

平成 21 年度 夏入試

東京大学情報理工学系研究科創造情報学専攻

専門科目

注意事項

1. 試験開始の合図まで, この問題冊子を開かないこと.
2. この表紙の下部にある受験番号欄に受験番号を記入しなさい.
3. 4問中3問 を選択して, 日本語ないし英語で解答すること.
4. 解答用紙は3枚配られる. 1問ごとに必ず一枚の解答用紙を使用すること. 解答用紙のおもて面に書ききれないときには, うら面にわたってもよい.
5. 解答用紙の指定された箇所に, 受験番号およびその用紙で解答する問題番号を記入すること.
6. 解答に関係のない記号, 符号, 文言などを記入した答案は無効とする.
7. 解答用紙および問題冊子は持ち帰らないこと.

受験番号 _____

第1問

n 個のレコード r_1, r_2, \dots, r_n があり, それぞれ k_1, k_2, \dots, k_n をキーに持つとする. このとき, 特定のキーをもつレコードを探索したい. 問い合わせ v が与えられたとき, $v = k_j$ ($1 \leq j \leq n$) となるキー k_j があれば探索成功としてレコード r_j を返し, そうでなければ探索失敗として, 失敗を示す値を返すという処理について考える.

ここで, 問い合わせ v が k_j である確率を p_j とし, 問い合わせ v での探索が失敗する確率を p_{miss} とする. 探索の計算時間は v とキーとの比較回数にほぼ比例するため, 1度の探索に要する比較回数の平均値 \bar{C} と最大値 C_{max} について考える.

(1) キーを k_1 から k_n まで順番に探索する逐次探索を考える.

(a) $p_1 = p_2 = \dots = p_n = \frac{1}{2n}$, $p_{\text{miss}} = \frac{1}{2}$ であるとき, \bar{C} , C_{max} を求めよ.

(b) $p_1 = \frac{1}{2}$, $p_2 = \frac{1}{4}, \dots, p_n = \frac{1}{2^n}$, $p_{\text{miss}} = \frac{1}{2^n}$ のとき, \bar{C} が2以下であることを証明せよ.

(2) キーを事前に整列しておき, 2分探索を利用する場合を考える. なお, ここでは $n = 2^m - 1$ (m は自然数) であるとする. また, 1度の比較で $v < k_j, v = k_j, v > k_j$ のいずれかであることを判断できるとする.

(a) $p_1 = p_2 = \dots = p_n = p_{\text{miss}}$ であるとき, m が 1, 2, 3, 4 のときの \bar{C} , C_{max} をそれぞれ求めよ.

(b) $p_1 = p_2 = \dots = p_n = p_{\text{miss}}$ であるとき, \bar{C} , C_{max} を m の関数として求めよ.

(3) ハッシュ表を用いる場合を考える. サイズ $S = 17$ のハッシュ表, ハッシュ関数 $h(x) = x \bmod S$ を用いて, r_1 から r_8 のレコードをハッシュ表に挿入する. キーの値はそれぞれ $k_1 = 10, k_2 = 1, k_3 = 17, k_4 = 97, k_5 = 21, k_6 = 4, k_7 = 39, k_8 = 73$ とする.

(a) 構築したハッシュ表の構造を図示せよ. なお, 衝突回避法は適宜選択せよ.

(b) $p_1 = p_2 = \dots = p_8, p_{\text{miss}} = 0$ であるとき \bar{C} , C_{max} を求めよ.

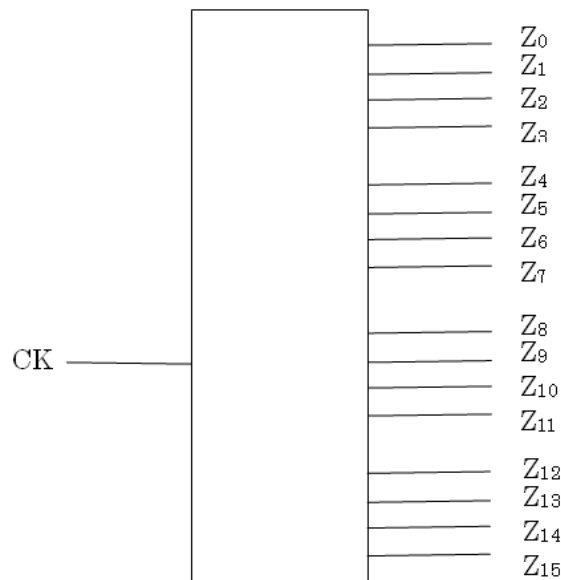
(4) 逐次探索, 2分探索, ハッシュ表を用いる探索について, その一般的な得失を述べよ.

