COMPUTER SCIENCE TRIPOS Part IB

Tuesday 6 June 2017 1.30 to 4.30

COMPUTER SCIENCE Paper 4

*Answer **five** questions.*

*Submit the answers in five **separate** bundles, each with its own cover sheet. On each cover sheet, write the numbers of **all** attempted questions, and circle the number of the question attached.*

You may not start to read the questions printed on the subsequent pages of this question paper until instructed that you may do so by the Invigilator

STATIONERY REQUIREMENTS
Script paper
Blue cover sheets
Tags

SPECIAL REQUIREMENTS
Approved calculator permitted

1 Artificial Intelligence

`EvilRobot` has two dogs called `Fido` and `Fifi`. All three of them enjoy `pie` and `sausages` so much that they like to steal them. At the beginning of the day the `butcher` has some `sausages` and the `pieShop` has some `pie`. Also, `EvilRobot` and his pets are at home, but they aim to end the day having relieved the local businesses of their products.

(*a*) Give a detailed definition of a *Constraint Satisfaction Problem (CSP)*. Include in your answer a definition of what it means for an assignment to be *consistent* and to be *complete*, and for an assignment to be a *solution*. [4 marks]

(*b*) Consider the constraint C on four variables $\{V_1, V_2, V_3, V_4\}$ each of which has the domain $\{\texttt{true}, \texttt{false}\}$, with

$$\begin{aligned} C = \{&(\texttt{true}, \texttt{true}, \texttt{true}, \texttt{true}) \\ &(\texttt{true}, \texttt{false}, \texttt{true}, \texttt{false}) \\ &(\texttt{false}, \texttt{true}, \texttt{false}, \texttt{false}) \\ &(\texttt{false}, \texttt{false}, \texttt{false}, \texttt{true}) \\ &(\texttt{false}, \texttt{false}, \texttt{true}, \texttt{true})\}. \end{aligned}$$

Explain how this constraint can be replaced by a collection of binary constraints having an identical effect. [3 marks]

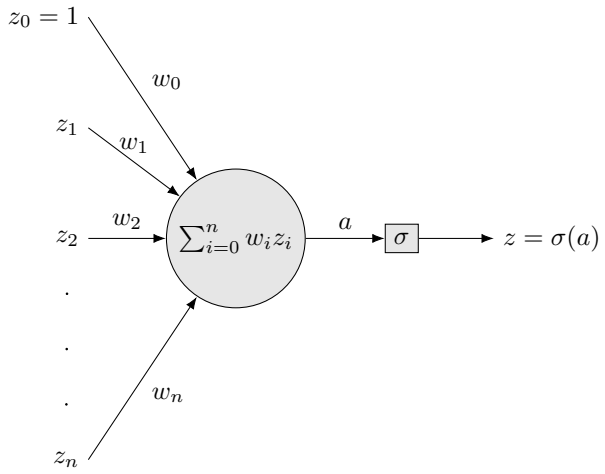
(*c*) Describe the *state-variable representation* for planning problems. Illustrate your answer by showing how the action of `EvilRobot` (or one of his pets) stealing something could be represented in the scenario set out at the beginning of this question. [5 marks]

(*d*) Explain how the *state* of a planning problem can be represented in the state-variable representation. [2 marks]

(*e*) Using your example of the stealing action provided in part (c), explain how this planning problem might be translated into a CSP. You should include in your explanation examples of the translation for actions, state variables, and action preconditions and action effects, but you need not describe the translation for frame axioms. [6 marks]

2 Artificial Intelligence

This question is about *neural networks*. We consider initially *multilayer perceptrons* with nodes of the following kind.



(*a*) Derive an expression for the gradient $\frac{\partial E_i(\mathbf{w})}{\partial w_j}$ for weight w_j in an output node when $E_i(\mathbf{w})$ is the error for the ith example

$$E_i(\mathbf{w}) = \frac{1}{2}(y_i - h(\mathbf{w}; \mathbf{x}_i))^2,$$

$h(\mathbf{w}; \mathbf{x}_i)$ is the output of the complete network for the ith example, and $\sigma(a) = a$. You need only derive the expression for the output node. [3 marks]

(*b*) Derive an expression for the gradient $\frac{\partial E_i(\mathbf{w})}{\partial w_j}$ for weight w_j in an output node when $\sigma(a) = 1/(1 + \exp(-a))$ and the error for the ith example is

$$E_i(\mathbf{w}) = -y_i \log h(\mathbf{w}; \mathbf{x}_i) + (1 - y_i) \log(1 - h(\mathbf{w}; \mathbf{x}_i)).$$

You may use the fact that $d\sigma(a)/da = \sigma(a)(1 - \sigma(a))$. You need only derive the expression for the output node. [7 marks]

(*c*) In the standard backpropagation algorithm the central quantity of interest for each node N is $\delta = \partial E_i(\mathbf{w})/\partial a$. It is proposed that, instead of using nodes in the form presented above, we introduce functions ϕ_i and construct multilayer networks from nodes that compute $z = \sigma(a)$ where

$$a = \sum_{i=0}^{n} w_i \phi_i(\mathbf{z}).$$

Here, $\mathbf{z}^T = \begin{bmatrix} z_0 & z_1 & \cdots & z_n \end{bmatrix}$ and the functions ϕ_i are fixed, having no further parameters. A multilayer perceptron is constructed from nodes of this kind. Give a detailed, general derivation of the formula for computing δ for a *non-output* node N in this network, assuming you know the values of δ for the nodes connected to the output of N. [10 marks]

3 Computer Graphics and Image Processing

Consider the calculation of light emanating from a point on a surface.

