QUESTION AND ANSWER SHEET

Examination, 26 January 2010

1. **Introductions**   (20 points)

1) The usual shape of an introduction is
   a. General-to-specific
   b. Specific-to-general
   c. General-specific-general
   d. Specific-general-specific

2) The main purpose of an introduction is
   a. To present important definitions, assumptions, and theories.
   b. To summarize previous research in your field
   c. To explain the relevance of your research to the field
   d. To identify gaps in previous research

3) Which of the following is **not** often included in an introduction
   a. Motivation for your research
   b. Outline of the paper
   c. Related work
   d. Implications of your approach

4) Which of the following is **not** a typical step in the third move of an introduction?
   a. Announce the main findings of your research
   b. Outline the contents of the paper
   c. Show how your research area is interesting or important
   d. Describe the purpose of your research

5) In which move of an introduction might you find the following phrase?
   The aim of this paper is to . . .
   a. The opening move
   b. The second move
   c. The last move
   d. None of these

6) Which of the following sentences seems most natural?
   a. *Recently, many studies are conducted on memory loss with aging [1–5].*
   b. *Recently, many studies have been conducted on memory loss with aging [1–5].*
   c. *Recently, many studies were conducted on memory loss with aging [1–5].*
   d. *Recently, many studies are being conducted on memory loss with aging [1–5].*

7) What step in move 2 do you think the following sentence is an example of?

    However, very little work has been done on using parallel processing in solving the assignment problem

    a. Indicate a gap in research

    b. Raise questions

    c. Make a counterclaim

    d. Show unresolved conflict

8) Which of the following sentences might be used to indicate a continuation of an existing research tradition?

    a. *Although considerable research has been done on computer recognition of visual and auditory data, there has been little work on smell recognition.*

    b. *An alternative approach might be for each thread to have its own "thread-local" ray buffer.*

    c. *Several studies have indicated that people generally prefer reading materials in print rather than on a screen.*

    d. *A question remains as to whether a tool producing packets according to TCP's congestion control algorithms can predict TCP performance.*

9) Which of the following phrases is "research-oriented"?

    a. *This paper reports on the results obtained . . .*

    b. *The main purpose of the work reported here was to*

    c. *The aim of the present paper is to give . . .*

    d. *In this paper, we give preliminary results of . . .*

10) Which of the following sentences would be most natural in a 2010 publication?

    a. *Colless (1970) has shown that simple clustering methods based on pairwise similarities can give inconsistent estimates of an evolutionary tree.*

    b. *Colless (1970) shows that simple clustering methods based on pairwise similarities can give inconsistent estimates of an evolutionary tree.*

    c. *Colless (1970) showed that simple clustering methods based on pairwise similarities can give inconsistent estimates of an evolutionary tree.*

    d. *It is shown (Colless 1970) that simple clustering methods based on pairwise similarities can give inconsistent estimates of an evolutionary tree.*

2. Reference words

(10 points)

What do the underlined words refer to?

1.1 Flocking & Schooling

Birds flock, fish school, cattle herd. The natural world has many examples of species that organise themselves into groups for some reason, for example to reduce predation. It has been shown [1] that many predators are "tuned" to hunting individuals, and are confused by large numbers of animals organised into a flock or school. Although the evolutionary advantages in flocks had been well characterised no simple models reproducing such behaviours had been demonstrated.

Reynolds created computer simulations of such flocks by modelling a few simple rules [2], and christened individuals that undergo such flocking "boids". Animations based upon boid-like motion have appeared in a number of Hollywood films[1].

The *emergent behaviour* of the flock is the result of the interaction of a few simple rules. In Reynolds' simulation, (1)these were:

  collision avoidance

  velocity matching (flying with the same speed and direction as the others)

  flock centring (trying to fly near the centroid of one's neighbours)

These rules are sufficient to reproduce natural behaviours, particularly if a predator is treated as an obstacle. However (2)their simplicity allows the use of such self-organised behaviour to be extended to serve a more useful purpose: data visualisation. In Information Flocking, a fourth rule is added which modifies the motion of the individuals on the basis of some similarity measure. (3)This measure can be derived from a set of data, with each individual boid associated with a single data item. The flocking motion then becomes a means of visualising the similarities between the individual data items.

A virtual world was created to display the flocking behaviour. Initially (4)this consists of a school of fish swimming around in 3D, but (5)it is easily extended to include such concepts as attractive and repellent objects (which might attract or repel specific items), and predators which might act as data filters.

The initial problem to which Information Flocking was applied is (6)that of visualising the interests of a group of people. Previously, hierarchical clustering techniques [5] have been applied to such data sets. Neural network approaches have also been used [6]. In particular, Orwig and Chen [7] have used Kohonen neural nets [8] to produce graphical representations of such data. While these representations proved to be much faster and at least as powerful as subjective human classification,

(7)<u>they</u> are essentially static. Information flocking is dynamic in that the fish in the simulation can change (8)<u>their</u> behaviour in response to changes in the underlying data set. The output is also dynamic (the fish "swim") which allows the human viewer to identify patterns more easily.

2. Methods

2.1 VRML

The prototype Information Flocking system was developed using VRML (Virtual Reality Modelling Language), version 2.0^2. (9)<u>This</u> is a powerful, emerging technology that allows rapid development of graphical programs. Objects in a VRML "world" can be controlled by means of a Java script. (10)<u>This</u> results in a system that can produce 3D, interactive graphical simulations that can be controlled using all the features of the Java programming language. VRML is also platform independent and allows easy creation of multi-user environments. A schematic of the system is given in Fig. 1.

1. these =
2. their =
3. This measure =
4. this =
5. it =
6. that =
7. they =
8. their =
9. This =
10. This =

3. <u>Ordering sentences in an abstract</u> (20 points)

Number the following sentences and phrases in order from 1–10. Which part of the abstract do you think each sentence is from? Label them as Title, Background, Purpose, Approach, Findings, or Conclusion. (Not all the parts may be present.)

___ ___ In this paper, we propose an approach for automated recovery of an input validation model from program source code.

___ ___ The two criteria can be used to guide the structural testing of input validation

4

in Web applications.

___ ___ We have discovered some empirical properties for characterizing input validation in Web applications

___ ___ Input validation is essential and critical in Web applications.

___ ___ Testing input validation in Web applications through automated model recovery

___ ___ Based on the model, we then formulate two coverage criteria for testing input validation.

___ ___ It is the enforcement of constraints that any input must satisfy before it is accepted to raise external effects.

___ ___ We have evaluated the proposed approach through case studies and experiments.

___ ___ The model recovered is represented in a variant of a control flow graph, called a validation flow graph, which shows essential input validation features implemented in programs.

4. Writing a title and abstract for an existing paper   (50 points)

Read the Japanese-language paper provided and write a suitable title and abstract .
The abstract should be 90–110 words. Your score will be determined by (1) number of
words, (2) accurate content, (3) understandability, (4) structure, and (5) formal style.

QUESTION AND ANSWER SHEET