Academic Year of 2015
Admission to the Master's Program
Department of Intelligence Science and Technology
Graduate School of Informatics, Kyoto University
(Fundamentals of Informatics)
(International Course)

13:30 - 15:00, February 11, 2015

NOTES

1. This is the Question Booklet in 3 pages including this front cover.
2. Do not open the booklet until you are instructed to start.
3. After start, check the number of pages and notify proctors (professors) immediately if you find missing pages or unclear printings.
4. **Answer all questions in this booklet.**
5. Write your answers in English.
6. Read carefully the notes on the Answer Sheets as well.

Master's Program	**Fundamentals of Informatics**	**Question Number**	**F-1**

Q. 1 Arithmetic expressions, as they are commonly written, use infix notation for binary operators (e.g. $+, -, *, /$). Infix notation uses parentheses to enclose parts of an expression that must be calculated first. To evaluate an arithmetic expression, a stack (a data structure) is often used in computers.

1.1 Describe the basic principle and operations of a stack (use figures if necessary).

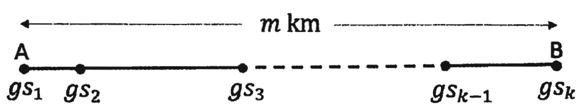
1.2 In addition to infix notation, postfix and prefix notations are used as well. What are the differences among them? What are the advantages of postfix and prefix notations over infix notation while using them in computer? Convert the following infix notation into postfix and prefix notations, where $*$ has higher precedence than $+$ and $-$.

$$(1+2)*(4-3)$$

1.3 Given the infix notation below, convert it to a postfix notation. Describe the procedure of evaluating the postfix notation using a stack.

$$6+((5+4)*(3*2)+1)$$

Q. 2 Suppose you will drive a car from city A, which has gas station gs_1, to city B, which has gas station gs_k, along a specific route with k gas stations, the total distance is m km, where $k > 3$ (see below figure). Before the trip, you know the distances between all the gas stations in the route. Assume that the car's gas tank, when full, holds enough gas to travel n km, which is not shorter than the distance between any two adjacent gas stations. Meanwhile, $n \ll m$.



2.1 Your goal is to minimize the number of stops along the route, without running out of gas at any point. Give an efficient strategy by which you can determine at which gas stations you should stop.

2.2 Prove that your strategy yields an optimal solution, and discuss its time complexity in terms of the number of gas stations.

Q.1 Consider a population of bears; suppose that bears are either **male** with probability 0.5 or **female** with probability 0.5; either **black** with probability 0.6 or **brown** with probability 0.4. Suppose that **male** bears are 3 times more likely to be **black** than **female** bears.

1. What is the probability that a **brown** bear is a **male**?
2. Rank from greatest to least, the entropies of the following random variables:
   - bear color;
   - bear gender;
   - bear color, conditionally to the fact that a bear is male;
   - bear gender, conditionally to the fact that a bear is brown.

Q.2 Consider an alphabet of 8 symbols whose probabilities are as follows:

A	B	C	D	E	F	G	H
$\frac{1}{128}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{16}$	$\frac{1}{128}$

Someone has drawn one of these symbols according to the probabilities above. Suppose you would like to discover that symbol, but that you can only ask questions that have yes/no answers, for instance, *'is the symbol in the set $\{A, B, C\}$?'*, *'is it H?'*.

1. Provide an efficient sequence of yes/no questions that you could ask in order to discover the selected symbol.
2. Following that sequence, how many such questions will you have to ask on average to discover the selected symbol?
3. What is the Shannon entropy of the above symbol set? Using Shannon's theorem, can you conclude that your sequence of questions is the most efficient on average? (if not, provide an optimal strategy).

Q.3

1. Let X be a nonnegative real-valued random variable. Prove that for any constant $T > 0$, the probability $P(X \geq T)$ that X is bigger than or equal to T is such that:
$$P(X \geq T) \leq \frac{E[X]}{T},$$
where $E[X]$, the expectation of X, is assumed to be finite. *hint: compare X with the random variable $T\mathbf{1}_{X\geq T}$, where*
$$\mathbf{1}_{X\geq T} := \begin{cases} 1 \text{ if } X \geq T \\ 0 \text{ if } X < T \end{cases}.$$
2. Let Y be a real-valued random variable with mean $\mu = E[Y]$ and variance $\sigma^2 = E[(Y-\mu)^2]$. Prove then that for any $T > 0$,
$$P\left(\frac{|Y-\mu|}{\sigma} \geq T\right) \leq \frac{1}{T^2}.$$
3. If Y has zero mean and variance 1, what is the upper bound that you can obtain from the result above on the probability that $|Y|$ is bigger than or equal to 2?