(*a*) What is meant by the following terms? Explain how their contribution to the overall amount of reflected light is calculated.

(*i*) Ambient illumination [2 marks]

(*ii*) Diffuse reflection [4 marks]

(*iii*) Specular reflection [4 marks]

(*b*) Suppose that the surface is represented as a polyhedral mesh with triangular faces. Explain how illumination is calculated across a face using each of the following.

(*i*) Gouraud shading [3 marks]

(*ii*) Phong shading [3 marks]

(*c*) Explain where the calculations for Gouraud and Phong shading should be performed when using OpenGL. [4 marks]

4 Computer Graphics and Image Processing

Consider display technologies for hand-held devices.

(*a*) Explain the principles of operation of each of the following.
Note: You may illustrate your answers with a diagram.

(*i*) Liquid crystal displays [5 marks]

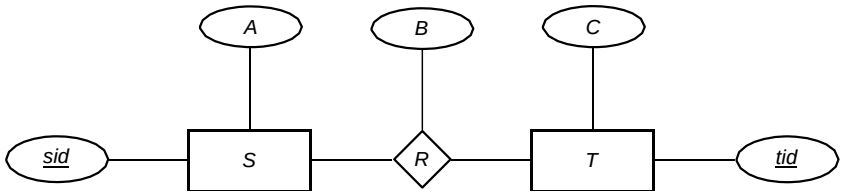
(*ii*) Electrophoretic (electronic paper) displays [5 marks]

(*b*) Compare and contrast their characteristics. [6 marks]

(*c*) Explain how liquid crystal and electrophoretic displays can show coloured images. [4 marks]

5 Databases

(a) Consider the following Entity-Relationship (ER) diagram.



Suppose we wish to implement this diagram in a relational database using three tables, $S(\underline{sid}, A)$, $T(\underline{tid}, C)$, and $R(\cdots)$. Describe the schema you would use for R depending on the cardinality of the relationship.

(i) When R is a many-to-many relationship between S and T. [2 marks]

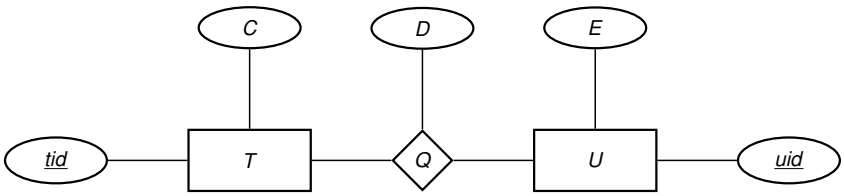
(ii) When R is a one-to-many relationship between S and T. [2 marks]

(iii) When R is a many-to-one relationship between S and T. [2 marks]

(iv) When R is a one-to-one relationship between S and T. [2 marks]

(b) Suppose R is a many-to-one relationship. Rather than implementing a new table for R, can we modify one of the tables representing S or T to implement this relationship? Discuss the advantages and disadvantages of such a representation. [4 marks]

(c) Suppose that we add the following diagram to our ER model.



Note that this implicitly defines a relationship between S and U resulting from the composition of relationships R and Q. Discuss the difficulties that you might encounter in attempting to implement this derived relationship directly in a table W. For example, would the results of evaluating this SQL

```
select sid, tid, B, D
from R
join Q on R.tid = Q.tid
```

always be equivalent to the contents of such a W? [8 marks]

6 Databases

Consider the following three tables.

$S(\underline{sid},\ A)$

sid	A
s1	a1
s2	a2
s3	a3

$R(\underline{sid},\ \underline{tid},\ B)$

sid	tid	B
s1	t1	b1
s1	t2	b2
s2	t1	b4
s2	t3	b5

$T(\underline{tid},\ C)$

tid	C
t1	c1
t2	c2
t3	c3
t4	c4

(a) Represent the information in these three tables in a single table, using NULL values where needed. [4 marks]

(b) Represent the information in these three tables as three JSON objects, each associated with one of the values of the *sid* key. Is any information lost? [4 marks]

(c) Represent the information in these three tables as four JSON objects, each associated with one of the values of the *tid* key. Is any information lost? [4 marks]

(d) We now have three distinct ways of representing the same information (the original tables, one big table, and the collection of JSON objects from parts (b) and (c)). Carefully compare and contrast these approaches and discuss their related advantages and disadvantages. [8 marks]

7 Economics, Law and Ethics

You are commissioned by a customer to design a toy robot that children will be able to control using a smartphone app. This app will also enable them to program the robot using a simple scripting language. To simplify the networking, all communications between the app and the robot flow over wifi via your server.

