# Final Exam: Th. Mar.23, 1.30-3.30pm

**Intro AI ICS 171 Winter 2006**
Instructor: Max Welling

- *This exam is closed book*
- *You may use a calculator*
- *Always explain your answer*
- *Spend your time wisely: get a shot at each question*
- *You can get a total of 70 points*
- *Write with a PEN.*
- *Good Luck !*

1.(10 pts) **Miscellaneous Questions**
Answer the following questions with True or False.

a.(2 pts) The task environment of a taxi-driving agent is episodic.
a) answer: False: it is sequential.
b.(2 pts) $A^*$ search has polynomial time complexity for all search problems.
b) answer: False: it has exponential space complexity
c.(2 pts) The alpha-beta pruning algorithm is an optimal search strategy.
c) answer: True.
d.(2 pts) The time complexity for solving a constraint satisfaction problem on a tree-structured constraint graph is linear in the number of variables of the problem.
d) answer: True.
e.(2 pts) A genetic algorithm is a local search algorithm.
e) answer: True.

2.(10 pts) **Propositional Logic**
Consider the following knowledge base: $KB_0 = \neg A \vee \neg B \vee C$ and the sentence $\alpha = \neg A \vee \neg B$

a.(2 pts) Use De Morgan's law to rewrite $\neg\alpha = \neg(\neg A \vee \neg B)$
a) answer: $\neg(\neg A \vee \neg B) = A \wedge B$.
b.(2 pts) Consider the updated KB: $KB_1 = KB_0 \wedge \neg\alpha$. Is the $KB_1$ in Horn form? Are all clauses in $KB_1$ definite clauses?
b) answer: Yes and yes. There are 3 clauses, each clause contains exactly 1 positive literal.
c.(4 pts) Use resolution to prove $C = true$.
c) answer: First Use $A$ and $\neg A \vee \neg B \vee C$ to conclude $\neg B \vee C$. Then use that and $B$ to conclude $C$.
d.(2 pts) Is $\alpha$ entailed by $KB_0$, i.e. $KB_0 \models \alpha$?
d) answer: No: we have just shown that $KB_1 = KB_0 \wedge \neg\alpha$ has a solution and is therefore not unsatisfiable.

3.(10 pts) **First Order Logic**

Consider the following relations: $F(x)$ is true when $x$ is female, $M(x)$ is true when $x$ is male, $D(x)$ is true when $x$ lives in Disneyland and $L(x,y)$ is true when $x$ likes $y$. Translate the following sentences into first order logic:

a.(5 pts) There is at least one male and female, both living in Disneyland, that like each other.

a) answer: $\exists x, y \; M(x) \wedge F(y) \wedge D(x,y) \wedge L(x,y) \wedge L(y,x)$.

b.(5 pts) All males and females living in Disneyland like each other.

b) answer: $\forall x, y \; M(x) \wedge F(y) \wedge D(x,y) \Longrightarrow L(x,y) \wedge L(y,x)$.

4.(20 pts) **Learning**

Consider a general binary classification problem, where we have $D$ discrete valued features and one binary class label. Consider a dataset with $N$ training cases.

A noise-free training set is a dataset for which the following holds: for any subset of data-cases that have exactly the same attribute values, the class labels are the same as well.

a.(4 pts) For an arbitrary noise-free training dataset, does there exist a decision tree that has 0% training error? (explain)

a) answer: Yes, consider the tree that first splits on the first attribute $A$, then for every value of $A$, it splits on $B$ etc. until the last attribute. At the last attribute all data-cases have the same attribute values for all attributes and hence must have the same class label. Hence, no training error.

b.(4 pts) Same as in a) but now for a perceptron classifier?

b) answer: No, the perceptron can only fit a linear surface and is not able to solve some very simple problems. For instance: XOR.

c.(4 pts) Now, lets consider a noisy dataset. This means, there is at least one pair of data-cases that have the same attribute values, but different class labels. Does there exist a decision tree that achieves 0% training error on a noisy dataset?

c) answer: No: by the time you have split on attributes there is a leaf node that has at least one positive and one negative example in it. Since you have to make a decision at the leaf, one of them is doomed to be wrongly classified.

d.(4 pts) Consider the XOR problem: $(A, B; Y) = \{(0,0;0), (0,1;1), (1,0;1), (1,1;0)\}$ where we have 4 data-cases and 2 attributes. Draw a decision tree that achieves 0% training error.

d)answer First split on A and then on B or reversed.

e.(4 pts) Determine the information gain after splitting on $A$, where $A$ is the first attribute.

e) answer: The information gain is zero.

5.(20 pts) **Probability**

John likes recognizing cars. He classifies cars into one of 3 classes: Car=[Ferrari,RollsRoyce,Other]. John observes 3 features: Color=[red,other], Speed=[fast,slow] and Weight=[heavy,light]. We will assume that the features Color, Speed and Weight are all conditionally

independent given Car. Furthermore, it is given that:
P(Color=red|Car=Ferrari)=0.5 ,
P(Speed=high|Car=Ferrari)=0.5,
P(Weight=light|Car=Ferrari)=0.9,
P(Color=red|Car=RollsRoyce)=0,
P(Speed=high|Car=RollsRoyce)=0.1,
P(Weight=light|Car=RollsRoyce)=0,
P(Color=red|Car=other)=0.1,
P(Speed=high|Car=other)=0.4,
P(Weight=light|Car=other)=0.5,
P(Car=Ferrari)=0.01 (John lives in Newport Beach),
P(Car=RollsRoyce)=0.01.

a.(4 pts) Use conditional independence to express P(Color,Speed,Weight,Car) as function of P(Color|Car), P(Speed|Car), P(Weight|Car) and P(Car).

a)answer: P(Color,Speed,Weight,Car)=P(Color|Car)P(Speed|Car)P(Weight|Car)P(Car).

b.(4 pts) How many entries does the joint probability table have for P(Color,Speed,Weight,Car)?

b)answer: 24 entries.

c.(4pts) Using the available information, compute the probability of:
P(Color=red,Weight=light,Speed=high,Car=Ferrari) and of
P(Color=other,Weight=heavy,Speed=low,Car=RollsRoyce).

c)answer: P(Color=red,Weight=light,Speed=high,Car=Ferrari)=0.5x0.9x0.5x0.01=0.00225
P(Color=other,Weight=heavy,Speed=low,Car=RollsRoyce)=1*1*0.9*0.01=0.009

d.(4 pts) Use Bayes rule to express P(Car|Color,Speed,Weight) in terms of the joint probability table. Note: this expression may involve terms where you need to sum over all possible values of certain variables.

d)answer: P(Car|Color,Speed,Weight)=P(Color,Speed,Weight,Car)/P(Color,Speed,Weight).
The denominator can be be expressed a sum over all values for Car of the joint probability table.

e.(4 pts) John sees a car and observes: Color=red, Speed=high, Weight=light. Compute the probability that the car is a Ferrari.

e) answer: Applying the equation in c: $0.0025/(0.0025 + 0 + 0.0196) = 0.113$