявить такую строку u ». Вы указываете (разумеется, вежливо), что такое объяснение совершенно несостоятельно: как узнать, что кратчайшая из таких строк u не имеет

длины, экспоненциальной по размеру входных данных, — ведь в этом случае она не может служить сертификатом полиномиального размера?

Но как выясняется, это заявление можно преобразовать в доказательство принадлежности NP . А именно, докажите следующее утверждение:

Если существует строка u , которая является конкатенацией и по A , и по B , то существует строка, длина которой ограничивается полиномиально по сумме длин строк в $A \cup B$.

24. Имеется двудольный граф $G = (V, E)$; допустим, его узлы разбиты на множества X и Y так, что один конец каждого ребра принадлежит X , а другой принадлежит Y . Определим (a, b) -скелет графа G как множество ребер $E' \subseteq E$, для которого не более a узлов в X инцидентно ребру в E' и не менее b узлов в Y инцидентно ребру в E' .

Покажите, что для заданного двудольного графа G и чисел a и b задача принятия решения о существовании (a, b) -скелета является NP -полной.

25. Для функций g_1, \dots, g_l функция $\max(g_1, \dots, g_l)$ определяется по формуле

$$[\max(g_1, \dots, g_l)](x) = \max(g_1(x), \dots, g_l(x)).$$

Рассмотрим следующую задачу: имеется n кусочно-линейных непрерывных функций f_1, \dots, f_n , определенных на интервале $[0, t]$ для некоторого целого t . Также задано целое число B . Требуется решить: существуют ли k функций f_{i_1}, \dots, f_{i_k} , для которых

$$\int_0^t [\max(f_{i_1}, \dots, f_{i_k})](x) dx \geq B?$$

Докажите, что эта задача является NP -полной.

26. Разъезжая по отдаленным частям света, вы со своим другом накопили большую грудку сувениров. В то время вы не задумывались над тем, какие сувениры останутся вам, а какие — ему, но, похоже, пришло время разделить приобретения.

Один из возможных способов дележа выглядит так: допустим, имеется n объектов $1, 2, \dots, n$ и объект i имеет согласованную ценность x_i . (Например, это может быть сумма, за которую его можно продать; будем считать, что у вас с друзьями нет расхождений по поводу оценки). Можно попытаться найти разбиение объектов на два множества, чтобы общая ценность объектов в каждом множестве была одинаковой.

Для этого следует решить следующую задачу численного разбиения. Даны положительные целые числа x_1, \dots, x_n ; требуется решить, возможно ли разбить числа на два множества S_1 и S_2 с одинаковой суммой:

$$\sum_{x_i \in S_1} x_i = \sum_{x_j \in S_2} x_j.$$

Покажите, что задача численного разбиения является NP -полной.

27. Рассмотрим следующую задачу. Даны положительные целые числа x_1, \dots, x_n , а также числа k и B . Требуется узнать, возможно ли разбить числа $\{x_i\}$ на k множеств S_1, \dots, S_k так, чтобы квадраты сумм множеств в совокупности давали не более B :

$$\sum_{i=1}^k \left(\sum_{x_j \in S_i} x_j \right)^2 \leq B.$$

Покажите, что эта задача является NP-полной.

28. Ниже описана разновидность задачи о независимом множестве. Имеется граф $G = (V, E)$ и целое число k . В контексте этой задачи множество $I \subset V$ будет называться *строго независимым*, если для любых двух узлов $v, u \in I$ ребро (v, u) не принадлежит E , а также не существует пути из двух ребер из u в v — то есть не существует узла w , для которого $(u, w) \in E$ и $(w, v) \in E$. Задача о *строго независимом множестве* требует решить, существует ли в G строго независимое множество с размером не менее k .

Докажите, что задача о строго независимом множестве является NP-полной.

29. Вы строите большую сеть рабочих станций, которая будет моделироваться ненаправленным графом G ; узлы G представляют отдельные рабочие станции, а ребра представляют прямые каналы связи. Всем рабочим станциям необходим доступ к центральной базе данных с информацией, необходимой для работы базовых функций операционной системы.

Конечно, базу данных можно реплицировать на каждой рабочей станции; в этом случае обращения с любой станции будут выполняться очень быстро, но вам придется поддерживать большое количество копий. Также можно сохранить одну копию базы данных на одной рабочей станции, к которой другие рабочие станции будут обращаться за информацией по сети, но это создаст большие задержки для рабочих станций, находящихся на расстоянии многих сетевых переходов от места хранения базы данных.

