A PROGRAMOZÁS ALAPJAI 1
(BMEVIEEAA00, 2024/25/1)
NAGYHÁZI FELADAT

Shanon-Fano kódoló és dekódoló program

készítette:

Ferencz Péter

(RFG7SN)

Budapesti Műszaki és Gazdaságtudományi Egyetem

Villamosmérnöki és Informatikai Kar

Mérnökinformatikus Bsc

2024 Október

1. Specifikáció **1**

1.1. A program célja . . . 1

1.2. Felhasználói interakció . . . 1

1.2.1. kódolás (kodol) . . . 1

1.2.2. dekódolás (dekodol) . . . 1

1.3. A program által elfogadott kapcsolók . . . 2

1.3.1. Bemenet . . . 2

1.3.2. Kimenet . . . 2

1.3.3. Kódtábla . . . 2

1.3.4. Statisztika . . . 2

1.3.5. Segítség . . . 3

1.4. A program kimenete . . . 3

2. Adatszerkezet-mutató **5**

2.1. Adatszerkezetek . . . 5

3. Fájlmutató **7**

3.1. Fájllista . . . 7

4. Adatszerkezetek dokumentációja **9**

4.1. Bits struktúrareferencia . . . 9

4.1.1. Részletes leírás . . . 9

4.1.2. Adatmezők dokumentációja . . . 9

4.1.2.1. b . . . 9

4.1.2.2. length . . . 10

4.2. CodeWord struktúrareferencia . . . 10

4.2.1. Részletes leírás . . . 10

4.2.2. Adatmezők dokumentációja . . . 10

4.2.2.1. bits . . . 11

4.2.2.2. codeWord . . . 11

4.3. codewordFrequency struktúrareferencia . . . 11

4.3.1. Adatmezők dokumentációja . . . 12

4.3.1.1. codeWord . . . 12

4.3.1.2. freq . . . 12

4.4. commandLineArguments struktúrareferencia . . . 12

4.4.1. Részletes leírás . . . 12

4.4.2. Adatmezők dokumentációja . . . 12

4.4.2.1. displayStatistics . . . 13

4.4.2.2. displayTable . . . 13

4.4.2.3. infile . . . 13

4.4.2.4. outfile . . . 13

4.5. InputFileBuffer struktúrareferencia . . . 13

4.5.1. Részletes leírás . . . 14

4.5.2. Adatmezők dokumentációja . . . 14
4.5.2.1. currentBit . . . 14
4.5.2.2. file . . . 14
4.6. Node struktúrareferencia . . . 14
4.6.1. Adatmezők dokumentációja . . . 15
4.6.1.1. codeword . . . 15
4.6.1.2. left_0 . . . 15
4.6.1.3. right_1 . . . 15
4.7. OutputFileBuffer struktúrareferencia . . . 15
4.7.1. Részletes leírás . . . 16
4.7.2. Adatmezők dokumentációja . . . 16
4.7.2.1. bits . . . 16
4.7.2.2. file . . . 16

5. Fájlok dokumentációja **17**
5.1. src/bin.c fájlreferencia . . . 17
5.1.1. Függvények dokumentációja . . . 18
5.1.1.1. bits_pushBit() . . . 18
5.1.1.2. getBitFromRight() . . . 19
5.1.1.3. print_bits() . . . 19
5.2. src/cla.c fájlreferencia . . . 19
5.3. src/codeword.c fájlreferencia . . . 20
5.3.1. Típusdefiníciók dokumentációja . . . 22
5.3.1.1. uchar . . . 22
5.3.2. Függvények dokumentációja . . . 22
5.3.2.1. bits_equ() . . . 22
5.3.2.2. isNullbit() . . . 22
5.3.3. Változók dokumentációja . . . 23
5.3.3.1. NULLBIT . . . 23
5.4. src/debug.h fájlreferencia . . . 23
5.4.1. Makródefiníciók dokumentációja . . . 24
5.4.1.1. PRINTDEBUG_CORRUPTEDFILE . . . 24
5.4.1.2. PRINTDEBUG_CUSTOM . . . 24
5.4.1.3. PRINTDEBUG_FILEERR . . . 24
5.4.1.4. PRINTDEBUG_MALLOCNULL . . . 24
5.5. src/decoder.c fájlreferencia . . . 25
5.5.1. Függvények dokumentációja . . . 26
5.5.1.1. appendCodeword() . . . 26
5.5.1.2. createNodeIfNotexists() . . . 26
5.5.1.3. decode() . . . 27
5.5.1.4. freeTree() . . . 27
5.6. src/encoder.c fájlreferencia . . . 27

