2006

Department of Creative Informatics
Graduate School of Information Science and Technology
The University of Tokyo

Information Science and Technology

Instructions

1. Do not open this brochure until the signal to begin is given.
2. Write your examinee ID below on this cover.
3. Answer three out of four problems.
4. Three answer sheets are given. Use a separate sheet for each problem. You may use the backside of the sheet.
5. Write down the examinee ID and the problem ID inside the top blanks of each sheet.
6. Do not take out the sheets and this brochure from this room.

Examinee ID _____

Problem 1

The *edit distance* between two strings $str1$ and $str2$ is defined as the minimum number of the following operations required to transform $str1$ into $str2$.

- insert one character
- delete one character
- substitute one character by another character

For example, $str1$="sport" is trasnformed into $str2$="sort" by deleting the character 'p'; therefore the edit distance is 1.

(1) Answer the edit distance between $str1$="commuter" and $str2$="computers".

(2) Let's denote str_n as the prefix of length n of a given str, $m(i, j)$ as the edit distance between $str1_i$ and $str2_j$. Write down the recursive formula that holds between $m(i, j)$ and $m(i-1, j)$, $m(i, j-1)$, $m(i-1, j-1)$.

(3) Describe an algorithm for calculating the edit distance between two given strings based on the recursive function described in (2). Show its complexity in both space and time.

(4) Calculate the edit distance between $str1$="abrabr" and $str2$="arbarb".

(5) Describe three applications of the edit distance.

Problem 2

Answer the following questions about a cache memory in a processor.

(1) Describe the reason within five lines why cache memory increases execution speed of programs.

(2) Suppose that the processor has an instruction cache and a data cache, the size of a cache block is 32 bytes. Calculate the cache hit ratio of a program that multiplies each element of a vector A whose size is sufficiently large N and whose elements are 32-bit long.

Example of the program (written in C programming language)

```
for(i=0; i < N; i = i + 1)
            A[i] = k * A[i];
```

(3) Show a block-diagram of a 2-way set associative cache memory, using memories, comparators, multiplexers, registers (flop-flops) etc.

(4) Describe five methods to increase cache hit ratio both in hardware technology and programming techniques. Each item should be described within two lines.

Problem 3

Consider a pinhole camera model as shown in Fig. 1, where the light from an object in three dimensional (3-D) space goes to the screen through a single point C (optical center). The vertical line to the screen through the optical center is called the optical axis, and the point where the optical axis crosses the screen is called the image center. Set the 3-D Camera coordinate system C-X,Y,Z at the optical center and the two dimensional (2-D) Image coordinate system I-x,y at the image center as shown in Fig. 1. When the distance from the optical center to the screen is f and the 3-D point $\mathbf{P}(\mathrm{X}, \mathrm{Y}, \mathrm{Z})$ is projected to the 2-D point $\mathbf{p}(\mathrm{x}, \mathrm{y})$ on the screen, the coordinates of $\mathbf{p}(\mathrm{x}, \mathrm{y})$ are described as follows:

$$\mathrm{x} = \mathrm{f}\frac{\mathrm{X}}{\mathrm{Z}}, \quad \mathrm{y} = \mathrm{f}\frac{\mathrm{Y}}{\mathrm{Z}}$$

Answer the questions below.

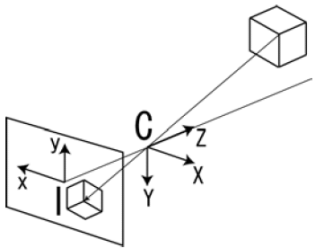

Fig. 1

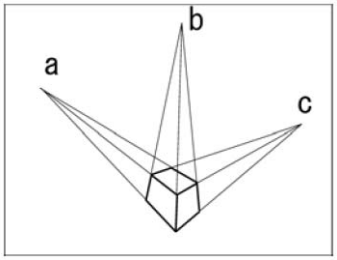

Fig. 2

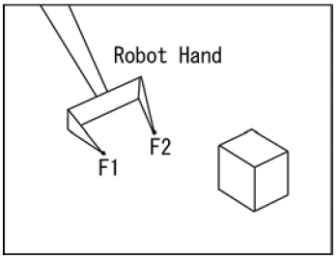

Fig. 3

(1) The point on the screen projected from the infinity point on a line L in 3-D space is called the vanishing point of L. When L goes through a 3-D point $\mathbf{P0}(\mathrm{X}_0, \mathrm{Y}_0, \mathrm{Z}_0)$ with an orientation vector $\mathbf{m}(\mathrm{m}_1, \mathrm{m}_2, \mathrm{m}_3)$, describe the 3-D coordinates of the point $\mathbf{P}(\mathrm{X}, \mathrm{Y}, \mathrm{Z})$ on L and calculate the coordinates of the vanishing point $\mathbf{p}(\mathrm{x}, \mathrm{y})$ of L.

(2) Fig. 2 shows the projected image of a rectangular solid and three vanishing points $\mathbf{a}(\mathrm{a_x}, \mathrm{a_y})$, $\mathbf{b}(\mathrm{b_x}, \mathrm{b_y})$, $\mathbf{c}(\mathrm{c_x}, \mathrm{c_y})$ of the parallel edge lines on the solid. Describe the equations that hold among the three vanishing points.

(3) When the projected image of a rectangular solid is given as a gray image, explain the method of image processing for getting the edge lines from the gray image.

(4) Fig. 3 shows an image of a robot hand controlable by specifying its position and orientation in the base coordinate system of the robot. Image processing provides the motion trajectories of the end points F1, F2 of the robot fingers on the screen. Consider the method to get the transformation matrix from the robot base coordinate system to the camera coordinate system by translating the robot hand in three orthogonal directions. Explain what kind of additional information other than the coordinates of three vanishing points is required to get the transformation matrix.

Problem 4

Select four items out of the following eight items regarding information systems, and explain each item in approximately $5 \sim 10$ lines. (When indicated, explain with giving an example.)

(1) "Join" operation in relational databases. (Explain with an example.)

(2) The difference between "process" and "thread" in operating systems or programs.

(3) Context-free grammar. (Explain with an example.)

(4) "Gouraud shading" and "Phone shading" in computer graphics (CG).

(5) Image coding schemes "GIF" and "JPEG" universally used in the WWW.

(6) Digital Signature method using the public-key encryption.

(7) Knowledge representation using the "Semantic Network", and its feature. (Explain with an example.)

(8) "PID control", and the roles of its components: "P", "I" and "D".