

## Assignment 2

The programme fully meets the specification and implements both the monoalphabetic and the vigenere cipher.

The programme performs several validations for both the keyword and the filename inputs. For the keyword, it is checked that the string is not empty. Furthermore, it is checked that the input only contains letters and that none of the letters are repeated. In any of these cases, an appropriate JOptionPane error message is shown and the text fields are cleared. The design decision was made to allow lower-case letters. In this case, the lower-case letters are simply converted to upper case letter before they are used to form the cipher.

Regarding the filename input, it is checked that the string is not empty. Furthermore, the filename is validated to ensure that it ends in either P or C. Again, in each of these cases an appropriate JOptionPane error message is shown and the text fields are cleared. This is also the case if there is a file handling error, such as the input file not being found.

There is no validation on the length of the string in the filename text field and it is assumed that a valid string length will be entered. Due to the nature of the keyword, the maximum string length is 26 unique letters.

It is assumed that the user only wishes to encode upper-case letters. For any invalid characters in the input file, such as lower-case letters, spaces and comas; the character is simply not encoded/decoded and will remain the same in the output file.

The programme also fully implements the letter frequencies functionality. A report is outputted for each encode/decode operation. Examples of these can be seen in appendix 1.

Several sets of data were used to test the programmes functionality. A selection of the invalid test data can be seen below in table 1. Where appropriate, a reference is made to the corresponding figure number of the error message that is produced, these can be seen in appendix 2. A small selection of the valid data that was tested can be seen in table 2 and in the screen dumps in Appendix 1.

*Table 1. Invalid test Data*

Scenario	Keyword	Filename	Cipher	Plain text	Result
Repeating letter in keyword	TREE	messageP	Mono	REPEATED	Error message (13)
Empty keyword		messageP	Vigenere	NO KEYWORD	Error message (14)
Invalid characters in keyword	H,L;	messageP	Mono	WRONG KEYWORD	Error message (15)
Empty filename	BUCK		Vigenere	NO FILENAME	Error message (16)
Filename does not end in C or P	SING	message	Mono	WRONG FILENAME	Error message (17)
File handling error (File does not exist)	SYMBOL	otherP	Vigenere		Error message (18)

*Table 2. Valid test data*

Scenario	Keyword	Filename	Cipher	Plain text	Encrypted	Decrypted	Result
Correct data	FIG	messageP/ messageC	Mono	GOOD DATA	WOOZ ZFJF	GOOD DATA	Success
Correct data	FRIEND	messageP/ messageC	Vigenere	EVEN BETTER DATA	JMMR OHYKMV QDYR	EVEN BETTER DATA	Success

## Appendix 1 - Screen Dumps

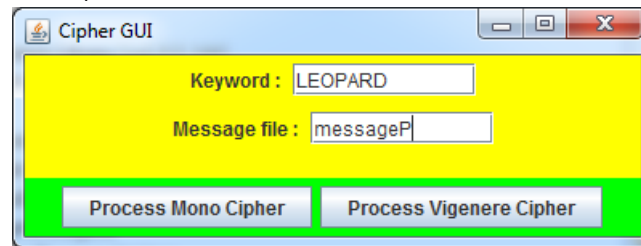


Figure 1. GUI with valid inputs

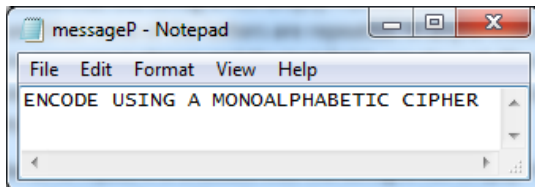


Figure 2. The un-encoded message

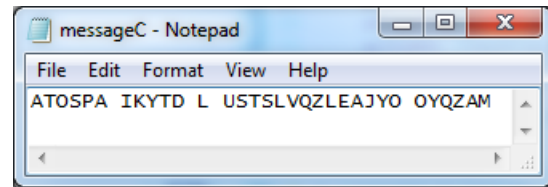


Figure 3. The encoded message

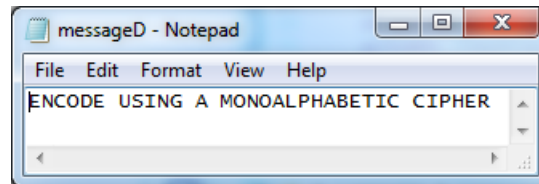


Figure 4. The decoded message

Letter	Freq	Freq%	AvgFreq%	Diff
A	4	12.5	8.2	4.3
B	0	0.0	1.5	-1.5
C	0	0.0	2.8	-2.8
D	1	3.1	4.3	-1.2
E	1	3.1	12.7	-9.6
F	0	0.0	2.2	-2.2
G	0	0.0	2.0	-2.0
H	0	0.0	6.1	-6.1
I	1	3.1	7.0	-3.9
J	1	3.1	0.2	2.9
K	1	3.1	0.8	2.3
L	3	9.4	4.0	5.4
M	1	3.1	2.4	0.7
N	0	0.0	6.7	-6.7
O	3	9.4	7.5	1.9
P	1	3.1	1.9	1.2
Q	2	6.3	0.1	6.2
R	0	0.0	6.0	-6.0
S	3	9.4	6.3	3.1
T	3	9.4	9.1	0.3
U	1	3.1	2.8	0.3
V	1	3.1	1.0	2.1
W	0	0.0	2.4	-2.4
X	0	0.0	0.2	-0.2
Y	3	9.4	2.0	7.4
Z	2	6.3	0.1	6.2

