

**PRINCE OF SONGKLA UNIVERSITY**  
**FACULTY OF ENGINEERING**

**Midterm Examination:** Semester 2

**Academic Year:** 2013

**Date:** 7 January 2014

**Time:** 09.00-12.00 (3 hours)

**Subject Number:** 242-500

**Room:** S101

**Subject Title:** Research and Development Methodologies

---

**Exam Duration:** 3 hours (180 minutes)

**This paper has 11 pages, 8 questions 130 marks (25%).**

**Authorised Materials:**

- Writing instruments (e.g. pens, pencils).
- Textbooks, a notebook, handouts, and dictionaries are permitted.

**Instructions to Students:**

- Scan all the questions before answering so that you can manage your time better.
- Write your answers in Thai only.
- Write your name and ID on every page.
- Any unreadable parts will be considered wrong.

**Cheating in this examination**

Lowest punishment: Failed in this subject and courses dropped for next semester.

Highest punishment: Expelled.

NO	Time (Min)	Marks	Collected	NO	Time (Min)	Marks	Collected
1	20	15		5	20	20	
2	40	30		6	15	10	
3	30	20		7	10	5	
4	20	20		8	15	10	
<b>Total</b>	<b>170</b>	<b>130</b>		<b>Raw</b>		<b>25%</b>	

**Question 1**

**(15 marks; 20 minutes)**

From the following abstract, find the following items.

การออกแบบและพัฒนาซอฟต์แวร์เหมืองข้อมูลสำหรับการวิเคราะห์กระบวนการผลิตฮาร์ดดิสก์ไดรฟ์

โครงการนี้เสนอการออกแบบและการพัฒนาอัลกอริทึมและซอฟต์แวร์สำหรับวิเคราะห์การเพิ่มปริมาณผลผลิตในอุตสาหกรรมฮาร์ดดิสก์ไดรฟ์ โดยเน้นการหาความสัมพันธ์ของคุณลักษณะ (Attributes), พารามิเตอร์ (Parameters) และผลผลิต (Yield) ในการผลิต อาศัยข้อมูลจากคลังข้อมูลกลางของบริษัทซึ่งเก็บข้อมูลการผลิตจากโรงงานต่างๆ โดยการดึงข้อมูลมาใช้ประโยชน์จะผ่านอาศัยการจัดการข้อมูลแบบ ETL (Extract Transform Loader) ของซอฟต์แวร์ที่มีชื่อว่า Informatica ซึ่งข้อมูลที่จะศึกษาคือข้อมูลค่าพารามิเตอร์เดี่ยว (Individual Data) ข้อมูลแสดงความสัมพันธ์ (Correlation Data) และข้อมูลที่อยู่ในรูปการรวมผลคำนวณ (Computation Aggregation) ของการวิเคราะห์ส่วนประกอบชุดหัวอ่าน (HGA) โดยใช้หลักการทำเหมืองข้อมูล (Data Mining) โครงข่ายประสาทเทียม (Neural Networks) และหลักการสร้างต้นไม้เพื่อใช้ประกอบการตัดสินใจ (Decision Tree)

