

Тема 13. Запросы для поисковых систем с использованием логических выражений (задание 17)

При решении задач данного типа, необходимо повторить «Тема 7. Построение таблиц истинности логических выражений».

Пример №1 с решением

Ниже приведены запросы к поисковому серверу. Расположите номера запросов в порядке возрастания количества страниц, которые найдет поисковый сервер по каждому запросу. Для обозначения логической операции «ИЛИ» в запросе используется символ |, а для логической операции «И» – &.

- 1) принтеры & сканеры & продажа
- 2) принтеры & продажа
- 3) принтеры | продажа
- 4) принтеры | сканеры | продажа

Решение

Чем больше «&», тем меньшее количество страниц найдет поисковик, чем больше «|», тем большее.

- 1) меньше всего результатов выдаст запрос с наибольшими ограничениями — первый (нужны одновременно принтеры, сканеры и продажа)
- 2) на втором месте – второй запрос (одновременно принтеры и продажа)
- 3) далее – третий запрос (принтеры или продажа)
- 4) четвертый запрос дает наибольшее количество результатов (принтеры или сканеры или продажа)

Ответ – 1234 .

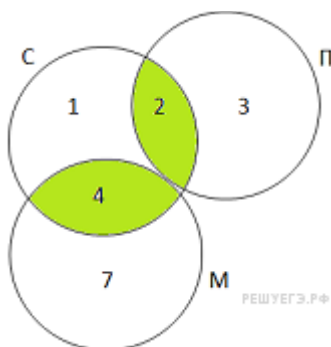
Пример №2 с решением

Некоторый сегмент сети Интернет состоит из 5000 сайтов. Поисковый сервер в автоматическом режиме составил таблицу ключевых слов для сайтов этого сегмента. Вот ее фрагмент:

Ключевое слово	Количество сайтов, для которых данное слово является ключевым
принтеры	400
сканеры	300
мониторы	500

Сколько сайтов будет найдено по запросу **(принтеры | мониторы) & сканеры** если по запросу **принтеры | сканеры** было найдено 600 сайтов, по запросу **принтеры | мониторы** – 900, а по запросу **сканеры | мониторы** – 750.

Решение



Для сокращения записи обозначим через С, П, М высказывания «ключевое слово на сайте – сканер» (соответственно принтер, монитор) и нарисуем эти области виде

диаграммы (кругов Эйлера). Заметим, что поскольку по запросу принтеры | мониторы было найдено 900 страниц, по запросам принтеры — 400, мониторы — 500, а $900 = 500 + 400$, области П и М не пересекаются. Интересующему нас запросу $(П | М) \& С$ соответствует объединение областей 4 и 2 («зеленая зона» на рисунке). Количество сайтов, удовлетворяющих запросу в области i , будем обозначать через N_i .

Из условия:

$$N_1 + N_4 + N_7 + N_2 = 750,$$

$$N_1 + N_4 + N_2 + N_3 = 600,$$

$$N_2 + N_3 = 400,$$

$$N_1 + N_2 + N_4 = 300,$$

$$N_4 + N_7 = 500.$$

Тогда из первого и пятого уравнений получаем, что $N_1 + N_2 = 250$, а из четвертого:

$$N_4 = 300 - 250 = 50.$$

Из второго и четвертого уравнений получаем, что $N_3 = 300$, а из третьего:

$$N_2 = 400 - 300 = 100.$$

Ответ: $N_2 + N_4 = 150$.

Пример №3 с решением

В языке запросов поискового сервера для обозначения логической операции «ИЛИ» используется символ «|», а для логической операции «И» — символ «&».

В таблице приведены запросы и количество найденных по ним страниц некоторого сегмента сети Интернет.

Запрос	Найдено страниц (в тысячах)
Мороз Солнце	3300
Солнце	2000
Мороз & Солнце	200

Какое количество страниц (в тысячах) будет найдено по запросу Мороз? Считается, что все запросы выполнялись практически одновременно, так что набор страниц, содержащих все искомые слова, не изменялся за время выполнения запросов.

Решение

По формуле включений и исключений имеем:

$$m(\text{Мороз} | \text{Солнце}) = m(\text{Мороз}) + m(\text{Солнце}) - m(\text{Мороз} \& \text{Солнце}) = m(\text{Мороз}) + 2000 - 200 = 3300. \Rightarrow m(\text{Мороз}) = 1500.$$

Ответ: 1500.

Задания для тренировки

1) В таблице приведены запросы к поисковому серверу. Расположите обозначения запросов в порядке убывания количества найденных страниц. В данной поисковой системе: символ & обозначает обязательное вхождение слов в одно предложение (логическое И); символ | обозначает поиск любого из заданных слов (логическое ИЛИ).

1) автобусы

2) (грузовики & легковые)|автобусы

3) (грузовики & легковые & мотоциклы) | автобусы

4) грузовики | легковые | мотоциклы | автобусы

2) В языке запросов поискового сервера для обозначения логической операции «ИЛИ» используется символ «|», а для логической операции «И» – символ «&». В таблице приведены запросы и количество найденных по ним страниц некоторого сегмента сети Интернет:

Запрос	Найдено страниц (в тысячах)
Чацкий & Молчалин & Фамусов	150
Чацкий & Фамусов	350
Чацкий & Молчалин	270

Компьютер печатает количество страниц (в тысячах), которое будет найдено по следующему запросу: *Чацкий & (Молчалин|Фамусов)* Укажите целое число, которое напечатает компьютер. Считается, что все запросы выполнялись практически одновременно, так что набор страниц, содержащих все искомые слова, не изменялся за время выполнения запросов.

3) В таблице приведены запросы к поисковому серверу. Расположите номера запросов в порядке убывания количества страниц, которые найдет поисковый сервер по каждому запросу. Для обозначения логической операции «ИЛИ» в запросе используется символ |, а для логической операции «И» – &.

- 1) гуси & утки
- 2) гуси & (утки | индюки)
- 3) гуси & утки & индюки
- 4) утки | индюки

4) Некоторый сегмент сети Интернет состоит из 1000 сайтов. Поисковый сервер в автоматическом режиме составил таблицу ключевых слов для сайтов этого сегмента. Вот ее фрагмент:

Ключевое слово	Количество сайтов, для которых данное слово является ключевым
сканер	200
принтер	250
монитор	450
принтер сканер	450
принтер & монитор	40
сканер & монитор	50

Сколько сайтов будет найдено по запросу

(принтер | сканер) & монитор

если по запросу **принтер | сканер** было найдено 450 сайтов, по запросу **принтер & монитор** – 40, а по запросу **сканер & монитор** – 50.

Ответы к заданиям для тренировки

- 1) 4231
- 2) 470
- 3) 4213
- 4) 90