Похоже, стоит поискать компромисс: ограничиться небольшим количеством копий базы данных, но разместить их так, чтобы любая рабочая станция либо содержала собственную копию, либо была соединена прямым каналом с рабочей станцией, на которой имеется копия. В терминологии графов такое множество узлов называется *доминирующим*.

Задача о доминирующем множестве формулируется следующим образом: для заданной сети G и числа k возможно ли разместить k копий базы данных в k разных узлах так, чтобы каждый узел либо содержал собственную копию базы данных либо был соединен прямым каналом с узлом, содержащим копию?

Покажите, что задача о доминирующем множестве является NP-полной.

30. Порой даже в традиционных задачах «непрерывной» математики совершенно неожиданно проявляются дискретные, комбинаторные аспекты из изучаемой нами области, которые проникают даже во вполне стандартные на первый взгляд задачи.

Для примера рассмотрим традиционную задачу минимизации функции одной переменной — допустим, $f(x) = 3 + x - 3x^2$ на интервале $x \in [0, 1]$. Производная рав-

на нулю в точке $x = 1/6$, но это точка максимума, а не минимума функции; чтобы получить минимум, следует выполнить стандартную проверку значений на границах интервала (причем минимум в данном случае достигается на границе $x = 1$).

Проверка границы не создает серьезных проблем для функций одной переменной; но предположим, что в задаче требуется минимизировать функцию n переменных x_1, x_2, \dots, x_n на единичном кубе, где все $x_1, x_2, \dots, x_n \in [0, 1]$. Минимум может достигаться как внутри куба, так и на его границах; и последняя перспектива выглядит устрашающе, так как куб состоит из 2^n «углов» (для которых каждое из значений x_i равно либо 0, либо 1), а также из различных фрагментов других измерений. Из-за этой сложности книги по матанализу подозрительно нечетко излагают этот вопрос при описании многофакторных задач оптимизации.

И тому есть веская причина: минимизация функции с n переменными в единичном n -мерном кубе относится к классу NP -полных задач. Чтобы это утверждение стало более конкретным, рассмотрим специальный случай многочленов с целочисленными коэффициентами по n переменным x_1, x_2, \dots, x_n . На всякий случай освежим в памяти терминологию: *одночленом* называется произведение вещественного коэффициента c и каждой из переменных x_i , возведенной в неотрицательную целую степень a_i ; это можно записать в виде $cx_1^{a_1}cx_2^{a_2}\cdots cx_n^{a_n}$. (Например, $2x_1^2x_2x_3^4$ является одночленом). *Многочленом* называется сумма конечного набора одночленов (например, $2x_1^2x_2x_3^4 + x_1x_3 - 6x_2^2x_3^2$ является многочленом).

Задача минимизации многофакторного многочлена формулируется так: для заданного многочлена по n переменным с целочисленными коэффициентами и заданной целой границей B возможно ли выбрать вещественные числа $x_1, x_2, \dots, x_n \in [0, 1]$, с которыми многочлен достигает значения $\leq B$?

Выберите в этой главе задачу Y , которая заведомо является NP -полной, и покажите, что $Y \leq_p$ *Минимизация многофакторного многочлена*.

31. Для заданного ненаправленного графа $G = (V, E)$ множеством *обратной связи* называется множество $X \subseteq V$, обладающее тем свойством, что $G-X$ не содержит циклов. Задача ненаправленного множества обратной связи формулируется так: для заданных G и k содержит ли G множество обратной связи с размером не более k ? Докажите, что задача ненаправленного множества обратной связи является NP -полной.
32. Расшифровка генома сопряжена с множеством сложных вычислительных задач. На простейшем уровне каждая хромосома организма может рассматриваться как чрезвычайно длинная строка (возможно, состоящая из миллионов символов) четырехбуквенного алфавита $\{A, C, G, T\}$. Один из классов методов расшифровки генома базируется на генерировании большого количества коротких, перекрывающихся фрагментов хромосомы, с последующим построением полной длинной строки, представляющей хромосому, из этого набора перекрывающихся подстрок.

