

ТЕОРИЯ ГРАФОВ

ДЕРЕВО РЕШЕНИЙ

Рассмотрим дерево с точки зрения иерархической структуры. **Деревом с корнем** называется дерево с одной выделенной вершиной. Именно эта выделенная вершина и является **корнем дерева**.

Каждую вершину дерева с корнем можно рассматривать как корень другого дерева, которое «растет» из него. Мы будем называть его **поддеревом** дерева T (рис. 20).

ТЕОРИЯ ГРАФОВ ДЕРЕВО РЕШЕНИЙ

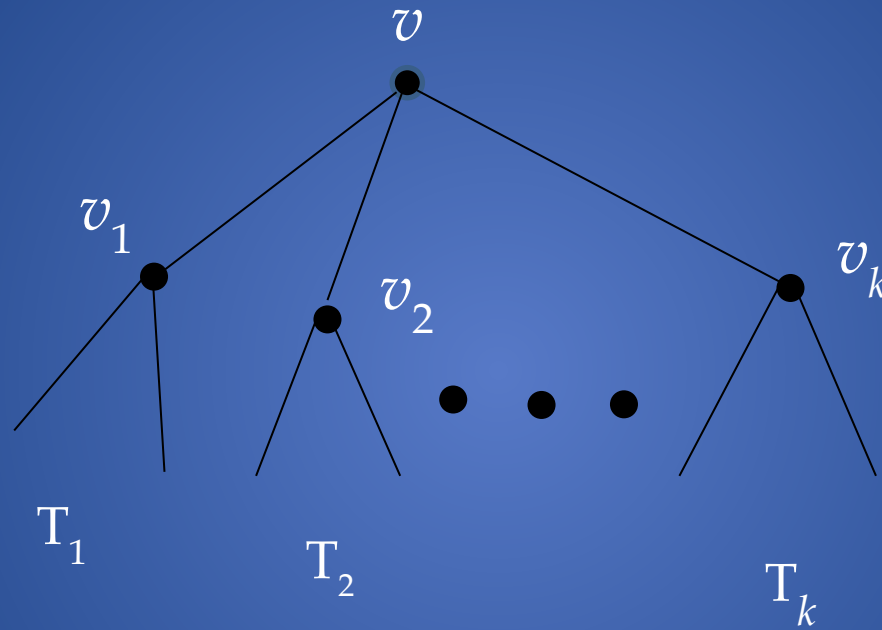


Рисунок 20

ТЕОРИЯ ГРАФОВ

ДЕРЕВО РЕШЕНИЙ

Вершины v_1, \dots, v_k графа T – это **сыновья корня** v (рис. 20). Мы изображаем такое дерево с корнем, расположенным наверху, и сыновьями, стоящими ниже, непосредственно под корнем.

Вершина на самом верху дерева – его корень (v), а вот те, которые находятся в самом низу дерева (и не имеют сыновей) принято называть **листьями**.

Остальные вершины, отличные от корня и листьев, называют **внутренними**.

Двоичные или **бинарные** деревья с корнем. Двоичное дерево отличается от остальных тем, что каждая его вершина имеет не более двух сыновей. В двоичном дереве с корнем вниз от каждой вершины идет не более двух ребер.