(*a*) Discuss the ethical and legal implications. [4 marks]

(*b*) Your customer decides to incorporate a microphone so that the robot can also recognise spoken commands. To save battery life, the speech recognition will be done in the server. What effect does this have on the ethical and legal situation? [4 marks]

(*c*) What practical advice can you give your customer about mitigating the legal risks? [4 marks]

(*d*) Your customer now wants to include a camera so that the robot can recognise gestures as well. Does this create any further ethical or legal risks, and if so, what might be done about them? [4 marks]

(*e*) How might the situation be affected by Brexit? [4 marks]

8 Security I

(*a*) *NybbleCrypt* is a block cipher optimized for use in exam questions. It has a block size of 4 bits and a key length of 64 bits. Each block can be written as a single hexadecimal digit, for example `5` $\oplus$ `9` = `c`.

(*i*) The *NybbleCrypt* encryption function for a particular key K is given in the following table:

m	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
$E_K(m)$	c	8	2	7	d	0	6	1	a	e	f	4	b	9	5	3

Decrypt the following messages, which were encrypted using E_K under the following modes of operation, respectively:

(A) ECB mode: `c994f88` [2 marks]

(B) CBC mode: `b144f` [3 marks]

(C) OFB mode: `eae26` [3 marks]

(*ii*) Calculate the CBC-MAC of the following message, using the same key K as in part (*a*)(*i*) above: `face` [2 marks]

(*iii*) *NybblePay* point-of-sale card terminals send 4-digit customer PINs to the bank's transaction-processing centre for verification. The bank's reply to the terminal consists of a 7-digit message in the following format:

(A) 4-digit PIN $m_1m_2m_3m_4$

(B) 2-digit result code m_5m_6: `10` if the PIN was correct, `e1` if not

(C) check digit $m_7 = m_1 \oplus \cdots \oplus m_6$ (the bit-wise XOR of previous digits)

This reply is sent OFB-encrypted using the *NybbleCrypt* blockcipher. You have intercepted such a ciphertext message: `a59defc2`. You are confident that it contains the result code m_5m_6 = `e1` for an incorrect PIN. Without knowing the encryption key K, modify the ciphertext message such that after decryption it shows the result code for a correct PIN, and a matching check digit, while preserving the included PIN. [5 marks]

(*b*) *NybbleShuffle* is a transposition cipher that operates on blocks of 32768 bytes. It splits each such block into 4-bit subblocks, and then rearranges these subblocks in pseudo-random order, under the control of a secret key K, in order to form the 32768-bytes long ciphertext block that it outputs. What is the smallest number of test blocks that you have to feed into an instance of the *NybbleShuffle* cipher in order to unambiguously reconstruct the permutation of subblocks that it applies, and how do you construct these test blocks? [5 marks]

9 Security I

(*a*) Let $\mathsf{Enc}_{K_\mathrm{E}}$ be the encryption function of an encryption scheme that provides indistinguishability under chosen plaintext attack (CPA security). Let $\mathsf{Mac}_{K_\mathrm{M}}$ be a message-authentication-code function that provides existential unforgeability. Named below are three techniques for applying these two functions together to a message M. For each of them

- briefly explain how $\mathsf{Enc}_{K_\mathrm{E}}$ and $\mathsf{Mac}_{K_\mathrm{M}}$ are combined, and
- state whether the resulting construct is likely to provide indistinguishability under chosen ciphertext attack (CCA security):

(*i*) encrypt-and-authenticate [2 marks]

(*ii*) authenticate-then-encrypt [2 marks]

(*iii*) encrypt-then-authenticate [2 marks]

(*b*) How can an attacker calling the C function `parse_text` below cause a buffer overflow? Explain how and why this works. [6 marks]

```
#include <stdlib.h>
#include <string.h>
#define BUFLEN 4096
int check(int n) {
  if (n > BUFLEN) abort();
  return n;
}
void parse_text(char *text, size_t len) {
  char buf[BUFLEN];
  memcpy(buf, text, check(len));
  /* ... */
}
```

(*c*) Many Unix system administrators create a personal group for each of their users with this user as the sole member.

(*i*) What is the purpose of such a group? [2 marks]

(*ii*) Such personal groups typically have the same name and integer identifier as the corresponding user identifier. Is this practice compatible with the Windows NT mechanism for identifying users and groups? [2 marks]

(*d*) Give two examples for resources where an operating system is expected to implement residual information protection and two alternative mechanisms for implementing it. What are their tradeoffs and threat assumptions? [4 marks]

END OF PAPER