

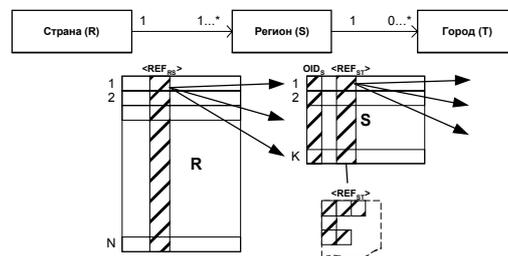
Математический метод оценки времени выполнения запросов к ООСУБД: предпосылки анализа и оценка алгоритма Forward Join

1. Особенности объектно-ориентированных СУБД

- объектные идентификаторы
- поддержка ассоциативных связей между экстендами объектов

```
select R.название_страны, R.S.название_региона
from R
where R.S.T.название_города = "Houston"
```

- поддержка иерархии наследования
- поддержка функций
- композиционные объекты



2. Предпосылки

Рассмотрим экстенды классов R и S, связанные ассоциативной связью.

Число объектов в R равно N, а число объектов в S равно K.

Каждый объект из R и S имеет свой уникальный идентификатор:

$$OID_R \quad OID_S$$

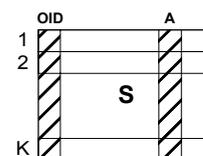
Каждый объект из R может иметь произвольное число ссылок на объекты S.

Домены ссылок за счет ссылочной целостности подчиняются следующему выражению:

$$D_{REF} \subseteq D_{OID}$$

Вероятность независимых в совокупности событий ссылки объекта R на объект S определяется как:

$$\eta_i \quad \forall i = 1..|D_{REF}|, \exists j = 1..|D_{OID}|: a_{REF_i} = a_{OID_j}$$



Алгоритм прямого соединения (Forward Join)

Лемма 1: Пусть соединение выполняется между экстендами R и S по алгоритму Forward Join. Пусть, также, выбор j-го объекта по условию F_S независим от выбора объекта по ссылке REF, тогда:

ПФСВ числа отобранных объектов из S:

$$\Omega(z) = [z^0]V_S(z) + \sum_{k=1}^K [z^k]V_S(z) \binom{K}{k}^{-1} \sum_{(j_1, \dots, j_k) \in Q_k} \prod_{j=j_1, \dots, j_k} (1 - (1 - [z^0]V_R(1 - \eta_j(1 - z))) (1 - z))$$

математическое ожидание числа отобранных объектов из S:

$$M(\omega) = \sum_{k=1}^K \left([z^k]V_S(z) \binom{K}{k}^{-1} \sum_{(j_1, \dots, j_k) \in Q_k} \sum_{j=j_1, \dots, j_k} (1 - [z^0]V_R(1 - \eta_j(1 - z))) \right)$$

$V_S(z)$ - ПФСВ числа выбранных по условию объектов из S

$V_R(z)$ - ПФСВ числа выбранных по условию объектов из R

Q_k , $|Q_k| = \binom{K}{k}$ - множество комбинаций по k элементов идентификаторов объектов из S

Следствие 1: Пусть вероятности ссылок равны между собой и домены значений ссылок совпадают:

$$\forall j \in 1..K: \eta_j = \eta = \frac{1}{|D_{REF}|} = \frac{1}{|D_{OID}|} = \frac{1}{K}$$

ПФСВ числа отобранных объектов из S:

$$\Omega(z) = \left(1 - \left(1 - \left(1 - \frac{1}{K} \right)^{N_R} \right) (1 - z) \right)^{K_S}$$

математическое ожидание числа отобранных объектов из S:

$$M(\omega) = K_S \left(1 - \left(1 - \frac{1}{K} \right)^{N_R} \right)$$

N_R и K_S - число отобранных объектов из R и S по условиям

F_R и F_S

Теорема 1: Предположим, что выполнены предпосылки леммы 1 и соединение выполняется по физическому и логическим указателям, тогда:

ПЛС времени выполнения соединения:

$$\Xi(s) = \Psi(\Theta(s)\Omega(s))$$

математическое ожидание времени выполнения соединения:

$$M = M(\xi)(M(\theta) + M(\omega))$$

$\Theta(s)$ - ПЛС времени поиска физического указателя

$M(\theta)$ - математическое ожидание времени поиска физического указателя

$\Omega(s)$ - ПЛС времени доступа к объекту

$\Psi(z) = \prod_{j=1}^{|D_{OID}|} V_R(1 - \eta_j(1 - z))$ - ПФСВ числа ссылок на объекты из S

$M(\omega)$ - математическое ожидание времени доступа к объекту