ТЕОРИЯ ГРАФОВ

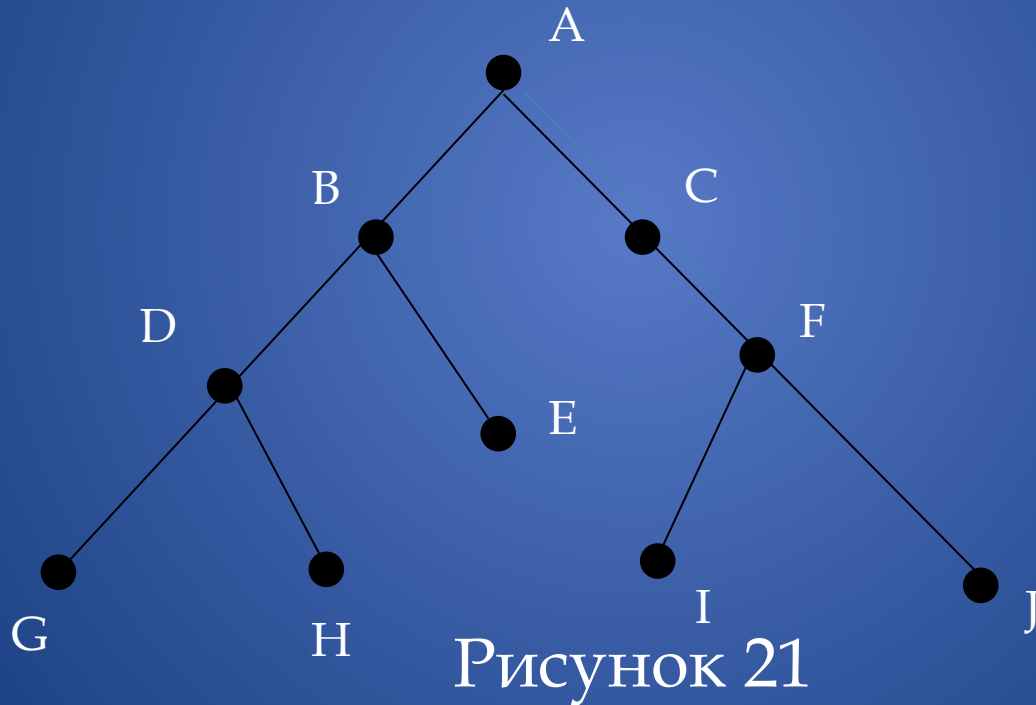
ДЕРЕВО РЕШЕНИЙ

Каждой вершине двоичного дерева с корнем соответствует не более, чем два поддерева, которые принято называть левым и правым поддеревьями этой вершины.

Если оказалось, что у какой-то вершины дерева отсутствует потомок слева, то ее левое поддерево называют нулевым деревом (т.е. нулевое дерево – это дерево без единой вершины). Аналогично, если у вершины отсутствует правый потомок, то ее правое поддерево будет нулевым.

ТЕОРИЯ ГРАФОВ ДЕРЕВО РЕШЕНИЙ

Пример. Пусть T – двоичное дерево с корнем, изображенное на рис. 21.



ТЕОРИЯ ГРАФОВ ДЕРЕВО РЕШЕНИЙ

Определите

а) корень T ;

б) корень левого поддеревья вершины B ;

в) листья T ;

г) сыновей вершины C .

Нарисуйте двоичное дерево с корнем T' , полученное из T перестановкой левых и правых поддеревьев у каждой вершины.

ТЕОРИЯ ГРАФОВ

ДЕРЕВО РЕШЕНИЙ

Дерево решений – используется в области анализа данных для прогнозных моделей. Структура дерева представляет собой следующее: «листья» и «ветки». На ребрах («ветках») дерева решения записаны атрибуты, от которых зависит целевая функция, в «листьях» записаны значения целевой функции, а в остальных узлах – атрибуты, по которым различаются случаи. Чтобы классифицировать новый случай, надо спуститься по дереву до листа и выдать соответствующее значение. Цель состоит в том, чтобы создать модель, которая предсказывает значение целевой переменной на основе нескольких переменных на входе.

ТЕОРИЯ ГРАФОВ

ДЕРЕВО РЕШЕНИЙ



ТЕОРИЯ ГРАФОВ ДЕРЕВО РЕШЕНИЙ

Предпринимать
научные исследования
и разработки?

