Школа лингвистики, 2018-19 уч. год
Дискретная математика для лингвистов
Графы-1 (21 октября 2017г.)

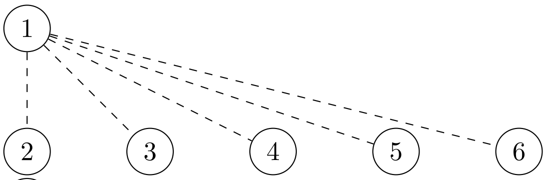
В. В. Кочергин, А. В. Михайлович

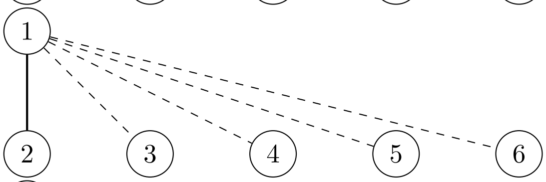
В окружающем нас мире существует большое количество всевозможных связей между различными объектами. Так, например, дороги соединяют между собой города, провода соединяют электроприборы, дружба соединяет между собой людей, а ссылки соединяют страницы в интернете. Часто структуры связей между объектами определенного вида становятся предметом научных исследований, в том числе лингвистических, социологических, кибернетических, логистических. Если число рассматриваемых объектов велико, структура связей между ними может оказаться достаточно сложной и запутанной. Тем не менее, необходимо на основе ее свойств искать эффективные алгоритмы ее изучения. При этом возникает потребность в построении математической модели структуры связей. Такой математической моделью является граф.

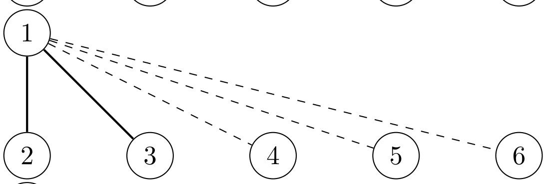
Начнём с одной простой задачи, ставшей уже классической.

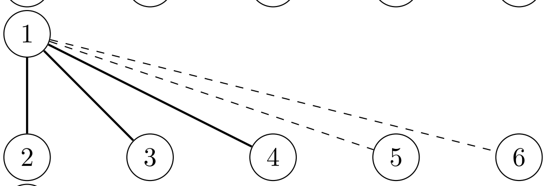
Рассмотрим следующую задачу. Нужно показать, что в любой компании, состоящей из 6 и более человек, всегда есть либо трое попарно знакомых, либо трое попарно незнакомых.

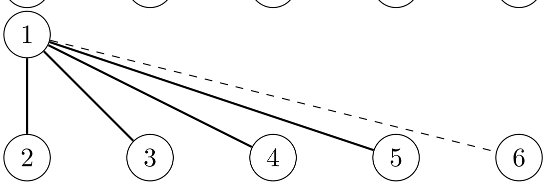
Для начала рассмотрим вариант этой задачи с 6 людьми в компании. Людей будем изображать точками на плоскости, знакомство — жирной линией, соединяющей соответствующие точки, а незнакомство — пунктиром. Посмотрим, как соединена точка, соответствующая человеку 1, с другими точками. Поскольку всего 5 линий, то в любом случае есть либо 3 (или больше) жирных, либо 3 (или больше) пунктирных.

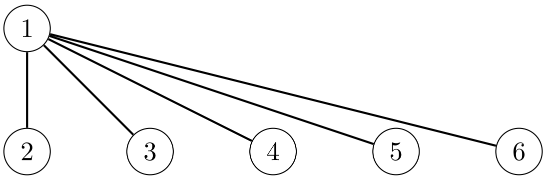




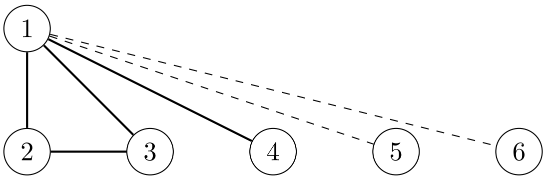




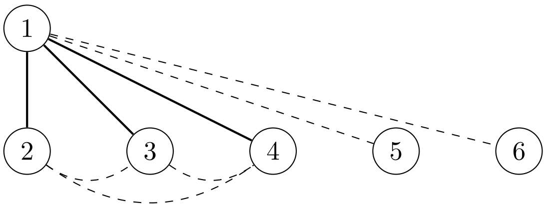




Рассмотрим случай с тремя жирными линиями (случай с тремя пунктирными линиями рассматривается аналогично). Пусть они соединяют точку 1 с точками 2, 3 и 4. Если среди линий, объединяющих точки 2, 3 и 4 есть жирная линия (например, соединяющая точки 2 и 3), то образуется треугольник из жирных линий, соответствующий трем попарно знакомым людям.



