Cifras de Beale

De acordo com a lenda, no início do século XIX, um homem chamado Thomas J. Beale enterrou um tesouro no valor de milhões de dólares em algum lugar da Virgínia. Beale deixou três textos cifrados com um dono de hotel, Robert Morriss, antes de desaparecer, e instruiu Morriss a abrir o pacote contendo as cifras apenas se ele não retornasse dentro de 10 anos. O primeiro texto cifrado descreve o conteúdo do tesouro, enquanto os outros dois descrevem a localização exata do local de enterro. Apesar de muitas tentativas de decodificar as cifras, seu significado permanece um mistério até hoje. A Figura 1 abaixo mostra um dos textos cifrados.

71, 194, 38, 1701, 89, 76, 11, 83, 1629, 48, 94, 63, 132, 16, 111, 95, 84, 341,
975, 14, 40, 64, 27, 81, 139, 213, 63, 90, 1120, 8, 15, 3, 126, 2018, 40, 74,
758, 485, 604, 230, 436, 664, 582, 150, 251, 284, 308, 231, 124, 211, 486, 225,
401, 370, 11, 101, 305, 139, 189, 17, 33, 88, 208, 193, 145, 1, 94, 73, 416,
918, 263, 28, 500, 538, 356, 117, 136, 219, 27, 176, 130, 10, 460, 25, 485, 18,
436, 65, 84, 200, 283, 118, 320, 138, 36, 416, 280, 15, 71, 224, 961, 44, 16, 401,
39, 88, 61, 304, 12, 21, 24, 283, 134, 92, 63, 246, 486, 682, 7, 219, 184, 360, 780,
18, 64, 463, 474, 131, 160, 79, 73, 440, 95, 18, 64, 581, 34, 69, 128, 367, 460, 17,
81, 12, 103, 820, 62, 116, 97, 103, 862, 70, 60, 1317, 471, 540, 208, 121, 890,
346, 36, 150, 59, 568, 614, 13, 120, 63, 219, 812, 2160, 1780, 99, 35, 18, 21, 136,
872, 15, 28, 170, 88, 4, 30, 44, 112, 18, 147, 436, 195, 320, 37, 122, 113, 6, 140,
8, 120, 305, 42, 58, 461, 44, 106, 301, 13, 408, 680, 93, 86, 116, 530, 82, 568, 9,
102, 38, 416, 89, 71, 216, 728, 965, 818, 2, 38, 121, 195, 14, 326, 148, 234, 18,
55, 131, 234, 361, 824, 5, 81, 623, 48, 961, 19, 26, 33, 10, 1101, 365, 92, 88, 181,
275, 346, 201, 206, 86, 36, 219, 324, 829, 840, 64, 326, 19, 48, 122, 85, 216, 284,
919, 861, 326, 985, 233, 64, 68, 232, 431, 960, 50, 29, 81, 216, 321, 603, 14, 612,
81, 360, 36, 51, 62, 194, 78, 60, 200, 314, 676, 112, 4, 28, 18, 61, 136, 247, 819,
921, 1060, 464, 895, 10, 6, 66, 119, 38, 41, 49, 602, 423, 962, 302, 294, 875, 78,
14, 23, 111, 109, 62, 31, 501, 823, 216, 280, 34, 24, 150, 1000, 162, 286, 19, 21,
17, 340, 19, 242, 31, 86, 234, 140, 607, 115, 33, 191, 67, 104, 86, 52, 88, 16, 80,
121, 67, 95, 122, 216, 548, 96, 11, 201, 77, 364, 218, 65, 667, 890, 236, 154, 211,
10, 98, 34, 119, 56, 216, 119, 71, 218, 1164, 1496, 1817, 51, 39, 210, 36, 3, 19,
540, 232, 22, 141, 617, 84, 290, 80, 46, 207, 411, 150, 29, 38, 46, 172, 85, 194,
39, 261, 543, 897, 624, 18, 212, 416, 127, 931, 19, 4, 63, 96, 12, 101, 418, 16, 140,
230, 460, 538, 19, 27, 88, 612, 1431, 90, 716, 275, 74, 83, 11, 426, 89, 72, 84,
1300, 1706, 814, 221, 132, 40, 102, 34, 868, 975, 1101, 84, 16, 79, 23, 16, 81, 122,
324, 403, 912, 227, 936, 447, 55, 86, 34, 43, 212, 107, 96, 314, 264, 1065, 323,
428, 601, 203, 124, 95, 216, 814, 2906, 654, 820, 2, 301, 112, 176, 213, 71, 87, 96,
202, 35, 10, 2, 41, 17, 84, 221, 736, 820, 214, 11, 60, 760.

Figura 1: Texto cifrado de Beale

Acredita-se que os textos cifrados foram criados usando um **livro cifra**, que é um tipo de criptografia que usa um livro ou texto pré-acordado como chave. Em um livro cifra, a mensagem em texto simples é primeiro convertida em números, e então cada número é usado para referenciar uma palavra ou frase específica no livro. A sequência resultante de palavras ou frases se torna o texto cifrado, como o exemplo da Figura 1. (Para mais detalhes, https://en.wikipedia.org/wiki/Beale_ciphers)

Gerando as cifras

Beale usou uma variante de um livro cifra. Ele escolheu um texto longo como **texto cifra**, numerou cada uma das palavras do texto sequencialmente, começando em 0 (zero), e formou uma nova mensagem em que cada caracter desta mensagem é a primeira letra de alguma palavra do texto cifra. A mensagem codificada final consiste em uma lista dos números de sequência das palavras escolhidas.