5.6.1. Függvények dokumentációja . . . 29
5.6.1.1. codewordToBits() . . . 29
5.6.1.2. compare_by_bitlength() . . . 29
5.6.1.3. compare_by_freq() . . . 29
5.6.1.4. encode() . . . 31
5.6.1.5. setCodeWord() . . . 31
5.7. src/fileBuffer.c fájlreferencia . . . 31
5.7.1. Függvények dokumentációja . . . 33
5.7.1.1. buff_createInputFileBuffer() . . . 33
5.7.1.2. buff_createOutputFileBuffer() . . . 33
5.7.1.3. buff_destroyInputFileBuffer() . . . 33
5.7.1.4. buff_destroyOutputFileBuffer() . . . 33
5.7.1.5. buff_flush() . . . 34
5.7.1.6. buff_readBit() . . . 34
5.7.1.7. buff_readBits() . . . 34
5.7.1.8. buff_readChar() . . . 34
5.7.1.9. buff_readInt() . . . 34
5.7.1.10. buff_rewind() . . . 34
5.7.1.11. buff_writeBit() . . . 34
5.7.1.12. buff_writeBits() . . . 35
5.7.1.13. buff_writeChar() . . . 35
5.7.1.14. buff_writeInt() . . . 35
5.8. src/graph.c fájlreferencia . . . 35
5.9. src/main.c fájlreferencia . . . 36
5.9.1. Enumerációk dokumentációja . . . 37
5.9.1.1. MODE . . . 37
5.9.2. Függvények dokumentációja . . . 37
5.9.2.1. main() . . . 37
5.9.2.2. parseCLA() . . . 37
5.9.2.3. printHelp() . . . 38

Meta **39**

5.10. Források, felhasznált irodalom . . . 39
5.11. Felhasznált segédprogramok . . . 39

1. fejezet

Specifikáció

1.1. A program célja

A program célja tetszőleges adat tömörítése majd ezek kitömörítése információvesztés nélkül. Ennek megvalósítására a Shanon-Fano tömörítő algoritmust [1] [2] alkalmazza.

1.2. Felhasználói interakció

A felhasználó két üzemmódot választhat ki a program futtatásakor: kódolás vagy dekódolás. Ezeket az első parancssori argumentumban a 'kodol' és 'dekodol' kulcsszavakkal tudja kiválasztani.

1.2.1. kódolás (kodol)

Kódoló üzemmódban a bemenetet (lásd Bemenet) a Shanon-Fano kódoló algoritmust alkalmazva írja a kimenetre (lásd Kimenet) a kódolt adatot.

```
program kodol --bemenet <fájl> --kimenet <fájl>
```

1.2.2. dekódolás (dekodol)

Dekódoló üzemmódban a bemenetet (lásd Bemenet) a Shanon-Fano dekódoló algoritmust alkalmazva írja a kimenetre (lásd Kimenet) a dekódolt adatot.

```
program dekodol --bemenet <fájl> --kimenet <fájl>
```

[1] C. E. Shannon, „A Mathematical Theory of Communication", 1948
[2] Robert M. Fano, „The Transmittion of Information", 1949

1.3. A program által elfogadott kapcsolók

A program futása során tetszőleges futtatást befolyásoló kapcsolókat (flageket) beállíthatunk. Ezek sorrendje tetszőlegesen választható.

