

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

Date: March 15, 2004

Instructor: Dr. V. Raghavan

Time: 3:30 - 4:45

Total Marks: 75

PART A (20 Marks)

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

Q1. Characterization vs. Clustering.

Q2. Pattern Validity.

Q3. Fact Constellation Schema.

Q4. Numerosity Reduction.

Q5. Algebraic Facts.

PART B (55 marks)

Answer *all* questions.

Q6.

Customer ID	Height	Hair	Eyes	Credit Rating
e_1	short	dark	blue	B
e_2	tall	dark	blue	B
e_3	tall	dark	brown	B
e_4	tall	red	blue	A
e_5	short	blond	blue	A
e_6	tall	blond	brown	B
e_7	tall	blond	blue	A
e_8	short	blond	brown	B

Table 1

a) Write a DMQL query to find the *characteristics* of students who have a credit rating of B.

b) Write a DMQL query to *compare* (discriminate) customers having credit rating A with customers having credit rating B.

Q7. Use data from Q6.

a) Find t-weights to characterize concepts corresponding to credit-rating = 'A' and credit-rating = 'B'.

b) Let us divide concept credit-rating = 'A' into subconcepts $A^+ = \{e_1, e_2, e_3\}$ and $A^- = \{e_6, e_8\}$. *Discriminate* by computing d-weights, among the two concepts using the generalized tuple Height = "tall".

c) Compute *information gain* associated with attribute Hair, with respect to the concepts credit-rating = 'A' and credit-rating = 'B'.

d) Show a flow chart of the algorithm to perform attribute subset selection that uses information gain as the criterion and step-wise backward elimination strategy.

e) What is the *naive* prediction with respect to credit-rating = 'A⁺' and credit-rating = 'A⁻'. Provide the prediction accuracy.

f) By how much does the predictive accuracy improve, when the attribute Hair is known?
What are the predictions?

Q8. (a) (Formal concepts) Using the Table of Q6 and restricting it to attributes Hair and Eyes, obtain the concept lattice. Label each node (corresponding to feasible concepts) with its intension and extension.

b) Briefly describe two ways in which the creation of concept lattice can be useful in understanding the patterns in data.