

情報 個別問題 (田中哲朗) 2009 年度前期試験

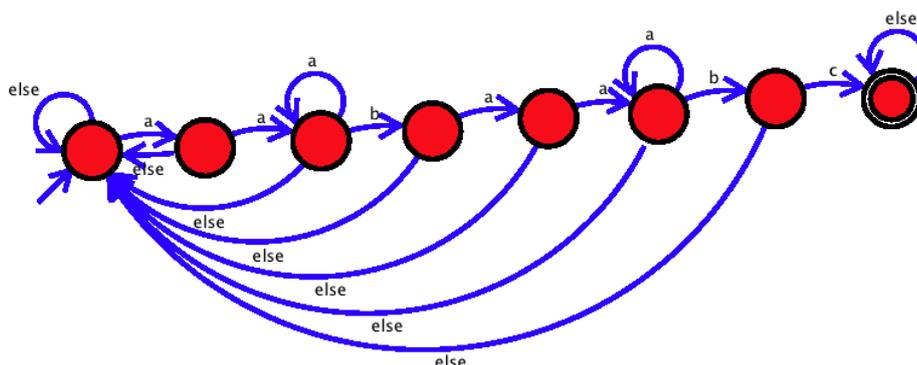
[科目名: 情報, 教員名: 田中哲朗, 7月28日 2限 (10:55-12:25), 試験時間 (共通問題と合わせて): 90分, クラス名: 理科一類 22, 28, 32組, 解答用紙計算用紙各1枚 (共通問題と合わせて)]

個別問題 1

(1) 以下の電子錠をオートマトン (有限状態機械) としてモデル化したい。

- 途中でどのような入力があっても, 最後7文字の入力が暗証番号「aabaabc」なら解錠する。
- 一度解錠した後は何を入力し続けても解錠状態となる。

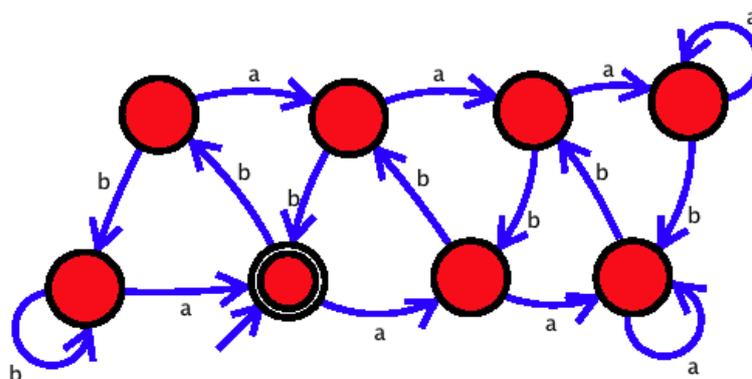
解錠状態が終了状態となるように以下のオートマトンを作ってみた。



このオートマトンには誤りがある。以下の (a)-(c) に解答しなさい。

- このオートマトンに解錠状態になるべきでない入力を与えた時に終了状態になることがある。このような入力のなるべく短い具体例を「aacaaaabbbbb」のような形式で一つ答えなさい。
- このオートマトンに解錠状態になるべき入力を与えても終了状態にならないことがある。このような入力のなるべく短い具体例を「aacaaaabbbbb」のような形式で一つ答えなさい。
- 誤りを修正したオートマトンの図を書きなさい。修正部分でなく全体図を書くこと。

(2) 条件 A を「a と b だけからなる入力で, それまでに現われた b の個数が a の個数の 2 倍」とする。条件 A を満たすかどうかを判定する (条件 A を満たす入力に対しては必ず終了状態, 満たさない入力に対しては必ず非終了状態となる) オートマトンを作ることを試みた。以下が作成してみたオートマトンである。



以下の (d)-(e) に解答しなさい。

- このオートマトンに条件 A を満たしていない入力を与えた時に終了状態になることがある。このような入力のなるべく短い具体例を「aacaaaabbbbb」のような形式で一つ答えなさい。

(e) このオートマトンに条件 A を満たしている入力を与えても終了状態にならないことがある．このような入力のなるべく短い具体例を「aacaaaaabbbbb」のような形式で一つ答えなさい．

実は、「条件 A を満たすかどうかを判定するオートマトン」を作ることはできない．以下の簡単な証明の (f) を埋めて、証明を完成させなさい．

- 条件 A を満たすかどうかを判定する状態数  $N$  のオートマトンが作成できると仮定する．
- $a$  が  $i$  個続いた入力を  $a^i$  と書くことにする．入力の集合  $\{a^0, a^1, \dots, a^N\}$  を考えると、集合のサイズは  $N+1$  なので、 $0 \leq i < j \leq N$  で  $a^i$  の入力後の状態と、 $a^j$  の入力後の状態が等しい  $i, j$  が必ず存在する．
- この状態は  $a^i$  の入力後の状態なので、この後に  $b$  が続けて (f) 個入力されると終了状態にならなくてはいいけない．しかし、 $a^j$  の入力後の状態でもあるので、 $a$  が  $j$  個、 $b$  が (f) 個の入力で終了状態になることになる．これは条件 A を満たさない入力である．
- 以上により、「条件 A を満たすかどうかを判定する状態数  $N$  のオートマトンが作成できる」という仮定に誤りがあったことがわかる．任意の  $N$  に関して成り立つので、「条件 A を満たすかどうかを判定するオートマトン」はできない．

## 個別問題 2

古代の遺跡から  $\{0, 1\}$  の二つの記号からなる謎のデータの一部で長さ 20 の記号列が発見された．

10000111000100010000

この謎のデータが  $\{0, 1\}$  の二つの記号が決まった確率で生成されるデータであり、確率は観測されたデータ中の出現頻度であると仮定して平均情報量を求めてみる．以下の表の (a)-(e) を埋めなさい (値を簡単に計算できない場合は、式の形で表現すること) ．

記号	出現回数	出現確率	情報量	平均情報量
0	14	0.7	(a)	$0.7 \times (a)$
1	(b)	(c)	(d)	$(c) \times (d)$
計	20	1.0		(e)

その後の研究でこの謎のデータは単なる 2 進数の列ではなく、 $A \rightarrow 00, B \rightarrow 01, C \rightarrow 10, D \rightarrow 11$  のようにして  $A, B, C, D$  の 4 つの文字を符号化した列だったことが分かった．発見された記号列は、

CABDABABAA

を表現したいものであるということになる．そこで、 $\{A, B, C, D\}$  の四つの記号が決まった確率で生成されるデータであり、確率は観測されたデータ中の出現頻度であると仮定して、平均情報量を計算し直してみた．以下の表および文の (f)-(n) を埋めなさい (値を簡単に計算できない場合は、式の形で表現する) ．なお、(n) には「と等しい」、「より大きい」、「より小さい」のうちのどれかが入るものとする．

記号	出現回数	出現確率	情報量	平均情報量
A	5	0.5	(f)	$0.4 \times (f)$
B	3	0.3	(g)	$0.3 \times (g)$
C	1	(h)	(i)	$(h) \times (i)$
D	(j)	0.1	(k)	$0.1 \times (k)$
計	(l)	1.0		(m)

(e) と (m) を比較すると、2 進数の表現で 2 文字分なので (m) が (e) のちょうど 2 倍になることを期待した．実際に計算してみたところ、(m) は (e) の 2 倍 (n) ことが分かった．