Хотя мы не будем рассматривать задачи сборки строк во всех подробностях, ниже приведена упрощенная задача, которая дает некоторое представление

о вычислительной сложности задач этой области. Допустим, имеется множество $S = \{s_1, s_2, \dots, s_n\}$ коротких строк ДНК из q -буквенного алфавита; каждая строка s_i имеет длину 2ℓ для некоторого числа $\ell \geq 1$. Также имеется библиотека дополнительных строк $T = \{t_1, t_2, \dots, t_m\}$ из символов того же алфавита; каждая из этих строк также имеет длину 2ℓ . Пытаясь определить, может ли строка s_b следовать прямо за строкой s_a в хромосоме, мы проверяем, содержит ли библиотека T строку t_k , чтобы первые ℓ символов t_k совпадали с последними ℓ символами в s_a , а последние ℓ символов t_k совпадали с первыми ℓ символами в s_b . Если это возможно, мы говорим, что t_k *скрепляет* пару (s_a, s_b) . (Другими словами, t_k может быть фрагментом ДНК, «перекрывающим» точку перехода s_a в s_b .)

Нам хотелось бы сцепить все строки S в некотором порядке без перекрытий, так чтобы каждая последовательная пара скреплялась некоторой строкой из библиотеки T . Следовательно, мы хотим упорядочить строки S в виде $s_{i_1}, s_{i_2}, \dots, s_{i_n}$, где i_1, i_2, \dots, i_n — такая перестановка $\{1, 2, \dots, n\}$, что для всех $j = 1, 2, \dots, n-1$ существует строка t_k , скрепляющая пару $(s_{i_j}, s_{i_{j+1}})$. (Одна строка t_k может использоваться для нескольких последовательных пар в конкатенации.) Если это возможно, говорят, что у множества S существует *идеальная сборка*.

Для заданных множеств S и T задача об идеальной сборке формулируется так: существует ли для S идеальная сборка по отношению к T ? Докажите, что задача об идеальной сборке является NP-полной.

Пример. Допустим, задан алфавит $\{A, C, G, T\}$, множество $S = \{AG, TC, TA\}$ и множество $T = \{AC, CA, GC, GT\}$ (так, что каждая строка имеет длину $2\ell = 2$). Тогда ответ на этот экземпляр задачи об идеальной сборке будет положительным: три строки S можно сцепить в порядке $TCAGTA$ (так, что $s_{i_1} = s_2$, $s_{i_2} = s_1$ и $s_{i_3} = s_3$). В этом порядке пара (s_{i_1}, s_{i_2}) скрепляется строкой CA из библиотеки T , а пара (s_{i_2}, s_{i_3}) сцепляется строкой GT из библиотеки T .

33. В натуральном хозяйстве люди обмениваются товарами и услугами напрямую, без участия денег как промежуточного шага. Обмен происходит тогда, когда каждая сторона рассматривает множество получаемых товаров как обладающее большей ценностью, чем множество товаров, которые отдаются взамен. Исторически человеческое общество обычно переходило от натурального хозяйства к денежной экономике; соответственно различные сетевые системы, экспериментирующие с бартером, могут рассматриваться как намеренные попытки вернуться к более ранней форме экономических взаимодействий. При этом обнаруживаются некоторые проблемы, изначально присущие натуральному обмену по сравнению с системами, использующими деньги. Одной из таких трудностей является сложность нахождения возможностей для обмена (даже когда такие возможности существуют).

Для моделирования этих сложностей мы введем концепцию *ценности*, которая назначается каждым человеком каждому объекту в реальном мире. Предположим, имеется множество из n людей p_1, \dots, p_n и множество m объектов a_1, \dots, a_m . Каждый объект принадлежит одному человеку. Далее, у каждого человека p_i имеется оценочная функция v_i , которая определяется так, что $v_i(a_j)$ является

неотрицательным числом, которое определяет ценность объекта a_j для p_i , — чем больше число, тем большую ценность объект представляет для человека.

Обмен между двумя людьми в такой системе возможен в том случае, если существуют люди p_i и p_j , и подмножества объектов A_i и A_j , принадлежащих p_i и p_j соответственно, так что для каждого участника предпочтительными являются объекты того подмножества, которое ему пока не принадлежит. А точнее, суммарная ценность объектов A_j для участника p_i превышает суммарную ценность объектов A_i , и суммарная ценность объектов A_i для участника p_j превышает суммарную ценность объектов A_j . (Обратите внимание: подмножество A_i не обязано включать все объекты, принадлежащие p_i ; то же относится к A_j .) A_i и A_j могут быть произвольными подмножествами принадлежащих им объектов, удовлетворяющим этим критериям.)