1.3.1. Bemenet

Parancssori megnevezés: `--bemenet <forrásfájl>`
Opcionális paraméter.
Ha nincs megadva, de a program egy figyelmeztető üzenet kíséretében folytatja a lefutást.
A fájl méretétől és tartalmától független a program lefutása.
Az azt követő paraméter megadja a forrásfájl elérési útvonalát. Ha nincs megadva, stdin-ról kér be új sorral lezárt szöveget.

1.3.2. Kimenet

Parancssori megnevezés: `--kimenet <célfájl>`
Opcionális paraméter.
Ha nincs megadva, de a program egy figyelmeztető üzenet kíséretében folytatja a lefutást.
A fájl méretétől és tartalmától független a program lefutása.
Az azt követő paraméter megadja a célfájl elérési útvonalát. Ha nincs megadva, stdout-ra írja ki a program a program kimenetét.

1.3.3. Kódtábla

Parancssori megnevezés: `--kodtabla`
Opcionális paraméter.
Azt szabályozza, hogy a kódtáblát kiírja-e a program a standard kimenetre.

1.3.4. Statisztika

Parancssori megnevezés: `--statisztika`
Opcionális paraméter.
Azt határozza meg, hogy a program kiírjon-e további számitásokat a program hatékonyságára vonatkozólag.
Az alábbi számítások történnek kiírásra:

- Tömörítés mértéke: bemenet mérete a tömörített adat méretéhez képest
- Kódtábal mérete: Elgymástól eltérő kódok száma
- Kódok mérete: legrövidebb kód, leghosszabb kód, kódok átlagos mérete
- Fa mérete: A generált fa mérete

1.3.5. Segítség

Parancssori megnevezés: `--help`
Opcionális paraméter.
A felhasználót tájékoztatja a program helyes használatáról. Ha ez a kapcsoló meg van adva, akkor a program nem ellenőrzi a többi kötelező kapcsoló jelenlétét, kiírja a szöveget majd kódolás / dekódolás nélkül befejezi a futást.
Az alábbi szöveg íródik ki:

```
program [üzemmód] <...kapcsolók...>
Üzemmód: kodol, dekodol
Kapcsolók:
--bemenet <forrásfájl>: Bemeneti fájl (ha üres akkor stdin)
--kimenet <célfájl>: Bemeneti fájl (ha üres akkor stdout)
--kodtabla <fájl>: A kódtábla fájl (kötelező)
--statisztika: A tömörítés hatékonyságát értékelő statisztika (opcionális)
--help: Ezt az üzenetet írja ki (opcionális)
```

1.4. A program kimenete

Sikeres futtatás esetén a program a A program által elfogadott kapcsolók pontban meghatározott viselkedés szerint működik. Sikertelen futtatás esetén a konzolra kiíródik a probléma és egy nem nullás kilépési kóddal a program megáll.
A fájl ami generálódik a következőképpen épül fel: Kódtábla karaktereinek száma: Hány darab karaktert és annak kódolását tartalmaz a kódtábla. Lehetséges értékei: 0-255 –> 1-256
Illeszkedés hossza: A fájl végén hány darab 0 bit van a 8 bites fájlmentés kielégítéséhez.
Kódtábla, melynek minden eleme az alábbiakból épül fel:

- Karakter: nyolc bit, melyet tömörítünk
- a karaktert reprezentáló kód hossza 8 biten
- a kód, mely nullás és eggyesek sorozata

Kódolt adat

<table>
<tr><td rowspan="2">Kódolt karakterek hossza
(1-256)</td><td rowspan="2">Illeszkedés
hossza
(0-7)</td><td colspan="4">Kódtábla</td><td rowspan="2">Kódolt adat</td><td rowspan="2">Illeszkedés
(0)</td></tr>
<tr><td>karakter
ASCII</td><td>karakterkód
hossza</td><td>kód</td><td>. . . . .</td></tr>
<tr><td></td><td></td><td>8 bit</td><td>n = 8 bit</td><td>n bit</td><td></td><td></td><td></td></tr>
<tr><td>l = 8 bit</td><td>i = 3 bit</td><td colspan="4">legalább l * (8 + 8 + 1) bit</td><td>legalább l bit</td><td>i bit</td></tr>
</table>

