

## Databases 2004 – Paper 5 Question 8 (GMB)

(a) (i)

```
SELECT name
FROM    MovieStar, StarsIn
WHERE   gender = 'M' AND
        name    = starName AND
        movieTitle = 'Red Squirrel';
```

The reader interested in Spanish movies will know that the correct answer to this query is Nancho Novo and Karra Elejalde.

(ii)

```
SELECT m2.title
FROM    Movie m1, Movie m2
WHERE   m1.title='Titanic' AND
        m2.length>m1.length;
```

(b) (i) It returns the names of producers of Nancho Novo's films.

(ii)

```
SELECT name
FROM    Producer, Movie, StarsIn
WHERE   cert#=prodC# AND
        title=MovieTitle AND
        year=movieYear AND
        starName='Nancho Novo';
```

(iii) The difference is when a producer has produced more than one movie for a star (which is the case of Nancho Novo, who has starred in two movies produced by Julio Medem). In the nested case we do not generate duplicates because we iterate over the **Producer** relation, which is a set. In the non-nested version we can generate an answer with repetitions. Thus to make them equivalent, the non-nested query requires a **DISTINCT** to be added to the **SELECT** clause.

(c) The query **EXISTS R** can be replaced with the query **1 IN (SELECT 1 FROM R)**. The trick is to notice that the inner **SELECT** will return an empty bag if **R** is empty and a bag of repeated 1s if not. We can then use the **IN** from the previous section to complete our encoding.