ซอฟต์แวร์ที่พัฒนาขึ้นจะสามารถใช้เพื่อวิเคราะห์ข้อมูลการผลิตเชิงวิศวกรรม โดยประกอบด้วยเครื่องมืออัตโนมัติสำหรับการวิเคราะห์ผลผลิต (Yield Analysis) และการหาจุดที่เหมาะสมที่สุด (Optimum Point) ที่ใช้ในการปรับเปลี่ยนคุณลักษณะการกระจาย (Distribution) ของพารามิเตอร์ตัวที่มีความสำคัญต่อผลผลิต เช่น การย้ายตามค่าเฉลี่ยของพารามิเตอร์ (Mean Shift) และการลดค่าเบี่ยงเบนของพารามิเตอร์ (Variance Reduction) การวิเคราะห์ความผันแปรของผลผลิต (Yield Variability) เมื่อฮาร์ดดิสก์ไดรฟ์ผ่านกระบวนการผลิตที่มีลักษณะคล้ายกัน แต่ได้ปริมาณผลผลิตต่างกัน ทั้งนี้กระบวนการพัฒนาซอฟต์แวร์จะคำนึงถึงประสิทธิภาพในการประมวลผล โดยอาศัยเทคโนโลยีการประมวลผลแบบขนานบนสถาปัตยกรรมแบบมัลติคอร์ (Multi-Core) นอกจากนี้เครื่องมือต่างๆ ที่พัฒนาขึ้นจะถูกนำไปบูรณาการกับระบบซอฟต์แวร์ที่ใช้อยู่ในปัจจุบันเพื่อความสะดวกในการใช้งานอีกด้วย

คำสำคัญ เหมืองข้อมูล, ฮาร์ดดิสก์ไดรฟ์, อุตสาหกรรม, ตัดสเปค, การย้ายตามค่าเฉลี่ย, การลดค่าเบี่ยงเบน

a) Contribution (5 marks)

.....

.....

.....

.....

.....

b) Objectives and Scopes (5 marks)

.....

.....

.....

.....

.....

Name \_\_\_\_\_ ID \_\_\_\_\_

c) Motivation/Problem Statements (5 marks)

.....  
.....  
.....  
.....  
.....  
.....

**Question 2 (40 marks; 30 minutes)**

Compare the following pairs.

a) Articles published in *conference proceedings* and articles published in *journals* (4 marks)

.....  
.....  
.....  
.....  
.....  
.....  
.....

b) *Research* and *innovation* (4 marks)

.....  
.....  
.....  
.....  
.....  
.....  
.....

c) *Basic research* and *applied research* (4 marks)

.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....

d) Items that should appear in a *proposal* and items that should appear in a *thesis* (4 marks)

.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....

e) *Impact factor* and *citation index* (4 marks)

.....  
.....  
.....  
.....  
.....  
.....  
.....

f) Front matter and back matter of a formal report (5 marks)

.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....

g) Selling points when working with *old data* and *new data*. (5 marks)

.....  
.....  
.....  
.....  
.....  
.....  
.....

**Question 3****(20 marks; 30 minutes)**

Tell whether the following statements are right (T) or wrong (F).

- \_\_\_\_\_ a) If the deadline is approaching, you must publish a paper that has your supervisor's name on it although your supervisor has not yet approved it.
- \_\_\_\_\_ b) You should prepare the slide presentation on your own. You should not bother your supervisor about this.
- \_\_\_\_\_ c) The supervisor is responsible for the content and scope of your thesis.
- \_\_\_\_\_ d) If we do not want to repeat the mistakes of the others, we need to read the literature.
- \_\_\_\_\_ e) Written academic English must contain contractions, hesitation fillers, and familiar Language.
- \_\_\_\_\_ f) The more number of pages for presentation, the better presentation you can show to the audience.
- \_\_\_\_\_ g) We should focus at the solution rather than the problem. How to build a solution is more important than understanding what the problem is.
- \_\_\_\_\_ h) A thesis is a technically feasible plan for solving a problem.
- \_\_\_\_\_ i) A thesis usually follows the chronology of how the research is conducted.
- \_\_\_\_\_ j) The proposal is a formal document of which purpose is to prove that you have made an original contribution to knowledge.
- \_\_\_\_\_ k) You should make it clear when the words or ideas that you are using are your own and when they are taken from another writer.
- \_\_\_\_\_ l) Outlines can be used to help you prepare to write a single paragraph, a composition, a paper, or even a book.
- \_\_\_\_\_ m) Abstract, Contents, Introduction, and Conclusions are read first by the readers or referees and therefore give very first impressions which are important.
- \_\_\_\_\_ n) Plagiarism is stealing or copying other people's work without a formal permission or does not give credit to nor refer to.
- \_\_\_\_\_ o) It is OK to paraphrase part of others' work without referring to the writer.
- \_\_\_\_\_ p) Use graphs when you want to show exact data.
- \_\_\_\_\_ q) To draw quantitative conclusions, you must use the actual numbers.
- \_\_\_\_\_ r) When working through something, where the process is important, use the white board
- \_\_\_\_\_ s) You should spend vast amounts of time drawing complex and beautiful diagrams for more impression of the readers.

