Summer Examination, 2009

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 the four problems in Japanese or English.
4. Three answer sheets are given. Use a separate sheet for each problem. You may continue to write your answer on the back of the answer sheet if you cannot conclude on the front.
5. Write down the examinee ID and the problem ID inside the top blanks of each sheet.
6. The answer is considered invalid if any unrelated marks, codes or phrases are included.
7. Do not remove the sheets and this brochure from this room.

Examinee ID _____

Problem 1

Given n records $r_1, r_2, \ldots, r_n$, each has a key $k_1, k_2, \ldots, k_n$, respectively. Given a key, consider searching for the corresponding record. For a given query key v, if there is a key k_j where $v = k_j$ $(1 \leq j \leq n)$ then the search is successful and the record r_j is returned, otherwise the search fails and the failure is returned.

Let p_j be the probability where the query key v is k_j, and p_{miss} be the probability of the search failure. The computation time is almost proportional to the number of comparisons with v. Let us calculate the average number of comparisons $\overline{C}$ and the maximum number of comparisons $C_{\max}$.

(1) Consider a sequential search that compares v with keys from k_1 to k_n.

(a) Obtain $\overline{C}$ and $C_{\max}$ when $p_1 = p_2 = ... = p_n = \dfrac{1}{2n},\ p_{\text{miss}} = \dfrac{1}{2}$.

(b) Prove that $\overline{C} \leq 2$ when $p_1 = \dfrac{1}{2},\ p_2 = \dfrac{1}{4}, ...,\ p_n = \dfrac{1}{2^n},\ p_{\text{miss}} = \dfrac{1}{2^n}$.

(2) Consider a binary search after sorting the keys. Let $n = 2^m - 1$ (m is a natural number), and moreover, one comparison will determine whether $v < k_j, v = k_j$ or $v > k_j$.

(a) Obtain $\overline{C}$ and $C_{\max}$ for each case of m=1, 2, 3, 4, when $p_1 = p_2 = ... = p_n = p_{\text{miss}}$.

(b) Obtain $\overline{C}$ and $C_{\max}$ as a function of m when $p_1 = p_2 = ... = p_n = p_{\text{miss}}$.

(3) Consider a search using a hash table. The records $r_1, \ldots, r_8$ are inserted into the hash table of size $S = 17$ using the hash function $h(x) = x \bmod S$. Let the key values each be $k_1 = 10, k_2 = 1, k_3 = 17, k_4 = 97, k_5 = 21, k_6 = 4, k_7 = 39, k_8 = 73$.

(a) Draw the structure of the hash table by choosing an appropriate method for avoiding collision.

(b) Obtain $\overline{C}$ and $C_{\max}$ when $p_1 = p_2 = ... = p_8,\ p_{\text{miss}} = 0$.

(4) Describe in general the advantages and disadvantages of sequential search, binary search and search using a hash table.

Problem 2

Consider an N digit decimal counter specified as follows:

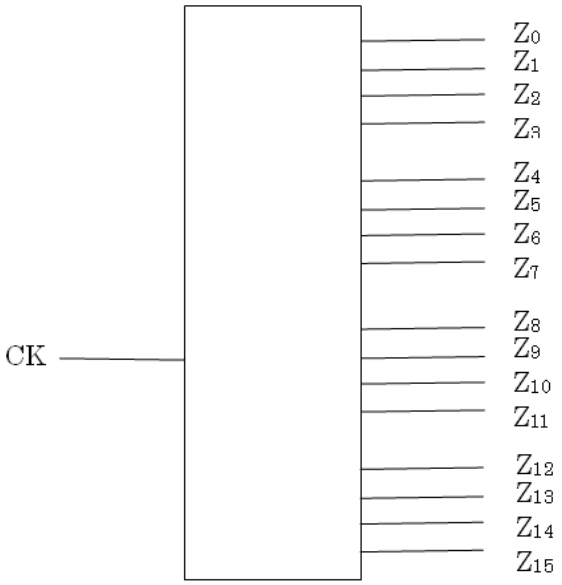
a. A one digit decimal is represented by 4 bits.

b. The counter is synchronous and has a clock CK, $4N$ bit outputs Z_i where $i = 0, \ldots, 4N-1$.

c. The initial value of the counter is 0, namely, $Z_i = 0$ where $i = 0, \ldots, 4N-1$.

d. The value of the counter increases by 1 at every input of the clock. When the value of the counter reaches the maximum value, the next clock input sets the output to be 0.

