

問題 2 次のハフマン符号に関する記述を読み、に入れるべき適切な字句を解答群から選べ。

[ハフマン符号について]

画像データや動画データ、音声データなどは、文字データに比べ多くのデータ量を必要とする。そのままでは、ネットワーク上で伝送する時の通信回線や保存する媒体に負担をかけてしまう。そこで、通常はデータの圧縮を行い、元のデータ量より少ないデータ量で利用する。圧縮方法のひとつに、ハフマン符号がある。

なお、この問題では、あらかじめデータの出現数を把握して符号化する静的ハフマン符号を用いる。

例えば、表 1 のような出現数の場合を考える。

表 1 データの出現数

データ	出現数
A	20
B	10
C	40
D	30

これらのデータを木構造の節を用意して格納する。なお，“:”に続く数値は出現数である。

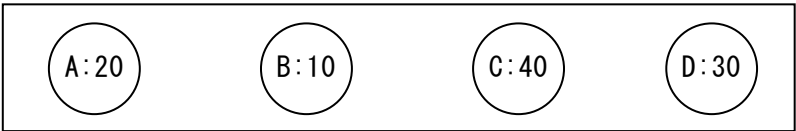


図 1 出現数を節に格納する

出現数の最も小さい節と 2 番目に小さい節をつなぐ節を追加する。つないだ節には、最も小さい節と 2 番目に小さい節に格納された出現数を合計した値を格納する。

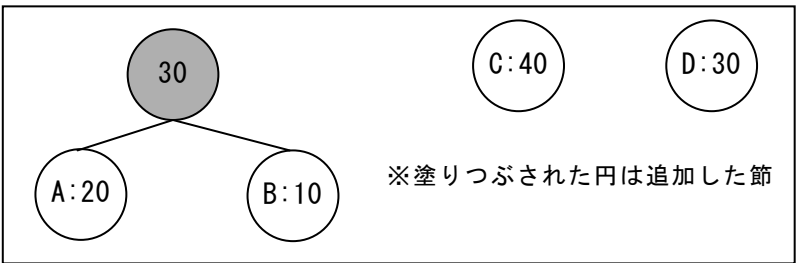


図 2 最も小さい節と 2 番目に小さい節をつなぐ節を作る

残った節と新しく追加した節の出現数の中から、同様に最も小さい値と2番目に小さい値を新しい節でつなぎ、その出現数の合計を新しい節に格納する。

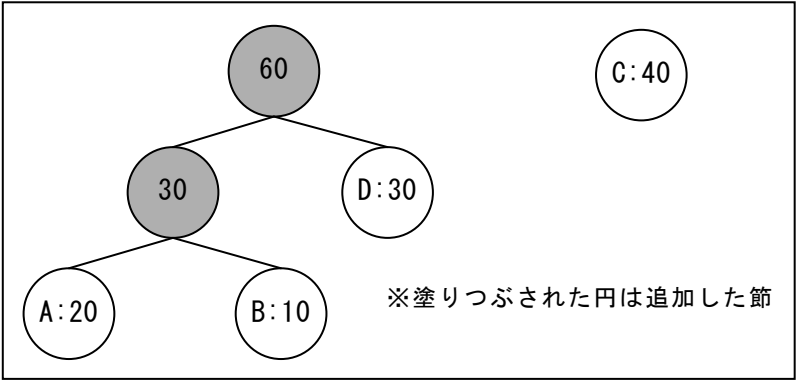


図3 さらに値の小さい2つの節をつなぐ節を作る

この動作を節が1つになるまで繰り返すとハフマン木が完成する。節と節をつなぐ枝に“0”と“1”を割り当て、枝をたどることでデータを符号化する。なお、左右のどちらかを“0”，“1”にするかは任意であるが、この問題では左を“0”，右を“1”とする。