Допустим, вы получили экземпляр задачи натурального обмена, определяемый приведенными выше данными ценности объектов (чтобы избежать проблем с представлением вещественных чисел, будем считать, что ценности всех объектов представляют собой натуральные числа). Докажите, что задача определения возможности обмена между двумя людьми является *NP*-полной.

34. В 1970-е годы ученые, работающие в области математической социологии (включая Марка Грановеттера и Томаса Шеллинга), взялись за разработку моделей некоторых видов коллективного человеческого поведения: почему одни модные «фишки» приживаются, а другие быстро отмирают? Почему одни технологические новшества распространяются повсеместно, а другие не выходят за пределы малых групп пользователей? Какова динамика беспорядков и грабежей, которые иногда проявляются в поведении разъяренной толпы? Ученые предположили, что все эти ситуации являются примерами *каскадных процессов*, при которых поведение участника сильно зависит от поведения его близкого окружения; процесс, запущенный несколькими участниками, распространяется на все большее количество людей и в конечном итоге получает очень широкое воздействие. Этот процесс можно сравнить с распространением слухов или эпидемии, передаваемой от человека к человеку.

Рассмотрим простейшую версию модели каскадных процессов. Имеется некоторое поведение (сбори́ние марок, покупка сотового телефона, участие в беспорядках), и в какой-то момент времени каждый человек либо принимает в нем участие, либо не принимает. Население представляется направленным графом $G = (V, E)$, в котором узлы соответствуют людям, а ребро (v, w) означает, что человек v влияет на поведение человека w ; если человек v участвует в поведении, это также способствует принятию поведения человеком w . У каждого человека w также имеется порог $\theta(w) \in [0, 1]$, который имеет следующий смысл: в любой момент, когда доля узлов, от которых выходят ребра в w , достигнет по меньшей мере $\theta(w)$, узел w также начинает участвовать в поведении.

Узлы с низким порогом быстрее «заражаются» поведением, узлы с более высоким порогом более консервативны. Узел w с порогом $\theta(w) = 0$ принимает поведение немедленно, без влияния окружения. Наконец, нужно договориться по поводу узлов без входящих ребер: будем считать, что они принимают поведение, если $\theta(w) = 0$, и не принимают его при более высоком пороге.

Для заданного экземпляра этой модели распространение поведения может быть смоделировано следующим образом.

Инициализировать все узлы w с $\theta(w) = 0$ как участников поведения.

(Все остальные узлы в исходном состоянии не участвуют в поведении.)

Пока множество участников поведения продолжает изменяться:

 Для всех людей, не участвующих в поведении, одновременно:

 Если среди узлов с ребрами к w доля участников не менее $\theta(w)$
 w становится участником

 Конец Если

Конец

Конец

Вывести итоговое множество участников

Этот процесс конечен, так как потенциальных участников всего n и при каждой итерации по крайней мере один человек становится участником поведения.

В последние годы исследователи, работающие в области маркетинга и анализа данных, изучали возможность применения этой модели для анализа эффектов «сарафанного радио» в успехе новых продуктов (так называемый *вирусный маркетинг*). Под поведением понимается использование нового продукта; нужно попытаться убедить нескольких ключевых представителей населения опробовать новый продукт и надеяться, что это вызовет по возможности значительный каскадный эффект.

Допустим, выбирается множество узлов $S \subseteq V$, а порог каждого узла в S обнуляется 0. (Убедив людей опробовать новый продукт, мы гарантируем их участие.) Затем запускается описанный выше процесс, и мы смотрим, насколько большим окажется итоговое множество участников. Обозначим размер итогового множества участников $f(S)$ (размер записывается в виде функции от S , так как он естественно зависит от выбора S). Величина $f(S)$ может рассматриваться как *влияние* множества S , так как она отражает распространение поведения, «посянного» в S .

При планировании маркетинговой кампании целью является нахождение малого множества S , имеющего как можно большее влияние $f(S)$. *Задача максимизации влияния* определяется так: для заданного направленного графа $G = (V, E)$ с пороговыми значениями для каждого узла, параметров k и b существует ли множество S , содержащее не более k узлов, для которых $f(S) \geq b$?

Докажите, что задача максимизации влияния является NP-полной.