\_\_\_\_\_t) Use handouts for complex tables, with lots of figures, equations, and anything that cannot be understood within 30 seconds

**Question 4**

**(20 marks; 20 minutes)**

a) Give at least 3 reasons why we need to do research. (3 marks)

.....  
.....  
.....  
.....  
.....

b) Explain different aims in writing *abstracts*, *introductions* and *conclusions* (6 marks)

.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....

c) Why do we need to do a literature survey? Give at least 3 reasons. (3 marks)

.....  
.....  
.....  
.....  
.....

d) What do you need to do at the meeting with your supervisor? (3 marks)

.....  
.....  
.....  
.....  
.....  
.....  
.....

e) What are pitfalls and shortcomings of using the *Copy-and-Paste* method?  
(5 marks)

.....  
.....  
.....  
.....  
.....  
.....

**Question 5** (20 marks; 20 minutes)

Answer the following questions about research documents.

a) What do successful proposals need? (4 marks)

.....  
.....  
.....  
.....  
.....

b) What are the pitfalls about writing a proposal? (4 marks)

.....  
.....  
.....  
.....  
.....  
.....

c) List the ethics in writing a research article. (4 marks)

.....  
.....  
.....  
.....  
.....

d) How do we write a literature survey? (4 marks)

.....  
.....  
.....  
.....  
.....

e) Explain, at least 4 ways, how to increase creativity. (4 marks)

.....

.....

.....

.....

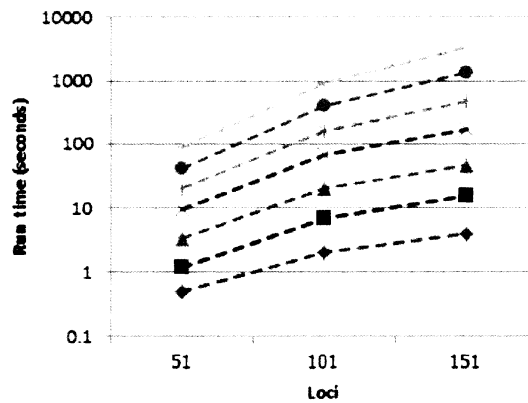
.....

**Question 6**

**(10 marks; 15 minutes)**

Spot what are wrong with the following data representations. Also inform what should be changed or added.

a) (2 marks)



SEQ Radix.

.....

.....

b) (2 marks)

Parameters	Values
Number of Genotypes ( $G$ )	500, 1,000, 2,000, 5,000, 10,000, 20,000, 40,000
Number of loci in Genotypes ( $L$ )	51, 101, 151
SNPHAP Versions	original, -O3, SEQ Radix, OMP Radix 4T, 8T, 16T

Table 1. Parameter sets of our experiments.

.....

.....



c) (2 marks)

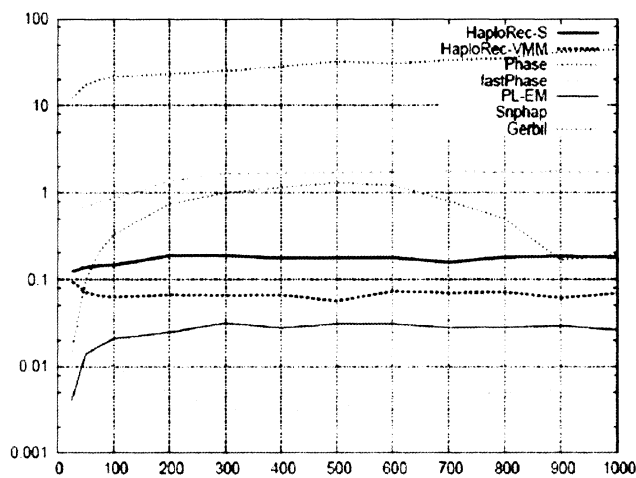


Fig. 1. Run time per genotype in seconds vs. Number of genotypes of each program (Loci=30)

d) (2 marks)

