



UNIVERSITY OF COLOMBO, SRI LANKA

UNIVERSITY OF COLOMBO SCHOOL OF COMPUTING
(Successor to the Institute of Computer Technology (ICT), University of Colombo))

DEGREE OF BACHELOR OF INFORMATION TECHNOLOGY (EXTERNAL)
Academic Year 2008/2009 – 3rd Year Examination – Semester 5

IT5303: Intelligent Systems
Structured Question Paper
17th May, 2009
(TWOHOURS)

To be completed by the candidate

BIT Examination Index No: _____

Important Instructions:

- The duration of the paper is **2 (Two) hours**.
- The medium of instruction and questions is English.
- This paper has **4 questions** and **16 pages**.
- **Answer all 4 questions. Each of the four questions carries 25 marks.**
- **Write your answers** in English using the space provided **in this question paper**.
- Do not tear off any part of this answer book. Under no circumstances may this book, used or unused, be removed from the Examination Hall by a candidate.
- Note that questions appear on both sides of the paper.
If a page is not printed, please inform the supervisor immediately.

Questions Answered

Indicate by a cross (X), (e.g.) the question numbers of the questions answered.

To be completed by the candidate by marking a cross (X)	1	2	3	4
To be completed by the examiners:				

- 1) (a) What are the four (4) factors which determine the “rationality” of an intelligent agent?

(04 marks)

ANSWER IN THIS BOX

- 1) The performance measure that defines the criterion of success
- 2) The agent’s prior knowledge of the environment
- 3) The actions which the agent can perform
- 4) The agent’s percept sequence to date

- (b) Define “Goal-based Agent” and “Utility-based Agent”. What is the main difference between them?

(07 marks)

ANSWER IN THIS BOX

*Goal-based agent is an intelligent agent that always makes decisions on the basis whether it can achieve its goal by taking a particular decision or not.

*Utility-based agent is an intelligent agent that always tries to optimize a utility when decisions are made.

*Goal-based agents have descriptions about the **goals** they are supposed to achieve and utility-based agents define the utilities that they are supposed to optimize as **utility functions**.

- (d) Consider the context free grammar given below and derive two (2) possible parse trees for the sentence “*The teacher punished a student with long hair*”.

