学習指導要領（2）- 知・技 - ア
学習指導要領（2）- 思・判・表 - ア
学習内容（2）- ア メディアとコミュニケーション

情報II

次の文章は、情報を圧縮する技術として、記号列を１つの数値に変換して表現する方法について説明したものである。空欄 (9)(10)(11) ～ (32) に入るもっとも適した数字を解答欄にマークしなさい。また、空欄 (33) ～ (35) に入るもっとも適した語を選択肢から選び、その番号を解答欄にマークしなさい。

4 種類の記号 A,B,C,D があるとし、これを半開区間 $[0,1)$ 内の区間に対応づけて、数値に変換して表現することにする。ただし半開区間 $[a,b)$ は、$\{x \mid a \leqq x < b,\ x$ は実数$\}$ で表わされる数の集合を示す。4 種類の記号を変換するために区間 $[0,1)$ を 4 分割し、各記号を次のように分割された各区間に対応づける。

記号	A	B	C	D
区間	$[0,0.25)$	$[0.25,0.5)$	$[0.5,0.75)$	$[0.75,1)$
区間幅	0.25	0.25	0.25	0.25

この場合、記号 A を 0 に変換したり、0.1 に変換したり、区間内の任意の数値に変換して表現できる。0 も 0.1 も区間 $[0,0.25)$ に含まれているので、どちらも元の記号 A に復元できる。また、記号 B の場合、例えば 0.25 に変換して表現できるが、0.25 を 2 進法の小数で表すと 0. (9)(10)(11) になる。

2 個以上の記号の並びを数値に変換して表現するには、次のような計算を行う。AB という記号列を変換するには、まず 1 個目の A に対して区間 $[0,0.25)$ を対応させる。2 個目の B は、この狭くなった区間 $[0,0.25)$ を最初の分割と同様に 4 分割し、その 2 番目の区間を対応させる。以下の説明では、対応付けの順序に変更はなく、区間の開始する値が小さい方から順番に A、B、C、D が対応するものとする。つまり、記号列 AA,AB,AC,AD に対応する区間は次のようになる。

記号列	AA	AB	AC	AD
区間	$[0,0.0625)$	$[0.0625,0.125)$	$[0.125,0.1875)$	$[0.1875,0.25)$
区間幅	0.0625	0.0625	0.0625	0.0625

ここで、$[0,1)$ 内の数値を 2 進法の小数のある桁数で切り捨てて表現したときに、それがどの記号列に対応するかを一意的に区別できるようにするために必要な桁数について考察する。上記のように 4 種類の記号を 2 個使って構成される記号列は、全部で (12)(13) とおりあるので、少なくとも 2 進法の小数第 (14) 位までが必要になる。

4種類の記号 A,B,C,D の出現頻度が等しくない場合には、出現頻度に比例した長さの区間を用いて対応づけを行なって変換すると、色々な記号列に対する数値表現に必要な小数の桁数の期待値が小さくなり、情報を圧縮する技術として使えることが知られている。以下の説明では、2進法の小数としては循環小数になり、有限の桁数では正確に表現できない数が扱われているが、説明を簡単にするため考慮せず、10進法の小数で正確に計算が行なわれて処理されているものとする。

例えば、記号列中の記号 A,B,C,D の出現頻度が次のようになっていたとする（この表の頻度を後の説明で使うので、頻度表 X と呼ぶ)。

記号	A	B	C	D
出現頻度	0.4	0.3	0.2	0.1

これに対し、次のように出現頻度に比例した長さの区間を割り当てる。

記号	A	B	C	D
区間	[0, 0. (15))	[0. (15) , 0. (16))	[0. (16) , 0. (17))	[0. (17) , 1)
区間幅	0.4	0. (18)	0.2	0.1

ここで、BAC の 3 個の記号からなる記号列を数値に変換して表現することを考える。1個目の B には、区間 [0. (15) , 0. (16)) が対応する。2個目までを含む BA はこの区間を同様に分割して、区間 [0. (15) , 0. (19) (20) (21) (22)) が対応する。また、3個目までを含む BAC は、区間 [0. (23) (24) (25) (26) , 0. (27) (28) (29) (30)) が対応する。

次に、この頻度表 X を使って変換された数値を記号列に戻す手順について考える。数値 0.8 が与えられたとする。この値は、区間 [0. (31) , 0. (32)) に含まれるので、最初の記号は (33) であることが分かる。同様に考えて、2個目まで元の記号列に戻すと、 (33) (34) が得られ、3個目まで元の記号列に戻すと、 (33) (34) (35) が得られる。ここから分かるように、この戻す手順は無限に続けることができるので、実際に利用するには、元々何個の記号を数値に変換したのかが分かるようにする、あるいは終端を示す記号を追加する、などの工夫が必要になる。この他にも、有限桁数での計算にとどめる工夫や、小数の計算にともなう丸め誤差などの考慮が必要になる。

【 (33) ～ (35) の選択肢】

(1) A　(2) B　(3) C　(4) D