Considere por exemplo o seguinte texto como **livro cifra**:

```
"Em 1892 o intelectual paranaense José Francisco da Rocha Pombo colocaria, no Largo
Ouvidor Pardinho, a pedra fundamental da Universidade do Paraná. O projeto foi
frustrado pelo Movimento Federalista que impediu a
criação da universidade.Vinte anos depois, em 1912, o estado contava com um
reduzido número de intelectuais,
apenas nove médicos e quatro engenheiros, mas se desenvolvia muito devido a
produção da erva-mate. Além
disso, a Guerra do Contestado fez com que as vastas lideranças políticas se
empenhassem pela criação de
uma universidade, de modo a dar uma identidade ao povo paranaense."
```

Usando a ideia de Beale, (considerando que apenas **espaço** e **linefeed** (mudança de linha) são usados para separar palavras) teríamos **arquivo de chaves** abaixo, em que cada linha contém 2 campos separados por **':'** . O segundo campo contém uma lista da posição de palavras cuja primeiro caracter é o caracter indicado no primeiro campo.

```
1: 38 1
a: 89 85 72 65 63 59 48 35 31 15
c: 79 70 68 42 41 32 10
d: 86 83 80 67 64 61 58 56 46 36 33 20 18 7
e: 77 62 53 51 40 37 0
f: 69 28 25 24 17 6
g: 66
i: 88 47 30 3
j: 5
l: 74 12
m: 84 57 54 50 27
n: 49 45 11
o: 39 22 13 2
p: 91 90 78 75 60 26 23 21 16 14 9 4
q: 71 52 29
r: 44 8
s: 76 55
u: 87 82 81 43 34 19
v: 73
```

Usando os códigos acima, a frase **"casa de papel"** poderia ser codificada da seguinte maneira:

```
42 15 76 85 -1 46 51 -1 91 48 75 77 12
```

Ou ainda:

```
79 85 55 72 -1 64 62 -1 90 89 90 0 12
```

Note que no exemplo acima, algumas letras são representadas por vários códigos. Quanto maior o livro cifra, mais cchaves serão atribuídos para a mesma letra, aumentando assim o número de representações possíveis do texto cifrado. Ainda no exemplo acima, vale mencionar que o espaço entre as palavras foi representado pelo código -1.

Implementação do Trabalho

Escreva um programa em C que codifique e decodifique uma mensagem usando a cifra de Beale. O programa deverá ter as seguintes funcionalidades:

- Codificar uma mensagem qualquer contida em um arquivo ASCII usando um **livro cifra**. O programa deve ter uma opção de salvar em um arquivo texto as chaves geradas no formato descrito anteriormente. A linha de execução do programa dever ser a seguinte:

```
./beale  -e  -b LivroCifra -m MensagemOriginal -o MensagemCodificada -c
ArquivoDeChaves
```

- Decodificar uma mensagem, usando um arquivo de códigos

```
./beale  -d  -i MensagemCodificada  -c ArquivoDeChaves  -o MensagemDecodificada
```

- Decodificar uma mensagem usando o livro cifra

```
./beale -d -i MensagemCodificada -b LivroCifra -o MensagemDecodificada
```

- Um exemplo de livro cifra pode ser encontrado aqui

Requisitos

Tendo em vista que um caractere pode ter um número variável de chaves (dependendo do tamanho do livro cifra), você deve armazenar as chaves numa lista, a qual deve ser alocada dinamicamente. Para o mesmo livro cifra e mensagem, o programa deve fornecer mensagens codificadas a cada execução.

Estrutura do Código Fonte

O projeto deve ter, além do arquivo `beale.c` (que contém o código da função main), pelo menos três arquivos `.c` com as funções para gerar o arquivo de chaves, codificação e decodificação de uma mensagem. As estruturas de dados necessárias devem estar definidas em um arquivo `.h`

Produto a ser entregue

Deve ser entregue ao professor um arquivo *.tar* ou *.zip* contendo pelo menos:

- Arquivos `.c` e `.h`
- Arquivo `Makefile`. Este arquivo para o projeto deve ter pelo menos:
  - Os alvos `all` (default), `clean` e `purge`.
  - CFLAGS = `-std=c99 -Wall`
  - ATENÇÃO: Deve ser OBRIGATORIAMENTE usada a opção de compilação `-std=c99`
  - Compilar e ligar separadamente (gerar arquivos `.o` intermediários)
- Arquivo `LEIAME` contendo nome e GRR do aluno, texto explicando resumidamente os módulos criados, as estruturas de dados e os algoritmos usados.
- Estes arquivos devem estar em um diretório de nome **<login>**, onde **<login>** é o login do usuário nos sistemas linux do DINF.

- O nome do arquivo a ser entregue deve ser **<login>.tar** ou **<login>.zip**
- Os trabalhos devem ser entregues através do Moodle C3SL:
    - Turma BCC1 (Prof. Luiz Oliveira)
    - Turma BCC2 (Prof. Armando)
    - Turma BCC3 (Prof. Vinícius)

Avaliação

Os itens de avaliação do trabalho e respectivas pontuações são:

- Modularização e organização do código-fonte (15 pontos)
- Funcionamento: corretude das respostas nos testes executados (40 pontos)
- Eficiência: algoritmos e estruturas de dados utilizados para obter um melhor desempenho e uso eficiente de alocação dinâmica de memória (45 pontos)

ATENÇÃO: programas que tiverem erros de compilação ou terminarem a execução de forma abrupta sem que tenha havido processamento adequado receberão nota **ZERO**