$S \rightarrow NP VP$

$NP \rightarrow DET NP$

$NP \rightarrow ADJ N$

$NP \rightarrow NP PP$

$VP \rightarrow V NP$

$VP \rightarrow V NP PP$

$PP \rightarrow PREP NP$

$DET \rightarrow a | the$

$N \rightarrow teacher | student | hair$

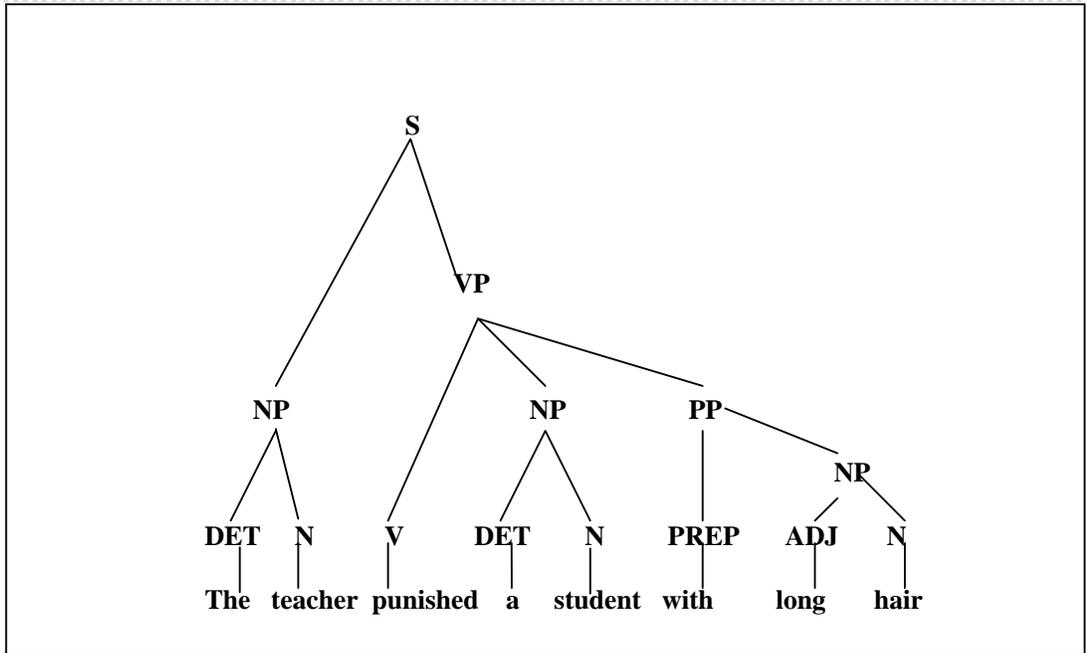
$V \rightarrow punished$

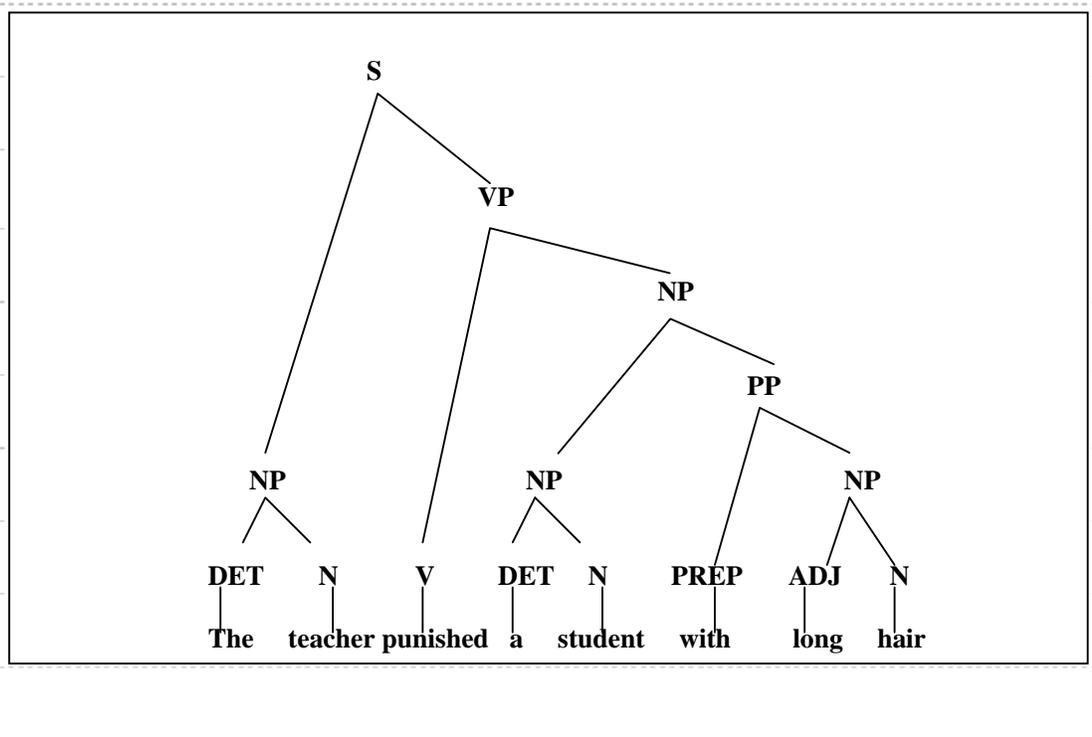
$ADJ \rightarrow long$

$PREP \rightarrow with$

(06 marks)

ANSWER IN THIS BOX





- (e) What steps can be taken to select the acceptable parse tree from the two parse trees you have given in part (d) ?