Пример. Допустим, граф $G = (V, E)$ состоит из пяти узлов $\{a, b, c, d, e\}$, четырех ребер (a, b) , (b, c) , (e, d) , (d, c) , а пороги всех узлов установлены равными $2/3$. В этом случае ответ на экземпляр задачи максимизации влияния, определяемой графом G , с $k = 2$ и $b = 5$, будет положительным: если выбрать $S = \{a, e\}$, то все три остальных узла также станут участниками. (В этой задаче такой вариант выбора S только один; например, если выбрать $S = \{a, d\}$, то b и c станут участниками, но e не станет; а если выбрать $S = \{a, b\}$, то ни один из узлов c , d или e не станет участником.)

35. Трое ваших друзей работают в крупной компании, занимающейся компьютерными играми. Вот уже несколько месяцев они работают над идеей новой игры «Droid Trader!», одобренной руководством компании. За это время им пришлось выслушать множество обескураживающих замечаний, от «Вам придется основательно поработать с отделом маркетинга над названием» до «А может, стоит добавить побольше крови?»

Но сейчас практически все уверены, что игра пойдет в производство, и друзья обращаются к вам за решением последней проблемы: а не получится ли игра слишком простой — игроки слишком быстро освоят ее и им станет скучно?

Вы не сразу разбираетесь в происходящем по их подробному описанию; но если отбросить все космические сражения, поединки и псевдомистицизм в духе «Звездных войн», основная идея выглядит так. Игрок управляет космическим кораблем и пытается заработать деньги, покупая и продавая дроидов на разных планетах. Всего существуют n типов дроидов и k планет. Каждая планета p обладает следующими свойствами: на ней продаются $s(j, p) \geq 0$ дроидов типа j по фиксированной цене $x(j, p) \geq 0$ за штуку, для всех $j = 1, 2, \dots, n$; также существует спрос на $d(j, p) \geq 0$ дроидов типа j по фиксированной цене $y(j, p) \geq 0$. (Будем считать, что на планете не может быть положительного спроса одновременно с положительным предложением на дроидов одного типа; таким образом, по крайней мере одно из значений $s(j, p)$ и $d(j, p)$ равно 0.)

Игрок начинает на планете s с z единицами денег и заканчивает на планете t ; существует направленный ациклический граф G для множества планет, в котором пути $s \rightarrow t$ соответствуют действительным маршрутам игрока. (Граф G выбран ациклическим, чтобы избежать бесконечных игр.) На протяжении пути $s \rightarrow t$ в G игрок может участвовать в сделках следующим образом. Прибывая на планету p в пути P , он может купить до $s(j, p)$ дроидов типа j по цене $x(j, p)$ за штуку (при условии, что у него хватит денег) и/или продать до $d(j, p)$ дроидов типа j по цене $y(j, p)$ за штуку (для $j = 1, 2, \dots, n$). Счет игрока определяется суммой денег, которую он имеет на руках при прибытии на планету t . (Также начисляются дополнительные очки за космические сражения и поединки, но для формулировки задачи они нас не интересуют.)

Итак, задача фактически сводится к получению наибольшего результата. Другими словами, для заданного экземпляра игры с направленным ациклическим графом G на базе множества планет, всех остальных параметров, описанных выше, и целевого ограничения B существует ли путь P в G и последовательность сделок в P , при которой игрок заканчивает со счетом не менее B ? Назовем этот экземпляр *задачей максимального счета в игре Droid Trader!*, или сокращенно ЗМCDT!

Докажите, что задача ЗМCDT! является NP -полной (в предположении $P \neq NP$). Тем самым вы гарантируете отсутствие простой стратегии для получения высокого счета в игре ваших друзей.

36. Даже если вы знаете людей много лет, понять их порой бывает нелегко. Для примера возьмем хотя бы Раджа и Аланис. Ни один из них не относится к «жаворонкам», но сейчас они ежедневно встают в 6 утра и отправляются на рынок за свежими фруктами и овощами для своего ресторана здорового питания.

Стремясь сэкономить на продуктах, они столкнулись с непростой проблемой. Существует большое множество из n ингредиентов I_1, I_2, \dots, I_n (пучки зелени, бутылки рисового уксуса и т. д.). Ингредиент I_j покупается в единицах по $s(j)$ граммов (с покупкой целого количества единиц) и стоит $c(j)$ долларов за единицу. Кроме того, он остается годным для использования в течение $t(j)$ дней с даты покупки.