2. fejezet

Adatszerkezet-mutató

2.1. Adatszerkezetek

Az összes adatszerkezet listája rövid leírásokkal:

Bits
: Tetszőleges hosszú bitsorozat eltárolására alkalmas struktúra 9

CodeWord
: Karakter, és az azt kódoló bitsorozat 10

codewordFrequency 11

commandLineArguments
: A program parancssori argumentumait rendező struktúra 12

InputFileBuffer
: Struktúra, mely lehetővé teszi a bitenkénti olvasást egy fájlból 13

Node 14

OutputFileBuffer
: Struktúra, mely lehetővé teszi a bitenkénti írást egy fájlba 15

3. fejezet

Fájlmutató

3.1. Fájllista

Az összes fájl listája rövid leírásokkal:

src/bin.c 17
src/cla.c 19
src/codeword.c 20
src/debug.h 23
src/decoder.c 25
src/encoder.c 27
src/fileBuffer.c 31
src/graph.c 35
src/main.c 36

4. fejezet

Adatszerkezetek dokumentációja

4.1. Bits struktúrareferencia

Tetszőleges hosszú bitsorozat eltárolására alkalmas struktúra.

Adatmezők

- long long unsigned int b

  A tárolt szám A bitek jobbról balra értelmezendőek.
- size_t length

  A tárolt bitsorozat hossza.

4.1.1. Részletes leírás

Tetszőleges hosszú bitsorozat eltárolására alkalmas struktúra.

4.1.2. Adatmezők dokumentációja

4.1.2.1. b

```
long long unsigned int Bits::b
```

A tárolt szám A bitek jobbról balra értelmezendőek.

4.1.2.2. length

```
size_t Bits::length
```

A tárolt bitsorozat hossza.

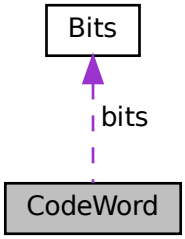
Ez a dokumentáció a struktúráról a következő fájl alapján készült:

- src/codeword.c

4.2. CodeWord struktúrareferencia

Karakter, és az azt kódoló bitsorozat.

A CodeWord osztály együttműködési diagramja:



Adatmezők

- uchar codeWord

  Egy byte, melyet a Shanon-Fano kódolás szerint kódolunk.
- Bits bits

  Egy bitsorozat, melyet a Shanon-Fano kódolás szerint a codeWord } kódolt változata.

4.2.1. Részletes leírás

Karakter, és az azt kódoló bitsorozat.

4.2.2. Adatmezők dokumentációja

4.2.2.1. bits

```
Bits CodeWord::bits
```

Egy bitsorozat, melyet a Shanon-Fano kódolás szerint a codeWord } kódolt változata.

4.2.2.2. codeWord

```
uchar CodeWord::codeWord
```

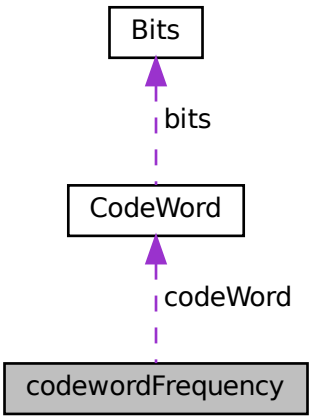
Egy byte, melyet a Shanon-Fano kódolás szerint kódolunk.

Ez a dokumentáció a struktúráról a következő fájl alapján készült:

- src/codeword.c

4.3. codewordFrequency struktúrareferencia

A codewordFrequency osztály együttműködési diagramja:



Adatmezők

- float freq
- CodeWord codeWord

4.3.1. Adatmezők dokumentációja

4.3.1.1. codeWord

```
CodeWord codewordFrequency::codeWord
```

4.3.1.2. freq

```
float codewordFrequency::freq
```

Ez a dokumentáció a struktúráról a következő fájl alapján készült:

- src/encoder.c

4.4. commandLineArguments struktúrareferencia

A program parancssori argumentumait rendező struktúra.