(02 marks)

ANSWER IN THIS BOX

- *The grammar can be disambiguated by the conversion of the CFG to a Probabilistic CFG by obtaining statistics from a sufficiently large corpus.
- *Use semantics or pragmatics to select the acceptable interpretation given the context.

- 2) (a) Give the corresponding Prolog statements for the following first order predicate logic statements:

love(saman, rani)

mother(sita, kamal)

$\forall x, y \text{ mother}(x, y) \rightarrow \text{love}(x, y)$

$\forall x, y \text{ teach}(x, y) \rightarrow \text{student}(y, x)$

(04 marks)

ANSWER IN THIS BOX

```
love (saman, rani).  
mother_of(sita, kamal)  
love(X,Y) :- mother_of(X,Y).  
student(Y,X) :- teach(X,Y)
```


ANSWER IN THIS BOX

Q (Start State)
Q -> U, S
S -> R
U ->P, R
R ->T (Goal State)

(d) Define the term “heuristic”.

(04 marks)

ANSWER IN THIS BOX

Heuristic is a rule of thumb that gives a rough estimate of the cost from a node to the goal state. Usually, it does not *overestimate* the cost.

(d) The table below gives the straight line distances to city P from other cities.

City	Straight Line Distance
P	0
Q	16,000
R	12,480
S	20,000
T	12,800
U	11,520

- (i) Define a suitable cost function to perform A* search.
- (ii) Find the optimal path from city S to city P using A* search.

(06 marks)

- (i) $f(n)=g(n)+h(n)$, where $g(n)$ is the distance from the start node to node n and $h(n)$ is the estimated cost from node n to the goal state.
- (ii) S, R, P

- 3) (a) Use the description given below to complete the script given in the answer box.
“A student has come to a teacher’s place to study. The teacher starts the lesson when he/she is ready. When the teacher finishes the lesson he/she will give some homework”

(06 marks)

ANSWER IN THIS BOX

Entry Conditions:

Teacher present, Student ready to learn.

Results:

Student learns, teacher gets tired, student gets tired.

Props:

Desks, books, pen

Roles:

Student, teacher

Scenes:

Student arrives, teacher teaches, student learns

Tracks:

Student has to do homework

- (b) Express the following first order predicate logic statements in natural language.
- (i) $\exists x(dog(x) \rightarrow \sim bark(x)) \wedge \forall x(dog(x) \wedge (male(x) \vee female(x)))$
 - (ii) $\exists x \exists y((love(x, y) \rightarrow love(y, x)) \vee (love(x, y) \rightarrow \sim love(y, x)))$

(04 marks)

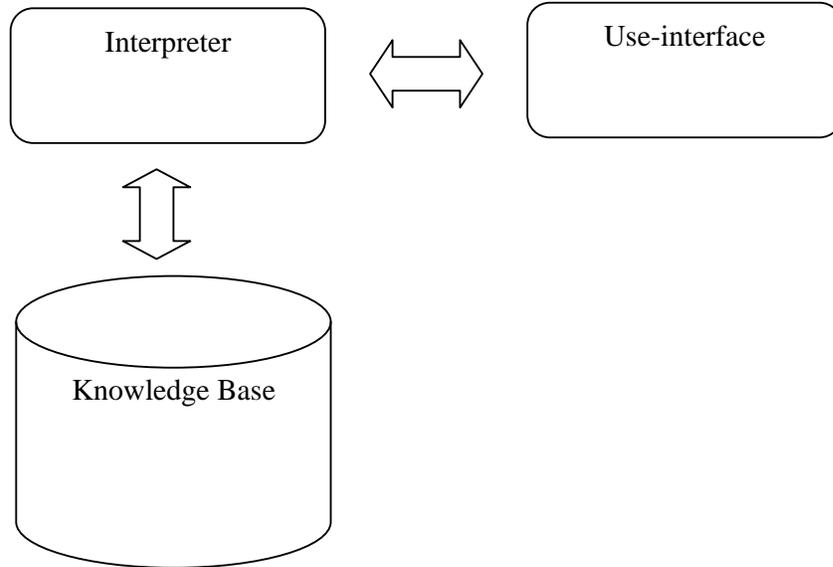
ANSWER IN THIS BOX

- (i) Although some dogs do not bark, they are either male or female.
- (ii) When someone loves another person, that person may or may not love him.