За следующие k дней ваши друзья должны подать k фирменных блюд, по одному в день. (Выбор дня, в который должно готовиться то или иное блюдо, остается на их усмотрение.) Фирменное блюдо с номером i использует подмножество $S_i \subseteq \{I_1, I_2, \dots, I_n\}$ ингредиентов, а именно $a(i, j)$ граммов ингредиента I_j . Наконец, существует еще одно ограничение: постоянные клиенты ресторана останутся постоянными только в том случае, если они получают еду, приготовленную из самых свежих продуктов. Для каждого фирменного блюда ингредиенты S_i делятся на два подмножества: те, которые должны покупаться в день изготовления фирменного блюда, и те, которые могут быть куплены в любой день, но их срок годности еще не истек. (Например, салат-мексиканский с базиликом должен готовиться из базилика, купленного в этот же день, а в соусе песто с рукколой, базиликом и корнелльским козьим сыром можно использовать базилик, купленный несколько дней назад, если он еще не завял).

Здесь-то и появляется возможность сэкономить на продуктах. Часто при покупке единицы ингредиента I_j для приготовления фирменного блюда на этот день им не нужен весь запас. Если они смогут на следующей день приготовить другое блюдо, которое тоже использует I_j , но не требует, чтобы ингредиент был куплен в тот же день, можно сэкономить за счет использования ранее купленных продуктов. Конечно, если блюда с базиликом будут следовать друг за другом, это может усложнить изготовление рецептов с козьим сыром и т. д., — собственно, здесь и возникает сложность.

Итак, *задача планирования фирменных блюд* выглядит так: для заданной информации об ингредиентах и рецептах, а также бюджета x возможно ли спланировать k фирменных блюд так, чтобы общая сумма, потраченная на ингредиенты за k дней, не превышала x ?

Докажите, что задача планирования фирменных блюд является NP-полной.

37. Как ни странно, даже в «Звездных войнах» кроется немало NP-полных задач.

Рассмотрим задачу, с которой столкнулись Люк, Лея и их друзья на пути со «Звезды Смерти» на скрытую базу повстанцев. Галактику можно рассматривать как ненаправленный граф $G = (V, E)$, в котором узлы представляют звездные системы, а ребро (u, v) обозначает возможность прямого перемещения из u в v . «Звезде Смерти» соответствует узел s , а скрытой базе повстанцев — узел t . Расстояния, представленные ребрами графа, не одинаковы; каждому ребру e назначается целочисленная длина $\ell_e \geq 0$. Кроме того, некоторые ребра представляют маршруты, активно патрулируемые имперскими кораблями; соответственно каждому ребру e также назначается целочисленный риск $r_e \geq 0$, который определяет прогнозируемые повреждения от насыщенных спецэффектами космических сражений в случае перемещения по этому ребру.

Самый безопасный маршрут пролегает по дальним секторам Галактики, от одной безопасной звездной системы к другой; но в этом случае топливо у корабля кончится задолго до того, как он доберется до места назначения. Самый быстрый маршрут — рвануть через населенный центр, но тогда придется отбиваться от слишком большого количества имперских кораблей. В общем случае общая длина любого пути P из s в t определяется как сумма длин всех его ребер, а общий риск — как сумма рисков всех его ребер.

Итак, Люк, Лея и компания пытаются решить для этого графа усложненную разновидность задачи о кратчайшем пути: им нужно добраться из s в t по пути с разумно малой общей длиной и общим риском. А конкретнее *задачу о кратчайшем галактическом пути* можно сформулировать так: для заданной конфигурации, описанной выше, и целочисленных ограничений L и R существует ли путь из s в t , в котором общая длина не превышает L и при этом общий риск не превышает R ?

Докажите, что задача о кратчайшем галактическом пути является NP -полной.

38. Рассмотрим разновидность задачи о дереве Штейнера, которую мы будем называть *графическим деревом Штейнера*. Имеется ненаправленный граф $G = (V, E)$, множество вершин $X \subseteq V$ и число k . Требуется решить, существует ли множество $F \subseteq E$, содержащее не более k ребер, для которого в графе (V, F) X принадлежит одной компоненте связности.

Покажите, что задача о графическом дереве Штейнера является NP -полной.