第2問

以下の仕様に従う10進数 N 桁のカウンタについて考えよ.

- 1桁の10進数は4ビットで表現される.
- カウンタは同期式であり, クロック入力 CK と, $4N$ ビットの出力 $Z_i, i=0, \dots, 4N-1$ を持つ.
- カウンタの初期値は0である. すなわち, $Z_i=0, i=0, \dots, 4N-1$ である.
- カウンタの値は, クロック入力がある毎に1ずつ増加する. カウンタの値がカウンタの最大値に達すると, 次のクロック入力により出力が0になる.

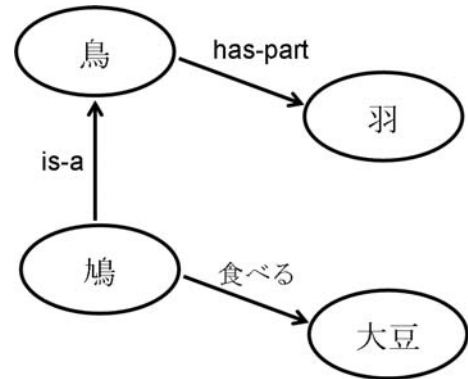
下記の図は, 入力がクロック CK であり, $N=4$ 桁の10進数を出力する回路を示している.



- $N=1$ の場合の状態遷移を表現する表または図を示せ.
- $N=1$ の場合のカウンタ論理回路を, AND, OR, NOT ゲートとDフリップフロップを用いて構成せよ.
- $N=4$ の場合のカウンタ論理回路を, (2) で求めた回路をベースとしたカウンタ4個と, AND, OR, NOT ゲートを用いて構成せよ.
- 任意の N に対し, 遅延時間が $O(\log N)$ であるカウンタ論理回路の構成法を示せ. なお, 遅延時間の近似値としてDフリップフロップの出力と入力の間に入る, AND, OR, NOT ゲートの段数を用いよ.

第3問

知識の表現法は幾つかあるが、ルール表現（プロダクション・システム）や論理式がモジュール性のある表現法であるのに対比すると、概念やエンティ（実在物）をノード、ノード間の関係を名前付きリンクで表す意味ネットワークは、個々の知識を相互に関連づけて表すことが特徴である。例えば以下の知識は右図のような意味ネットワークによって表される。



- ・鳩（ハト）は鳥である。
- ・鳥は羽を持つ。
- ・鳩（ハト）は大豆を食べる。

ここで、
is-a: は視覚視覚の一種であることの関係を表す。
has-a: は を部分として持つことの関係を表す。

問 1. 以下 1)~8) に記された内容を意味ネットワークとして表しなさい（必要なリンク名は新たに作成して良いが、リンク名は意味の分かるものにするか、注釈を付して意味が分かるようにすること）。

- 1) 太郎は人間であり、上野動物園で働く。
- 2) 太郎は上野動物園のペンギンの世話をしている。
- 3) ユキは上野動物園のペンギンであり、上野動物園のペンギンは上野動物園に住む。
- 4) ペンギンは鳥である。
- 5) 鳥は飛べる。
- 6) 太郎は1台のプリウスを所有する。
- 7) 健太は人間であり、別のプリウスを所有する。
- 8) プリウスはハイブリッドカーである。

問 2. (2-1) 上記問 1 の知識で作成した意味ネットワークでは、「ユキは飛べる」が導かれることになるが、このようになる理由を述べよ。

- (2-2) 上記 (2-1) と同様な機能は、現在広く普及しているどのようなプログラミング言語にどのような形で内在しているかを述べよ。
- (2-3) 上記 (2-1),(2-2) の機能は、知識あるいはプログラムの記述や管理において、どのようなメリットがあるかを述べよ。

問 3. 上記問 1 の意味ネットワークに対して「ユキの世話をし、ハイブリッドカーを所有する人は誰か？」を推論によって効率的に求めたい。どのような推論法があるかを説明し、この解を求めよ。

問 4. 上記問 1 で作成した意味ネットワークに以下の知識が追加されたとする。

9) ペンギンは飛べない。

この時、扱いによって導かれる事柄に次のようなケースがあり得る。

- a) 「ユキは飛べる」、「ユキは飛べない」のどちらも導ける。
- b) 「ユキは飛べる」、「ユキは飛べない」のどちらも導けない。
- c) 「ユキは飛べない」だけが導ける。