(c) Sketch a named diagram showing the main components of an expert system.

(03 marks)

ANSWER IN THIS BOX



- (d) What is meant by the term “knowledge acquisition bottleneck” in the context of expert system development?

(02 marks)

ANSWER IN THIS BOX

Knowledge should be acquired from an expert to build the knowledge base. This process is time consuming.
In order to acquire knowledge from the expert it is necessary to interview the expert by asking appropriate questions which are hard to develop.
Only a small portion of the expert knowledge can be extracted after spending a substantial amount of time.

(e) What are the main differences between rule-based and data-driven expert systems?

(04 marks)

ANSWER IN THIS BOX

Rule-based Systems:

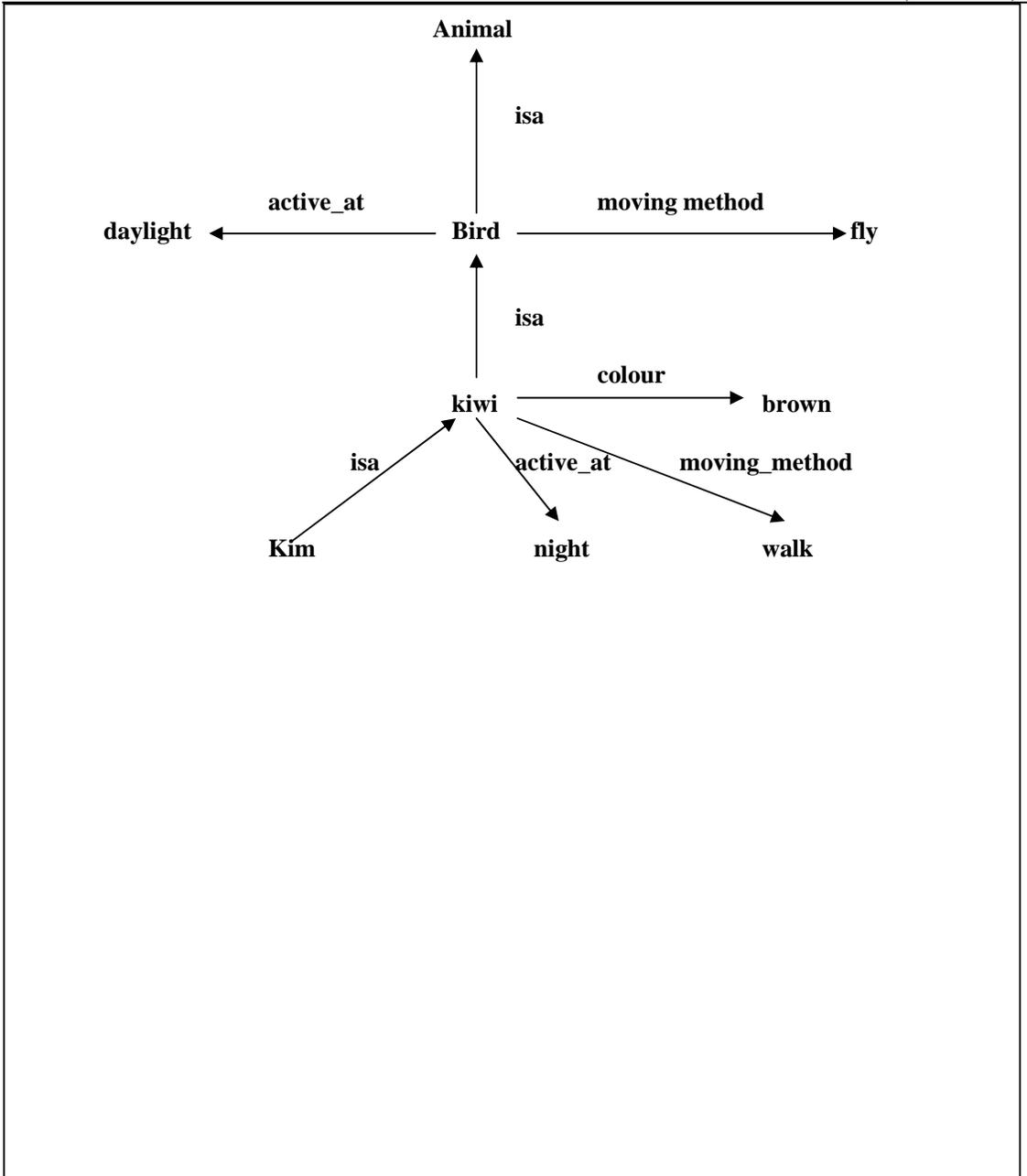
- Use expert knowledge
- Development is time consuming
- Produce a detailed justification for each decision
- Produce only one decision (alternatives are not given)