Успешные научные
исследования и разработки
собственными силами

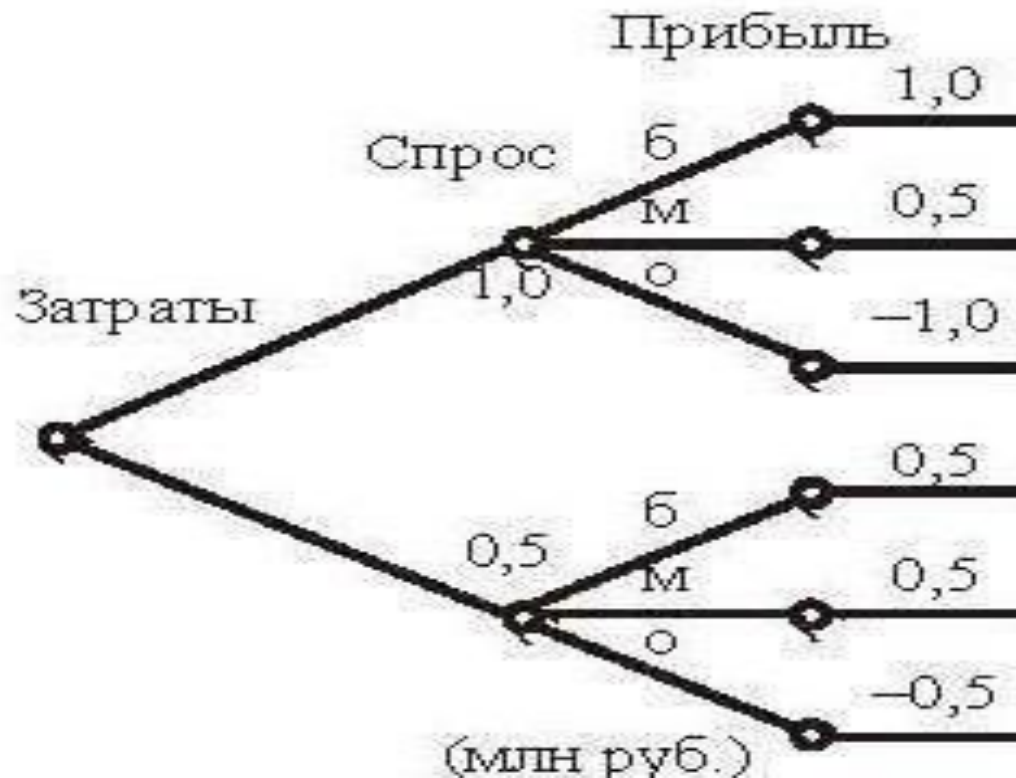


Организовывать
массовое производство?

Прибыль



ТЕОРИЯ ГРАФОВ ДЕРЕВО РЕШЕНИЙ



ТЕОРИЯ ГРАФОВ ДЕРЕВО РЕШЕНИЙ

Метод поиска с деревом решений

состоит в разбиении начальной задачи P_0 на некоторое число подзадач P_1, P_2, \dots, P_k (в целом представляющих всю задачу P_0) с последующей попыткой **разрешить** каждую из этих подзадач.

ТЕОРИЯ ГРАФОВ ДЕРЕВО РЕШЕНИЙ

Выражение **разрешить** понимаем так:

- найти оптимальное решение;
- показать, что значение оптимального решения хуже, чем для полученного до этого наилучшего решения;
- показать, что подзадача не является допустимой.

ТЕОРИЯ ГРАФОВ ДЕРЕВО РЕШЕНИЙ

Это разбиение описывается деревом (рис. 22),
вершины изображают подзадачи.