Если же среди этих линий нет жирных, то точки 2, 3 и 4 попарно соединены пунктиром. А значит, образуют тройку попарно незнакомых людей.



Таким образом, задачу для 6 человек решили.

Пусть у нас теперь в компании больше, чем 6 людей. Аналогичным образом соединяем первого со всеми остальными и рассматриваем те соединения, которых больше двух, и соответствующие им вершины.

Перейдём к строгим определениям.

Сопоставим рассматриваемые объекты с элементами некоторого (непустого) множества V. Эти элементы называются *вершинами.* Связь между двумя объектами мы будем представлять в виде пары соответствующих вершин. Такие пары называются *ребрами.* Множество ребер мы обозначим через E. *Графом* G называется совокупность множеств V и E (обозначается также $G = (V, E)$).

Пример 1. Рассмотрим следующие пары множеств V и E. Какие из них являются графами, а какие не являются и почему?

1. $V = \{v_1, v_2, v_3, v_4, v_5\}$, $E = \{(v_1, v_2), (v_1, v_3), (v_2, v_4), (v_3, v_5), (v_2, v_5)\}$.
2. $V = \{v_2, v_3, v_4, v_5\}$, $E = \{(v_2, v_4), (v_3, v_5), (v_2, v_5)\}$.
3. $V = \{v_1, v_3, v_4, v_5\}$, $E = \{(v_1, v_2), (v_1, v_3), (v_2, v_4), (v_3, v_5), (v_2, v_5)\}$.
4. $V = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$, $E = \{(v_1, v_2), (v_1, v_3), (v_2, v_4), (v_3, v_5), (v_2, v_5)\}$.
5. $V = \varnothing$, $E = \varnothing$. [1]

Поскольку в зависимости от постановки задачи исследуемые связи могут быть как двусторонними (например, договоры о сотрудничестве между компаниями), так и односторонними (например, авиаперелеты между городами), то необходимо условиться о том, считаем мы ребра упорядоченными парами вершин или неупорядоченными (т.е. различаются пары (u, v) и (v, u), или это одна и та

[1] Графами не являются третье и пятое множества. В третьем множестве ребра содержат вершину v_2, которой нет в множестве V, а в пятом множество вершин пустое.

же пара). В первом случае ребра называют *ориентированными ребрами*, а граф — *ориентированным графом*. Во втором случае граф называется *неориентированным*.

Пример 2. Что из перечисленного можно описать ориентированным графом, а что — неориентированным? Если связи не указаны или неочевидны, рассмотрите различные варианты.

1. Дороги в городе.[2]
2. Тротуары.[3]
3. Метро (Москва).[4]
4. Автобусные маршруты (Москва).[5]
5. Несколько людей в социальной сети.[6]
6. Группу людей, среди которых есть как знакомые, так и незнакомые.[7]
7. Генеалогическое древо.[8]
8. Сотрудники НИУ ВШЭ.[9]

Рассмотрим граф $G = (V, E)$. Пусть $u, v \in V$. Тогда, если $e = (u, v) \in E$, то говорят, что ребро e и вершины u и v *инцидентны* друг другу, а также, что ребро e соединяет вершины u и v, или что вершины u и v являются концами ребра e. Можно также сказать, что ребро e и вершина v инцедентны друг другу, если ребро e содержит вершину v. Если речь идет об ориентированном графе, то говорят, что ребро e выходит из вершины u и входит в вершину v.

Пример 3. Какие вершины и ребра являются инцедентными в следующем графе $G = (V, E)$, где $V = \{v_1, v_2, v_3, v_4, v_5\}$, $E = \{e_1, e_2, e_3, e_4, e_5\}$, $e_1 = (v_1, v_2)$, $e_2 = (v_1, v_3)$, $e_3 = (v_2, v_4)$, $e_4 = (v_3, v_5)$, $e_5 = (v_2, v_5)$?