Data-driven Systems:

- Use large amount of past (data)
- Development is relatively fast
- Justifications are not given for decisions
- Alternatives decisions are given with confidence score (ranking)

- (f) Draw a semantic network to represent the following scenario.
 “A bird is a kind of animal. The normal moving method of birds is flying. Birds are usually active at daylight. A kiwi is a bird whose moving method is walking and colour is brown. A kiwi is active at night. Kim is a kiwi”

(06 Marks)



- 4) (a) Compare and contrast the functions of the human brain and artificial neural networks.

(06 marks)

ANSWER IN THIS BOX

Similarities:

ANN

Human Brain

- Ability to gather knowledge
- Knowledge acquisition through learning and training
- Knowledge dissemination
- Ability to both supervised and unsupervised learning/training
- Ability to behave in changing environments

Differences:

ANN

Human Brain

- | | |
|---|--|
| <ul style="list-style-type: none">▪ Consistent behaviour▪ No decay of memory▪ Task specific | <ul style="list-style-type: none">Consistency not guaranteedMemory decay can happenNon-task specific |
|---|--|

(b) Briefly explain whether the “*perceptron convergence procedure*” can be used to solve the XOR problem.

ANSWER IN THIS BOX

As the XOR problem is not a linearly separable problem, perceptron convergence procedure does not converge to a weight vector.

- (c) Training Data, Validation Data and Test Data are the typical data sets used in training artificial neural networks. Briefly describe the role of each of these data sets.

(06 marks)

ANSWER IN THIS BOX

Training Data: Training data set is used to train ANNs. It is assumed that training data represent the problem domain adequately.

Validation Data: Validation data set is used to verify whether the ANN has been trained properly and is not overfitted to training data.

Test Data: The actual performance of the trained ANN is assessed with respect to the test data set.

(d) An artificial neural network has been trained to recognize characters of the English alphabet.

(i) Describe three suitable features to train the artificial neural network?

(03 marks)

ANSWER IN THIS BOX

- Height to Width ratio of a character
- Number of black pixels in a character
- Chain codes
- Spatial and frequency domain features

(ii) Assume that some characters are misclassified by the system. What are the steps which can be taken to minimize the misclassification error?

(02 marks)

ANSWER IN THIS BOX

- Increase the size of the training data set
- Use a post processing step to correct misclassified characters according to a language model

(iii) It was identified that the misclassification error on training data is significantly less than the misclassification error on test data. What might be the reason for this difference? What steps can be taken to rectify this situation?

(04 marks)

ANSWER IN THIS BOX

The ANN is “overtrained” or overfitted to the training data set.
Use the validation data set to determine the place where the ANN starts overfitting
and stop training further.

Index No:

--
