The Center for Advanced Computer Studies
University of Louisiana at Lafayette
CMPS 566
Term Test

Date: March 23, 2005 | Instructor: Dr. V. Raghavan

Time: 3:30 - 4:45 p.m. | Total Marks: 75

PART A (20 Marks)

NOTE: There are **five** parts. Answer *any 4*.

Q1. Outlier Analysis vs. Clustering.

Q2. Subjective measures of pattern interestingness.

Q3. Starnet Query Model

Q4. Operation-derived hierarchies.

Q5. Meta patterns.

PART B (55 marks)

Answer *all* questions.

Q6.

ID	Fuel	Cyl	Power	Tran	Mileage
T1	efi	4	high	manu	med
T2	efi	6	high	manu	med
T3	2-bbl	6	high	auto	low
T4	efi	6	med	manu	med
T5	efi	4	high	manu	high
T6	2-bbl	4	med	manu	high
T7	efi	6	high	auto	low
T8	efi	6	med	manu	low
T9	efi	4	med	auto	med
T10	2-bbl	4	high	manu	high
T11	efi	4	med	manu	med
T12	efi	4	high	auto	high
T13	2-bbl	4	low	manu	high
T14	efi	6	high	auto	med

Table 1

a) Write a DMQL query to find the *comparisons* of cars according to 'mileage.'

The class of "low mileage" is to be compared to the class of "high mileage."

b) Write a DMQL query to predict if a car's mileage class is "low", based on the car attributes of 'cyl' and 'Tran'.

Q7. Use data from A6.

a) Propose a concept hierarchy *and* specify its type for

(i) attribute 'Fuel', and

(ii) attribute 'Cyl'

Note:

Don't restrict your thinking to just the domain of values that are given for these two attributes in the table of Q6.

b) Using your answer to part a), for hierarchies on 'Fuel' and 'Cyl', and assuming that other attributes only have just two levels, determine how many distinct views will be associated with the data warehouse involving the 5 attributes as its dimensions.

c) Write an SQL query, assuming the data are stored in a relational DBMS and the fact table is named 'car-types', that lists number of cars produced, by 'mileage', for the generalized tuple cyl = "6" and Trans = "manu".

d) With respect to the generalized tuple Cyl = "4" and Trans = "manu", determine the d weights to compare the 'mileage' classes "med" and "high".

e) Compute a (bi-directional) quantitative description rule that connects class of mileage = “med” to attributes ’Cyl’ and ’ Tran’.

Q8. a) For the generalized table involving only the attributes ’Power’ and ’Mileage’, construct the context that would be obtained from the table of Q6. The context should be displayed in the *flattened* form.

b) What is the *naive* prediction with respect to mileage = “low” and mileage = “med”? Provide the prediction accuracy.

c) How much does the prediction accuracy improve, when attribute 'Power' is known? What are the predictions?

d) Using the context given in part a), illustrate

i) a feasible set of features (attribute values), *B*, such that $(\alpha(B), B)$ is a formal concept.

(ii) a feasible set of tuples, *A*, such that $(A, \beta(A))$ is a formal concept.

(iii) Identify two formal concepts (A_1, B_1) and (A_2, B_2) such that one is superconcept of the other.

(iv) Identify formal concept $((A_3, B_3)$ that is not comparable to the ones in part (iii).