The most frequent letter is A at 12.5%

Figure 5. Letter frequencies for the encoded file

Letter	Freq	Freq%	AvgFreq%	Diff
A	3	9.4	8.2	1.2
B	1	3.1	1.5	1.6
C	3	9.4	2.8	6.6
D	1	3.1	4.3	-1.2
E	4	12.5	12.7	-0.2
F	0	0.0	2.2	-2.2
G	1	3.1	2.0	1.1
H	2	6.3	6.1	0.2
I	3	9.4	7.0	2.4
J	0	0.0	0.2	-0.2
K	0	0.0	0.8	-0.8
L	1	3.1	4.0	-0.9
M	1	3.1	2.4	0.7
N	3	9.4	6.7	2.7
O	3	9.4	7.5	1.9
P	2	6.3	1.9	4.4
Q	0	0.0	0.1	-0.1
R	1	3.1	6.0	-2.9
S	1	3.1	6.3	-3.2
T	1	3.1	9.1	-6.0
U	1	3.1	2.8	0.3
V	0	0.0	1.0	-1.0
W	0	0.0	2.4	-2.4
X	0	0.0	0.2	-0.2
Y	0	0.0	2.0	-2.0
Z	0	0.0	0.1	-0.1

The most frequent letter is E at 12.5%

Figure 6. Letter frequencies for the decoded file

Monoalphabetic cipher array: [L, E, O, P, A, R, D, Z, Y, X, W, V, U, T, S, Q, N, M, K, J, I, H, G, F, C, B]

Vigenere cipher arrays: [[L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z, A, B, C, D, E, F, G, H, I, J, K], [E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z, A, B, C, D], [O, P, Q, R, S, T, U, V, W, X, Y, Z, A, B, C, D, E, F, G, H, I, J, K, L, M, N], [P, Q, R, S, T, U, V, W, X, Y, Z, A, B, C, D, E, F, G, H, I, J, K, L, M, N, O], [A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z], [R, S, T, U, V, W, X, Y, Z, A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q], [D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z, A, B, C]]

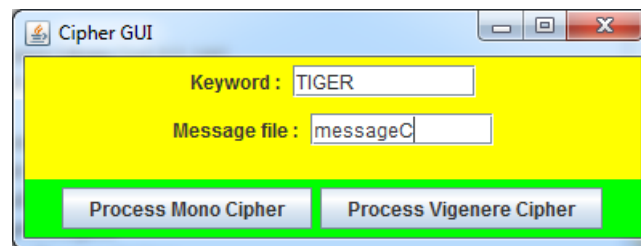


Figure 7. GUI with valid inputs

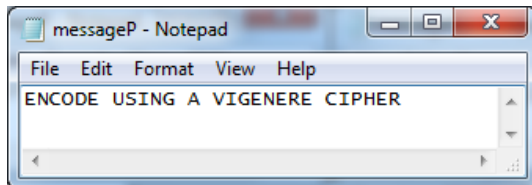


Figure 8. The un-encoded message

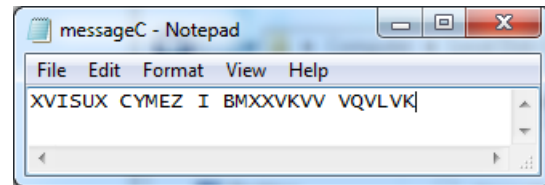


Figure 9. The encoded message

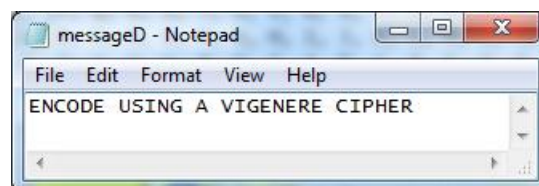


Figure 10. The decoded message

Letter	Freq	Freq%	AvgFreq%	Diff
A	0	0.0	8.2	-8.2
B	1	3.8	1.5	2.3
C	1	3.8	2.8	1.0
D	0	0.0	4.3	-4.3
E	1	3.8	12.7	-8.9
F	0	0.0	2.2	-2.2
G	0	0.0	2.0	-2.0
H	0	0.0	6.1	-6.1
I	2	7.7	7.0	0.7
J	0	0.0	0.2	-0.2
K	2	7.7	0.8	6.9
L	1	3.8	4.0	-0.2
M	2	7.7	2.4	5.3
N	0	0.0	6.7	-6.7
O	0	0.0	7.5	-7.5
P	0	0.0	1.9	-1.9
Q	1	3.8	0.1	3.7
R	0	0.0	6.0	-6.0
S	1	3.8	6.3	-2.5
T	0	0.0	9.1	-9.1
U	1	3.8	2.8	1.0
V	7	26.9	1.0	25.9
W	0	0.0	2.4	-2.4
X	4	15.4	0.2	15.2
Y	1	3.8	2.0	1.8
Z	1	3.8	0.1	3.7