For example, the following figure depicts the input (the clock CK) and the output ($Z_0 \sim Z_{15}$) representing a decimal number with $N = 4$ digits.



(1) Draw a table or a diagram showing the state-transition for the case N=1.

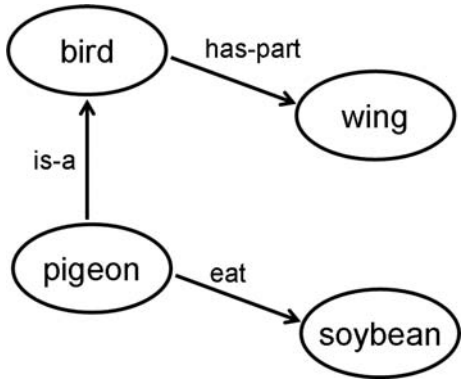
(2) Construct the logic circuit of the counter for the case $N = 1$ using AND, OR, NOT gates and D flip-flops.

(3) Construct the logic circuit of the counter for the case $N = 4$ using 4 counters based on (2) with AND, OR and NOT gates.

(4) For a given N, describe a method to construct the logic circuit of the counter whose delay time is $O(\log N)$. Approximate the delay time by the number of AND, OR and NOT gates between the output and the input of D flip-flops.

Problem 3

There are several methods for knowledge representation. In contrast to a system of production rules and logical formulae which are modular representations, the semantic network (where a node represents a concept or an entity and a named link represents a relation between the nodes) has the characteristics of representing the pieces of knowledge mutually interrelated. For example, the following knowledge can be represented in a semantic network shown on the right.



"is-a" represents a relation such that a pigeon is a bird. "has-part" represents a relation such that a bird has wings as its part.

- A pigeon is a bird.
- A bird has wings.
- A pigeon eats soybeans.

Q1. Represent the following pieces of knowledge 1)∼8) as a semantic network. (Here, you can create necessary link names. In this case, use a meaningful name for a link, or append an explanatory comment so that the meaning of the link name can be understood.)

1) Taro is a human, and works in the Ueno Zoo.
2) Taro takes care of Ueno Zoo's penguins.
3) Yuki is one of Ueno Zoo's penguins, and lives in the Ueno Zoo.
4) A penguin is a bird.
5) A bird can fly.
6) Taro owns a Prius.
7) Kenta is a human, and owns another Prius.
8) Prius is a hybrid car.

Q2. (2-1) In the semantic network made in Q1, "Yuki can fly" can be derived. Explain the reason why this derivation occurs.

(2-2) Describe which widely used programming languages include functionality similar to what occurs in (2-1).

(2-3) Explain the merits of the functionality mentioned in (2-1) and (2-2) from the viewpoint of the description and management of knowledge or programs.

Q3. We want to efficiently infer an answer to the query "who is the human that takes care of Yuki and owns a hybrid car?" in the semantic network made in Q1. Explain what type of inference method can be employed, and find an answer for this query.

Q4. Assume that the following knowledge is added to the semantic network made in Q1.

9) A penguin cannot fly.

In this case, there are the following possibilities.

a) Both "Yuki can fly" and "Yuki cannot fly" can be derived.
b) Neither "Yuki can fly" nor "Yuki cannot fly" can be derived.
c) Only "Yuki cannot fly" can be derived.

Why do such possibilities arise? Explain what type of knowledge manipulation is required to yield the case c).

Problem 4

Select four items out of the following eight items regarding information systems, and explain each item in approximately 4∼8 lines of text, using examples or images if necessary.

(1) Position feedback and torque feedback in a servo system

(2) Principle of proximity sensing (a particular type)

(3) Uncertainty in measurement

(4) Hidden Markov models

(5) Grid computing

(6) Phishing

(7) Reflection in programming

(8) Referential transparency