39. *Задача о направленных непересекающихся путях* определяется следующим образом: имеется направленный граф G и k пар узлов $(s_1, t_1), (s_2, t_2), \dots, (s_k, t_k)$. Требуется решить, существуют ли непересекающиеся по узлам пути P_1, P_2, \dots, P_k такие, что P_i идет из s_i в t_i .

Покажите, что задача о направленных непересекающихся путях является NP -полной.

40. Рассмотрим задачу, встречающуюся при проектировании схем широкополосной рассылки в сети. Имеется направленный граф $G = (V, E)$ с узлом $r \in V$ и назначенным множеством «целевых узлов» $T \subseteq V - \{r\}$. С каждым узлом v связано *время переключения* s_v , которое представляет собой положительное целое число.

Во время 0 узел r генерирует сообщение, которое должно быть получено каждым узлом в T . Для этого нужно построить схему, в которой r передает информацию некоторым из своих соседей (последовательно), которые в свою очередь сообщают некоторым из своих соседей и т. д., пока сообщение не будет получено каждым узлом T . На более формальном уровне схема рассылки определяется следующим образом: узел r может отправить копию сообщения одному из своих соседей во время 0; сосед получит сообщение во время 1. В общем случае во время $t \geq 0$ любой узел v , уже получивший сообщение, может отправить копию сообщения одному из своих соседей — при условии, что он еще не отправлял копию сообщения в любой из моментов $t - s_v + 1, t - s_v + 2, \dots, t - 1$. (В этом проявляется роль времени переключения: v необходима пауза в $s_v - 1$ шагов между последовательными отправками сообщения. Если $s_v = 1$, никакие дополнительные ограничения не добавляются.)

Временем завершения схемы рассылки называется минимальное время t , к которому все узлы T получают сообщение. Задача о времени рассылки определяется так: для входных данных, описанных выше, и границы b существует ли схема рассылки с временем завершения, не превышающим b ?

Докажите, что задача о времени рассылки является NP-полной.

Пример. Допустим, имеется направленный граф $G = (V, E)$ с множеством $V = \{r, a, b, c\}$ и ребрами (r, a) , (a, b) , (r, c) ; множество $T = \{b, c\}$ и время переключения составляет $s_v = 2$ для всех $v \in V$. В этом случае схема рассылки с минимальным временем завершения будет выглядеть так: r отправляет сообщение a во время 0; a отправляет сообщение b во время 1; r отправляет сообщение c во время 2; схема завершается во время 3, когда c получает сообщение. (Обратите внимание: узел a может отправить сообщение при получении его во время 1, так как это его первая отправка сообщения; но узел r не может отправить сообщение во время 1, так как $s_r = 2$, а сообщение было отправлено во время 0.)

41. Для направленного графа G *циклическим покрытием* называется множество циклов, не пересекающихся по узлам, в котором каждый узел в G принадлежит циклу. В задаче о *циклическом покрытии* требуется узнать, имеет ли заданный направленный граф циклическое покрытие.

(а) Покажите, что задача о циклическом покрытии может быть решена за полиномиальное время. (Подсказка: воспользуйтесь задачей о двудольном паросочетании.)

(б) Предположим, каждый цикл должен содержать не менее трех ребер. Покажите, что проверка существования такого циклического покрытия для графа G является NP-полной.

42. Компания, занимающаяся проектированием коммуникационных сетей, обращается к вам со следующей задачей. Изучается конкретная сеть, состоящая из n узлов и моделируемая направленным графом $G = (V, E)$. По соображениям отказоустойчивости G требуется разбить на максимально возможное количество виртуальных «доменов»: *доменом* в G называется множество X узлов с размером не менее 2, так что для каждой пары узлов $u, v \in X$ существуют направленные пути из u в v и из v в u , полностью содержащиеся в X .

Покажите, что следующая *задача доменной декомпозиции* является NP-полной. Имеется направленный граф $G = (V, E)$ и число k ; возможно ли разбить V не менее чем на k множеств, каждое из которых является доменом?

Примечания и дополнительная литература

В примечаниях к главе 2 упоминаются некоторые ранние работы по формализации вычислительной эффективности при полиномиальном времени; концепция NP-полноты сформировалась на основе этих работ и заняла центральное место в вычислительной теории после работ Кука (Cook, 1971), Левина (Levin, 1973) и Карпа (Karp, 1972). Эдмондс (Edmonds, 1965) обратил особое внимание на класс задач