The most frequent letter is V at 26.9%

Figure 11. Letter frequencies for the encoded file

Letter	Freq	Freq%	AvgFreq%	Diff
A	1	3.8	8.2	-4.4
B	0	0.0	1.5	-1.5
C	2	7.7	2.8	4.9
D	1	3.8	4.3	-0.5
E	6	23.1	12.7	10.4
F	0	0.0	2.2	-2.2
G	2	7.7	2.0	5.7
H	1	3.8	6.1	-2.3
I	3	11.5	7.0	4.5
J	0	0.0	0.2	-0.2
K	0	0.0	0.8	-0.8
L	0	0.0	4.0	-4.0
M	0	0.0	2.4	-2.4
N	3	11.5	6.7	4.8
O	1	3.8	7.5	-3.7
P	1	3.8	1.9	1.9
Q	0	0.0	0.1	-0.1
R	2	7.7	6.0	1.7
S	1	3.8	6.3	-2.5
T	0	0.0	9.1	-9.1
U	1	3.8	2.8	1.0
V	1	3.8	1.0	2.8
W	0	0.0	2.4	-2.4
X	0	0.0	0.2	-0.2
Y	0	0.0	2.0	-2.0
Z	0	0.0	0.1	-0.1

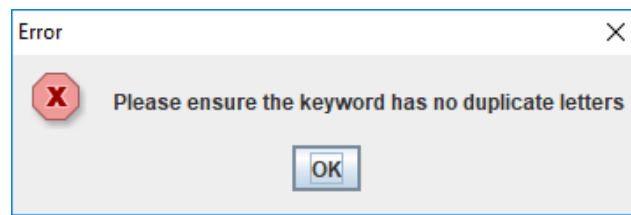
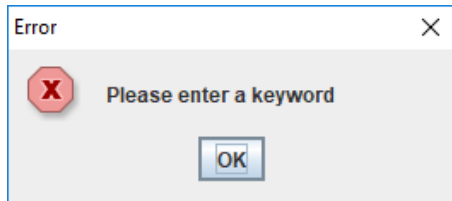
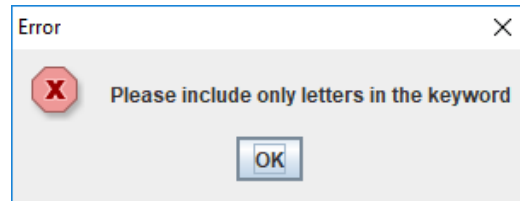
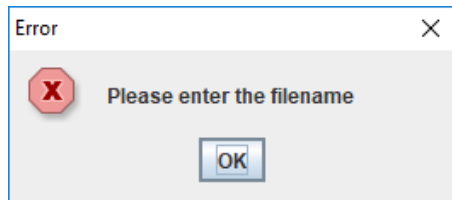
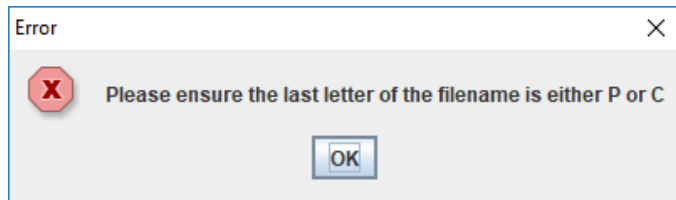
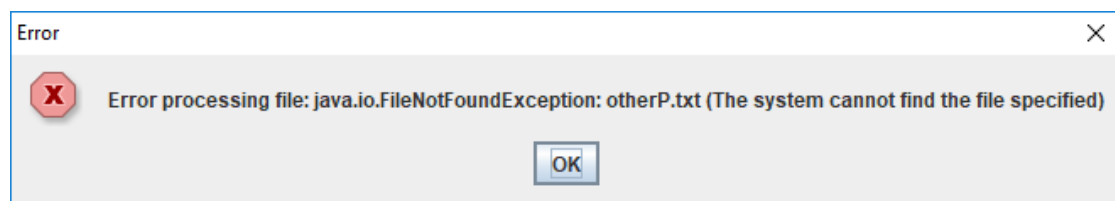
The most frequent letter is E at 23.1%

Figure 12. Letter frequencies for the decoded file

Monoalphabetic cipher array: [T, I, G, E, R, Z, Y, X, W, V, U, S, Q, P, O, N, M, L, K, J, H, F, D, C, B, A]

Vigenere cipher arrays: [[T, U, V, W, X, Y, Z, A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S], [I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z, A, B, C, D, E, F, G, H], [G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z, A, B, C, D, E, F], [E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z, A, B, C, D], [R, S, T, U, V, W, X, Y, Z, A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q]]

## Appendix 2 - Error Messages

*Figure 13. Duplicate letter error message**Figure 14. No keyword error message**Figure 15. Invalid character in keyword error message**Figure 16. No filename error message**Figure 17. Filename in wrong format error message**Figure 18. File not found error message*