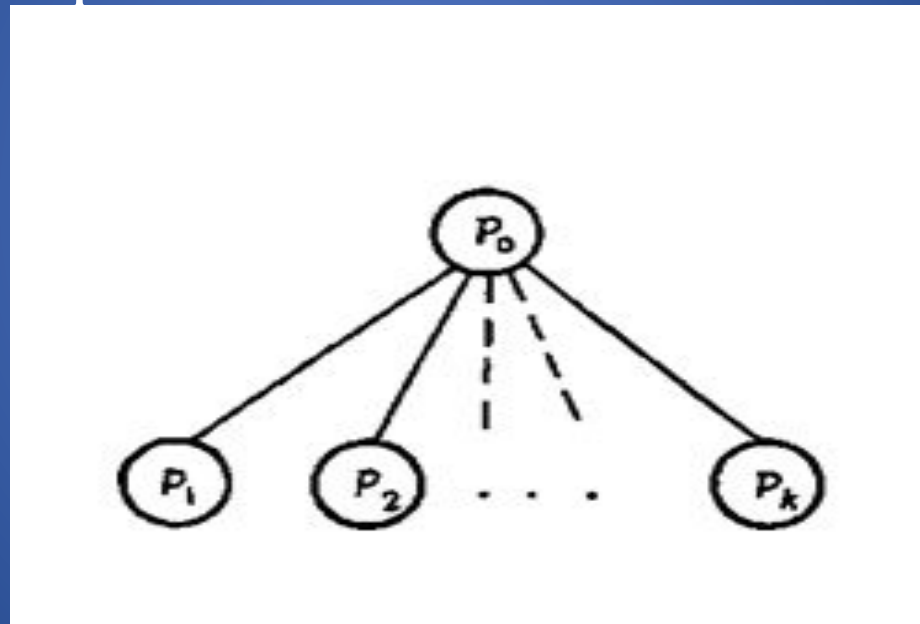


Рисунок 22

ТЕОРИЯ ГРАФОВ ДЕРЕВО РЕШЕНИЙ

Может оказаться, что подзадачу P_i нельзя разрешить, и эта подзадача сама разбивается на новые подзадачи $P_{i1}, P_{i2}, \dots, P_{ir}$ (рис. 23)

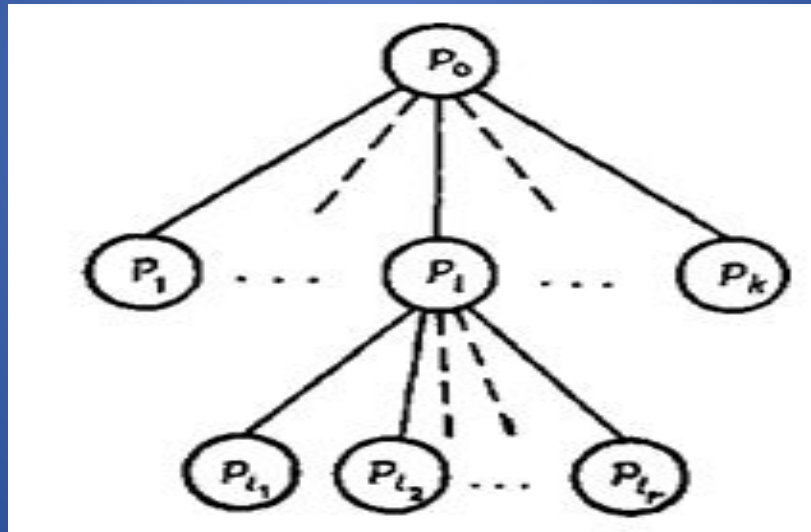


Рисунок 23

ТЕОРИЯ ГРАФОВ ДЕРЕВО РЕШЕНИЙ

Это разбиение, называемое **ветвлением**, повторяется для каждой подзадачи, которая не может быть разрешена.

На любом этапе полное множество подзадач, требующих разрешения, представляется множеством концевых вершин (т.е. вершин степени 1) всех цепей, выходящих из корня дерева решений. Эти вершины называются **висячими**, и на рис. 23 это $P_{1'} \dots, P_{i-1'}$ $P_{i1'} \dots, P_{ir'}$
 $P_{i+1'} \dots, P_k$

ТЕОРИЯ ГРАФОВ ДЕРЕВО РЕШЕНИЙ

Если поиск исчерпан, то очевидно, что множество подзадач, на которые разбита задача, должно представлять всю задачу. Таким образом, если задача P_i разбита на r подзадач P_{i1}, \dots, P_{ir} , то

$$\{P_{i1}\} \cup \{P_{i2}\} \cup \dots \cup \{P_{ir}\} = \{P_i\},$$

где $\{P\}$ обозначает множество всех допустимых решений задачи P .