При решении некоторых задач полезно также считать, что несколько различных ребер в множестве E могут представлять собой одну и ту же пару вершин (например, несколько различных способов доставки из одного пункта в другой). Такие ребра называются *кратными*. Граф, в котором допустимы кратные ребра, называется *мультиграфом*. Кроме того, при некоторых постановках приходится рассматривать ребра, у которых оба конца совпадают. Такие ребра называются *петлями*, а графы с петлями и кратными ребрами называются *псевдографами*. Если же постановка задачи не допускает использования кратных ребер и петель, то рассматриваемый граф называется *простым графом*. Отметим, что в разных источниках опускаются слова или приставки «мульти», «псевдо» или простой, в зависимости от тем и задач, которым посвящен этот источник. Например, в нашем курсе, если не сказано иное, рассматривается простой граф.

Пример 4. Что из перечисленного является простым графом, псевдографом, мультиграфом?

1. Граф $G = (V, E)$, где $V = \{v_1, v_2, v_3, v_4, v_5\}$, $E = \{e_1, e_2, e_3, e_4, e_5\}$, $e_1 = (v_1, v_2)$, $e_2 = (v_1, v_3)$, $e_3 = (v_2, v_4)$, $e_4 = (v_3, v_5)$, $e_5 = (v_2, v_5)$?
2. Граф, где вершинами являются аккаунты студентов вконтакте, посещающих курс дискретной математики, а ребрами — отношение «являться друзьями».
3. Московский метрополитен (станции пересадок считаются за одну вершину).
4. Граф $G = (V, E)$, где $V = \{v_1, v_2, v_3, v_4, v_5\}$, $E = \{e_1, e_2, e_3, e_4, e_5\}$, $e_1 = (v_1, v_2)$, $e_2 = (v_1, v_3)$, $e_3 = (v_2, v_4)$, $e_4 = (v_3, v_3)$, $e_5 = (v_2, v_5)$?
5. Граф $G = (V, E)$, где $V = \{v_1, v_2, v_3, v_4, v_5\}$, $E = \{e_1, e_2, e_3, e_4, e_5\}$, $e_1 = (v_1, v_2)$, $e_2 = (v_1, v_3)$, $e_3 = (v_2, v_4)$, $e_4 = (v_2, v_5)$, $e_5 = (v_2, v_5)$?
6. Граф, в котором вершинами являются все корпуса НИУ ВШЭ в г. Москва, а ребрами — дороги, их соединяющие. [10]

[2] Неориентированным графом можно описать только в том случае, когда все дороги двусторонние. В общем случае лучше использовать ориентированный граф.

[3] Тротуары, в отличие от дорог, двусторонние. Описание неориентированным графом подходит.

[4] Нериентированный граф.

[5] Ориентированный граф, так как не все остановки парные, не все маршруты ходят по одним и тем же дорогам.

[6] Если сеть допускает только двусторонние связи, то неориентированный, а если односторонние (например, вконтакте есть «подписчики»), то ориентированный.

[7] Поскольку знакомство является взаимным отношением, то можно описать ориентированным графом.

[8] Обычно в качестве связи используется отношение A является родителем B, поэтому граф ориентированный.

[9] Если в качестве связи рассматривать отношение A является начальником B, то ориентированный.

[10] 1,2,3 — простые графы, 4 — псевдограф, 5,6 — мультиграфы.

Визуально граф часто удобно представлять в виде схемы, на которой вершины изображены с помощью точек, а ребра — линий, соединяющих эти точки. Такое изображение графа называется его *геометрической реализацией*. С подобными схемами мы часто сталкиваемся в повседневной жизни, например, в метро. Вершинами графа метрополитена являются станции, а ребрами — перегоны и пересадки. Более простой пример дает граф $G_1 = (V, E)$, где $V = \{v_1, v_2, v_3, v_4, v_5\}$, а $E = \{(v_1, v_2), (v_2, v_3), (v_3, v_4), (v_4, v_5), (v_5, v_1)\}$. Геометрической реализацией этого графа является пятиугольник.

Несмотря на то, что задание графа с помощью геометрической реализации вполне наглядно для человека, в памяти компьютера граф намного удобнее хранить в виде матрицы.

Пусть в графе n вершин. Присвоим каждой вершине графа номер, и обозначим вершину с номером i (где $1 \le i \le n$) через v_i. Построим матрицу с n строками и n столбцами. На пересечении строки с номером i и столбца с номером j записывают количество ребер, которые соединяют эти вершины v_i и v_j. (Для простого графа это число может принимать только значения 0 и 1.) Такая матрица называется *матрицей смежности*.

Пример 6. Для графа $G = (V, E)$, где $V = \{v_1, v_2, v_3, v_4, v_5\}$, а $E = \{(v_1, v_2), (v_2, v_3), (v_3, v_4), (v_4, v_5), (v_5, v_1)\}$ матрицей смежности является следующая матрица.

