Question 3 Morse Code

The BIO has been receiving telegrams congratulating it on reaching its 10th anniversary. At least, we think it has. The telegrams have been sent in *Morse code* and, unfortunately, the gaps between letters have been left out.

In Morse code, each letter of the alphabet is replaced by a sequence of dots and dashes as follows:

a	. –	h ·	• • •	o		\mathbf{v}	
b		i ·	•	p		w	
c		j .		q		X	
d		k -	. –	r		y	
e	•	1 .		S	• • •	Z	
f		m -	_	t	-		
g		n -		u			

Every combination of between 1 and 4 dots and dashes is used, except for:

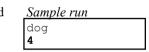
.-.-

Traditionally, dots were transmitted by a short note and dashes by a longer note, with pauses between different letters. This is why some mobile phones make the sound . . . -- . . . when receiving a message, since this is the Morse code for SMS.

If the gaps between letters are missed out, messages can be ambiguous. For example, even if we know the message -..---. is made up of three letters, it might mean **njg**, **dog**, **xmg** or **xon**.

3 (a) [26 marks]

Write a program which reads in a message (between 1 and 10 letters inclusive) and determines how many messages, with the same number of letters as the input, it might represent.



3 (b) [5 marks]

How many messages might ---- represent, if we do not know the number of letters in the message? How about -..---?

3 (c) [3 marks]

It is possible to come up with new ways of encoding the alphabet so that, even when the gaps between letters are missing, messages are unambiguous. The *size* of such an unambiguous encoding is the total number of dots and dashes in a message containing each letter once.

For example, we could encode each letter by some dots (indicating its position in the alphabet) followed by a dash; so. – would be **a**, ... – would be **b**, and 26 dots followed by a dash would be **z**. This encoding has a size of 377 (2 + 3 + ... + 27).

What is the smallest size an unambiguous encoding can have?

Total marks: 100

End of BIO 2004 Round One paper