ТЕОРИЯ ГРАФОВ ДЕРЕВО РЕШЕНИЙ

Любая подзадача, представляемая висячей вершиной и не поддающаяся разрешению, может быть в любой момент разбита на меньшие подзадачи.

Существует два основных типа поиска в зависимости от того, как выбирается следующая висячая вершина для продолжения процесса ветвления.

ТЕОРИЯ ГРАФОВ ДЕРЕВО РЕШЕНИЙ

Поиск в ширину (рис. 24)

При поиске по ширине ветвление происходит от уровня к уровню, так, что если на уровне 1 начальная задача P_0 разбивается на подзадачи P_1, P_2, \dots, P_k , то каждая из этих подзадач исследуется раньше, чем задачи уровня 2. Задачи уровня 1, которые не могут быть разрешены, разбиваются на подзадачи уровня 2, и опять все они исследуются до исследования какой-либо подзадачи, могущей возникнуть на уровне 3 и т.д.

ТЕОРИЯ ГРАФОВ ДЕРЕВО РЕШЕНИЙ

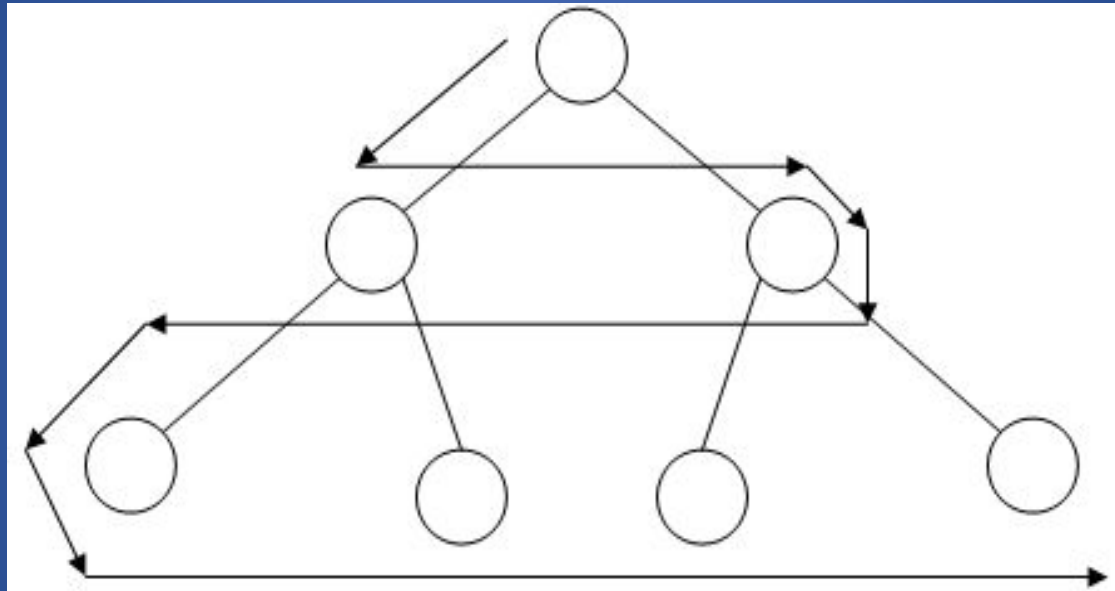


Рисунок 24

ТЕОРИЯ ГРАФОВ ДЕРЕВО РЕШЕНИЙ

Т.е. на каждой следующей итерации своей работы он расползается вширь по ребрам от вершин, до которых он дошёл к данной итерации. И так расползается он до тех пор, пока не побывает в каждой вершине текущей компоненты связности. Чаще всего поиск в ширину используется для нахождения кратчайшего пути от одной вершины до другой

ТЕОРИЯ ГРАФОВ ДЕРЕВО РЕШЕНИЙ

Поиск в глубину (рис. 25)

Этот алгоритм делает следующее:

поиск в глубину от начальной вершины не расползается вширь, а уходит вглубь по смежным рёбрам. Если из текущей вершины он не может дальше идти вглубь (все смежные вершины помечены), то он откатывается на предыдущую вершину (из которой он попал в текущую) и продолжает просматривать смежные вершины оттуда.