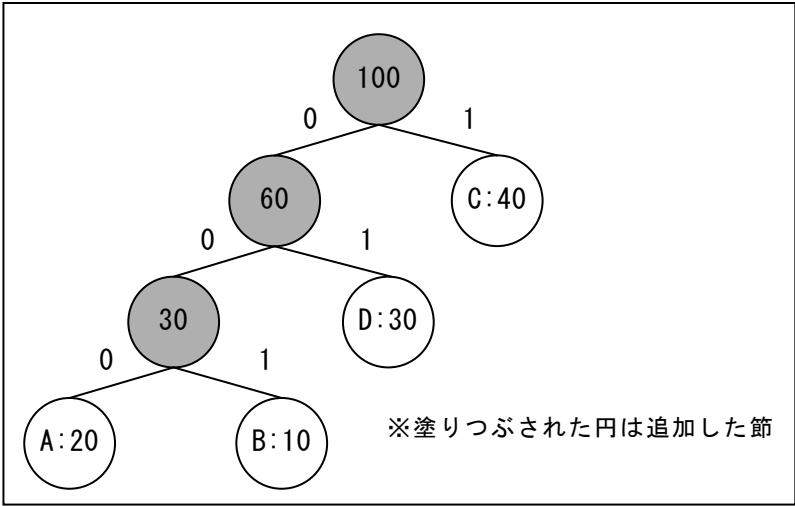


図4 完成したハフマン木

ハフマン木の枝をたどることで符号化した可変長ビット列が決定する。この場合は、次の表2のようになる。

表2 符号化したビット列

データ	ビット列
A	000
B	001
C	1
D	01

元のデータを符号化したビット列で置き換えることで圧縮が完了する。

[テキストファイルの圧縮]

あるテキストファイルをハフマン符号化し圧縮することを考える。

このテキストファイルはシフト JIS コードで保存されており，アルファベット "A"～"F"の6種類の半角文字（8ビット）で記録されている。記録されている文字数を数えたところ，表3のようになった。

表3 出現する文字数

データ（文字）	出現数
A	50
B	15
C	20
D	5
E	30
F	20

この表を基に，図5のようにハフマン木を作成した。

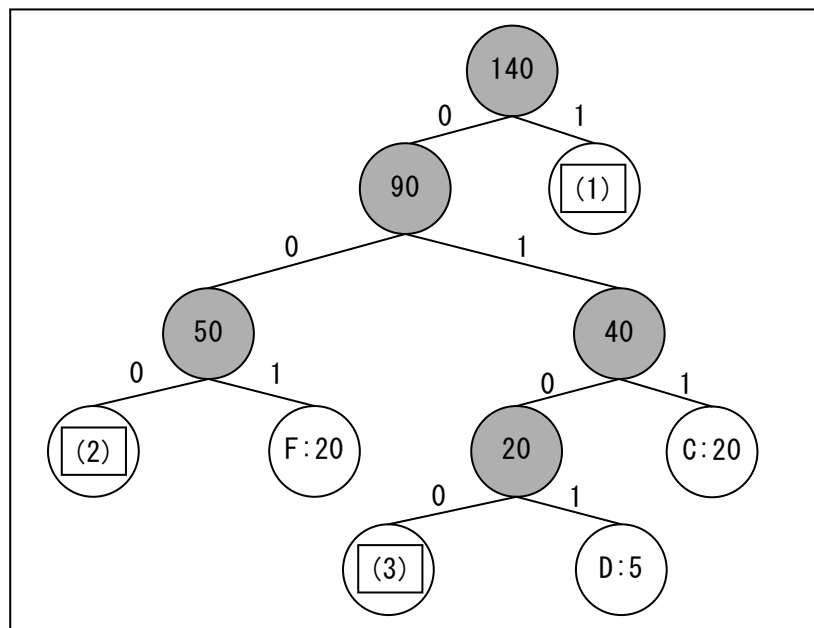


図5 ハフマン木

図5より、各データの値を符号化したビット列は、表4のようになる。

表4 出現する文字数

データ（文字）	符号化したビット列
A	1
B	0100
C	011
D	(4)
E	000
F	(5)

このことから、例えば、元のデータが“ACE”であれば、24ビット（3文字×8ビット）だったものが、圧縮すると(6)の7ビットになる。

また、このテキストファイルの元のデータ容量は(7)ビットであるが、符号化したビット列にすべて置き換えると(8)ビットになる。符号化する前に比べると、約(9)%の容量に圧縮されたことになる。ただし、1バイトは8ビットとし、制御情報は含めないものとする。

(1) ～ (3) の解答群

- | | | |
|---------|---------|---------|
| ア. A:50 | イ. B:15 | ウ. C:20 |
| エ. D:5 | オ. E:30 | カ. F:20 |

(4) , (5) の解答群

- | | | |
|---------|---------|---------|
| ア. 000 | イ. 001 | ウ. 010 |
| エ. 0100 | オ. 0101 | カ. 0111 |

(6) の解答群

- | | | | |
|------------|------------|------------|------------|
| ア. 0000111 | イ. 0001011 | ウ. 1000011 | エ. 1011000 |
|------------|------------|------------|------------|

(7) , (8) の解答群

- | | | |
|--------|---------|---------|
| ア. 140 | イ. 320 | ウ. 340 |
| エ. 360 | オ. 1120 | カ. 1400 |

(9) の解答群

- | | | | |
|-------|-------|-------|-------|
| ア. 24 | イ. 30 | ウ. 41 | エ. 63 |
|-------|-------|-------|-------|