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Следует отметить, что для простого неориентированного графа такая матрица будет являться симметричной, а по главной ее диагонали будут расположены нули. При этом сумма чисел в строке с номером i будет в точности равна степени вершины v_i. Для ориентированного графа симметричность может нарушаться. При этом при вычислении исходящей степени вершины v_i нужно вычислить сумму в строке с номером i, а для вычисления входящей степени — в столбце с номером i.

Кроме матриц смежности на практике иногда используются также *матрицы инцидентности*. Для построения матрицы инцидентности необходимо отдельно занумеровать ребра графа. Обозначим число ребер через m, а ребро с номером j (здесь $1 \le j \le m$) — через e_j. Тогда матрица инцидентности состоит из n строк и m столбцов. Для неориентированного графа на пересечении строки с номером i и столбца с номером j записывают единицу, если вершина v_i и ребро e_j инцидентны.

Пример 7. Для графа $G = (V, E)$, где $V = \{v_1, v_2, v_3, v_4, v_5\}$, а $E = \{(v_1, v_2), (v_2, v_3), (v_3, v_4), (v_4, v_5), (v_5, v_1)\}$ матрицей инцедентности является следующая матрица (ребра занумерованы в том пордяке, в котором они перечислены при определении графа).

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

Для простого неориентированного графа в каждом столбце матрицы инцидентности находится ровно две единицы, а сумма чисел в строке с номером i совпадает со степенью вершины v_i.

При задании матрицы инцидентности для ориентированного графа следует обращать внимание на то, является ребро e_j входящим в вершину v_i или исходящим. В первом случае на пересечении строки номер i и столбца номер j пишут минус единицу, а во втором — единицу. (При этом если ребро является одновременно входящим и исходящим, т.е. петлей, то принято писать ноль.) Таким образом, в каждом столбце матрицы инцидентности сумма чисел равна нулю. Количество единиц в строке с номером i равно исходящей степени вершины v_i, а количество минус единиц — входящей.

Пусть у нас есть граф $G = (V, E)$, $V = \{v_1, \ldots, v_n\}$, $E = \{e_1, \ldots, e_n\}$. Последовательность

$$v_{i_0} e_{j_1} v_{i_1} e_{j_2} v_{i_2} \ldots v_{i_{s-1}} e_{i_s} v_{i_s},$$

где $v_{i_0}, \ldots, v_{i_s}$ — вершины из V, а $e_{j_1}, \ldots, e_{j_s}$ — ребра из E, причем в каждой вида $v_{i_{k-1}} e_{j_k}$ и $e_{j_k} v_{j_k}$, $k = 1, \ldots, s$ вершина и ребро инцедентны, называется *путем.* Число ребер в последовательности называется *длиной пути* (в данном случае длина пути равна s). Путь, в котором все ребра различны, называется *цепью.* Путь, в котором все вершины различны, называется *простой цепью*

Вопрос. Может ли простая цепь содержать одинаковые ребра?

Путь называется *замкнутым,* если $v_{i_0} = v_{i_s}$. Замкнутый путь, в котором все ребра различны, называется циклом. Цикл, в котором все вершины различны, называется *простым циклом.*

Граф называется *связным* (также *односвязным*), если для любых двух различных его вершин существует путь, соединяющий эти вершины. *Расстоянием* между двумя вершинами называется длина кратчайшей цепи, соединяющей эти вершины.

Вопрос. Может ли цепь, соединяющая 2 вершины, длина которой равна расстоянию между вершинами, содержать одинаковые вершины?

Полным графом называется граф, в котором две различные вершины смежны (то есть соединены ребром).

Пример. Найти число ребер в полном графе с 7 вершинами.

По определению полного графа каждая вершина соединена с 6 другими вершинами. А значит, степень каждой вершины равна 6. Учитывая то, что сумма степеней вершин равна удвоенному числу ребер, получаем, что число ребер равно $6 \cdot 7/2 = 21$.

Рассуждая аналогичным образом можно показать, что число ребер в полном графе, содержащем n вершин, равно $\frac{n(n-1)}{2}$.

Дополнением графа $G = (V, E)$ называется граф $\overline{G} = (V, E')$, в котором вершины смежны в том и только том случае, когда они не смежны в графе G. Таким образом, при совмещении вершин графов G и $\overline{G}$ и объединении их ребер получаем полный граф ($G_1 = (V, E \cup E')$). Часто для решения задач гораздо удобнее перейти к рассуждениям про дополнение графа.