このような事態が生じる原因は何なのか？ また、c) にするにはどのような知識の扱いをすれば良いかを述べよ。

第4問

情報システムに関する以下の項目から 4項目 を選択し、各項目を 4~8 行程度で、必要に応じて例や図を用いて説明せよ。

- (1) サーボ系における位置フィードバックとトルクフィードバック
- (2) 近接覚センサーの原理 (1 種類でよい)
- (3) 計測における不確かさ (uncertainty)
- (4) 隠れマルコフモデル
- (5) グリッド・コンピューティング
- (6) フィッシング (phishing)
- (7) プログラミングにおけるリフレクション
- (8) 参照透明性

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, \dots, r_n , each has a key k_1, k_2, \dots, 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 \bar{C} and the maximum number of comparisons C_{max} .

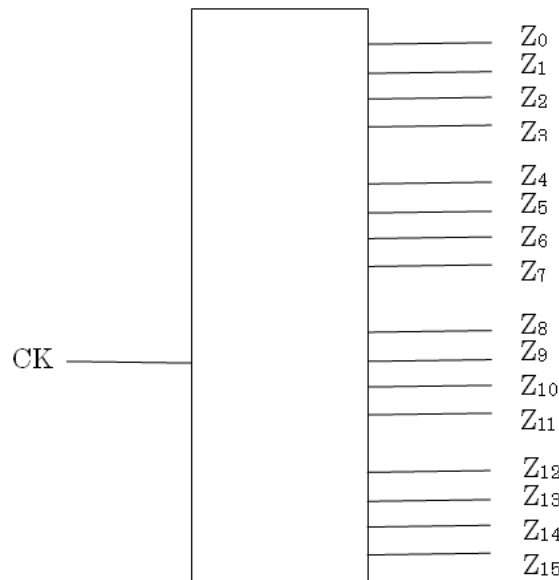
- (1) Consider a sequential search that compares v with keys from k_1 to k_n .
 - (a) Obtain \bar{C} and C_{max} when $p_1 = p_2 = \dots = p_n = \frac{1}{2n}$, $p_{\text{miss}} = \frac{1}{2}$.
 - (b) Prove that $\bar{C} \leq 2$ when $p_1 = \frac{1}{2}$, $p_2 = \frac{1}{4}, \dots, p_n = \frac{1}{2^n}$, $p_{\text{miss}} = \frac{1}{2^n}$.
- (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 \bar{C} and C_{max} for each case of $m=1, 2, 3, 4$, when $p_1 = p_2 = \dots = p_n = p_{\text{miss}}$.
 - (b) Obtain \bar{C} and C_{max} as a function of m when $p_1 = p_2 = \dots = p_n = p_{\text{miss}}$.
- (3) Consider a search using a hash table. The records r_1, \dots, 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 \bar{C} and C_{max} when $p_1 = p_2 = \dots = p_8$, $p_{\text{miss}} = 0$.
- (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 one digit decimal is represented by 4 bits.
- The counter is synchronous and has a clock CK, $4N$ bit outputs Z_i where $i = 0, \dots, 4N - 1$.
- The initial value of the counter is 0, namely, $Z_i = 0$ where $i = 0, \dots, 4N - 1$.
- 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.

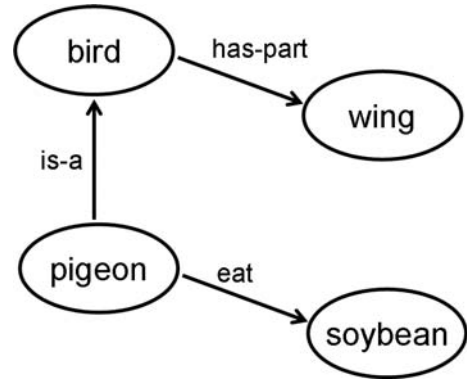
For example, the following figure depicts the input (the clock CK) and the output ($Z_0 \sim Z_{15}$) representing a decimal number with $N = 4$ digits.



- Draw a table or a diagram showing the state-transition for the case $N=1$.
- Construct the logic circuit of the counter for the case $N = 1$ using AND, OR, NOT gates and D flip-flops.
- Construct the logic circuit of the counter for the case $N = 4$ using 4 counters based on (2) with AND, OR and NOT gates.
- 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