ТЕОРИЯ ГРАФОВ ДЕРЕВО РЕШЕНИЙ

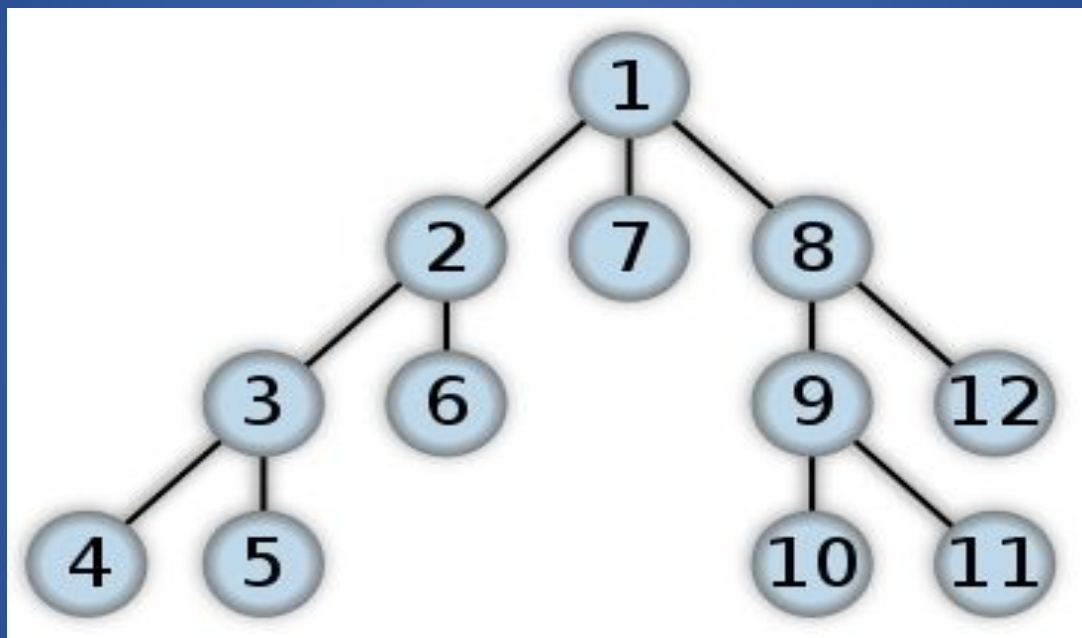


Рисунок 25

ТЕОРИЯ ГРАФОВ ДЕРЕВО РЕШЕНИЙ

Применение границ

Если задача P_0 подлежит решению как задача оптимизации, то безотносительно к типу поиска этот поиск завершается только тогда, когда будут разрешены все подзадачи, представляемые висячими вершинами. Для ускорения процесса разрешения для каждой из висячих вершин вычисляется нижние или верхние границы (для случаев минимизации или максимизации соответственно).

ТЕОРИЯ ГРАФОВ ДЕРЕВО РЕШЕНИЙ

Эти границы дают наименьшее (или наибольшее) возможное значение оптимального решения той подзадачи, которая соответствует рассматриваемой висячей вершине. Т.о. (для задачи минимизации), если окажется, что нижняя граница для вершины, соответствующей задаче P_i больше, чем величина наилучшего ответа, полученного ранее при поиске, то в P_i нет необходимости производить дальнейшее ветвление, т.к. в $\{P_i\}$ нет решения лучшего, чем текущий наилучший ответ.