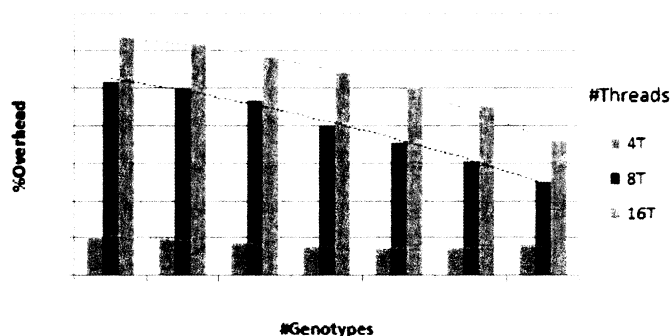


Fig. 5. %Overheads of OMP Radix SNPHAP on AMD A6-3650, L 51 loci

e) (2 marks)

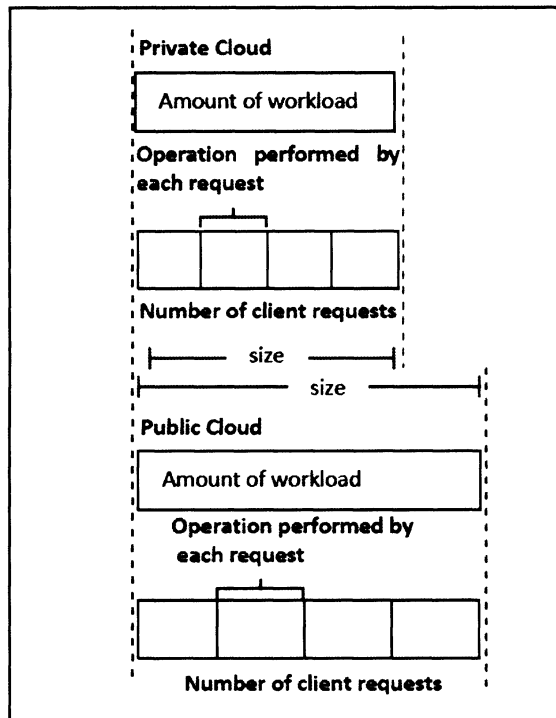


Fig. 1 Amount of workload in the Public Cloud and Private Cloud

.....  
 .....

**Question 7** (5 marks; 10 minutes)

From the following requirements, propose proper media.

- a) You want to compare groups or data with obvious differences.  
 .....
- b) You want to show data that relates to the total number, data ratio or percentage.  
 .....
- c) You want to show continuous data, the relationship between independent and dependent variables that have the same distance between data.  
 .....
- d) You want to show continuous data and the relationship between two interested variables. The distances between data are not the same and it is required to find the regression relationship.  
 .....
- e) You want to show just general data with some statistical analysis.  
 .....

**Question 8**

**(10 marks; 15 minutes)**

Explain how methodology is revealed in each part of a research report.

a) Introduction

(2.5 marks)

.....  
.....  
.....  
.....  
.....

b) Literature Review

(2.5 marks)

.....  
.....  
.....  
.....

c) Method

(2.5 marks)

.....  
.....  
.....  
.....  
.....

d) Result and Discussion

(2.5 marks)

.....  
.....  
.....  
.....  
.....

Pichaya Tandayya  
Lecturer

**An unanswered question  
is better than  
an unquestioned answer.**