CSE 562: Practice problems for test #2

Database

You are given the following relational schema:

Movie(Title,Dname)
Cast(underlineTitle,Aname,Role)
Actor(Aname,Addr)
Award(Aname,Year,Title)

Keys are underlined.

Problem 1

Consider the following query:

$\mathbf{Q_1} \equiv Movie \bowtie Cast \bowtie Actor \bowtie Award.$

You are given the following cardinality estimates:

$T(Movie) = V(Movie, Title) = 10000$
$V(Movie, Dname) = 1000$
$T(Cast) = 50000$
$V(Cast, Title) = 10000$
$T(Actor) = V(Cast, Aname) = V(Actor, Aname) = 4000$
$T(Award) = 500$
$V(Award, Aname) = 200$
$V(Award, Title) = 100$

Calculate the plan of least estimated cost using the dynamic programming algorithm.

Solution

Relations	*Cost*	*Best plan*
Movie, Cast	*0*	$Movie \bowtie Cast$
Cast, Actor	*0*	$Actor \bowtie Cast$
Actor, Award	*0*	$Award \bowtie Actor$
Movie, Award	*0*	$Award \bowtie Movie$
Cast, Award	*0*	$Award \bowtie Cast$
Movie, Cast, Actor	*50000*	$(Movie \bowtie Cast) \bowtie Actor$
Movie, Cast, Award	*500*	$(Award \bowtie Movie) \bowtie Cast$
Cast, Actor, Award	*500*	$(Award \bowtie Actor) \bowtie Cast$
Movie, Actor, Award	*500*	$(Award \bowtie Actor) \bowtie Movie$
Movie, Cast, Actor, Award	*501*	$((Award \bowtie Movie) \bowtie Cast) \bowtie Actor$

For multiple plans with the same cost only one is selected above.

Problem 2

Consider the following query:

$$\mathbf{Q_2} \equiv \sigma_{Cast.Role=''star''}(Movie \bowtie Cast).$$

Explain how indexes can be used to make the evaluation of the query $\mathbf{Q_2}$ more efficient. List all such indexes and specify whether they are primary or secondary, sparse or dense. Describe the evaluation plans that use the indexes.

Solution

1. Dense secondary index on $Cast.Role$.
2. Dense secondary index on $Cast.Title$.

Evaluation plans:

1. $HashJoin(TableScan(Movie), IndexScan(Cast, Cast.Role =''\ star''))$
2. $Filter(IndexJoin(TableScan(Movie), IndexScan(Cast, Cast.Title = Movie.Title)), Cast.Role =''\ star'')$

There are other possible evaluation plans but they requite joining the relation in a different order which is suboptimal ($Movie$ is smaller than $Cast$).