Adatmezők

- FILE * infile

  A –bemenet kapcsoló által megadott stream.
- FILE * outfile

  A –kimenet kapcsoló által megadott stream.
- bool displayTable

  Megajda, hogy a program kiírja-e a kódtáblát.
- bool displayStatistics

  Megajda, hogy a program kiírjon-e további számitásokat a program hatékonyságára vonatkozólag.

4.4.1. Részletes leírás

A program parancssori argumentumait rendező struktúra.

4.4.2. Adatmezők dokumentációja

4.4.2.1. displayStatistics

```
bool commandLineArguments::displayStatistics
```

Megajda, hogy a program kiírjon-e további számitásokat a program hatékonyságára vonatkozólag.

4.4.2.2. displayTable

```
bool commandLineArguments::displayTable
```

Megajda, hogy a program kiírja-e a kódtáblát.

4.4.2.3. infile

```
FILE* commandLineArguments::infile
```

A –bemenet kapcsoló által megadott stream.

4.4.2.4. outfile

```
FILE* commandLineArguments::outfile
```

A –kimenet kapcsoló által megadott stream.

Ez a dokumentáció a struktúráról a következő fájl alapján készült:

- src/cla.c

4.5. InputFileBuffer struktúrareferencia

Struktúra, mely lehetővé teszi a bitenkénti olvasást egy fájlból.

Adatmezők

- FILE * file

  A fájl, melyből olvasunk.
- uchar * currentBit

  Megadja, hogy az adott fájl olvasásánál hanyadik bitnél tartunk. Értéke 0 és 7 közötti.

4.5.1. Részletes leírás

Struktúra, mely lehetővé teszi a bitenkénti olvasást egy fájlból.

4.5.2. Adatmezők dokumentációja

4.5.2.1. currentBit

```
uchar* InputFileBuffer::currentBit
```

Megadja, hogy az adott fájl olvasásánál hanyadik bitnél tartunk. Értéke 0 és 7 közötti.

4.5.2.2. file

```
FILE* InputFileBuffer::file
```

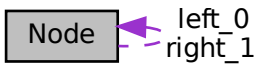
A fájl, melyből olvasunk.

Ez a dokumentáció a struktúráról a következő fájl alapján készült:

- src/fileBuffer.c

4.6. Node struktúrareferencia

A Node osztály együttműködési diagramja:



Adatmezők

- char codeword
- struct Node * left_0
- struct Node * right_1

4.6.1. Adatmezők dokumentációja

4.6.1.1. codeword

```
char Node::codeword
```

4.6.1.2. left_0

```
struct Node* Node::left_0
```

4.6.1.3. right_1

```
struct Node* Node::right_1
```

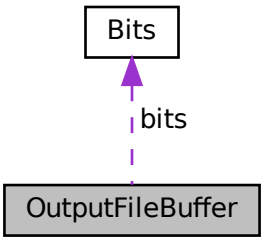
Ez a dokumentáció a struktúráról a következő fájl alapján készült:

- src/graph.c

4.7. OutputFileBuffer struktúrareferencia

Struktúra, mely lehetővé teszi a bitenkénti írást egy fájlba.

Az OutputFileBuffer osztály együttműködési diagramja:



Adatmezők

- FILE * file

  A fájl, melybe írunk.
- Bits * bits

  A még nem a fájlba beírt bitek.

4.7.1. Részletes leírás

Struktúra, mely lehetővé teszi a bitenkénti írást egy fájlba.

4.7.2. Adatmezők dokumentációja

4.7.2.1. bits

```
Bits* OutputFileBuffer::bits
```

A még nem a fájlba beírt bitek.

4.7.2.2. file

```
FILE* OutputFileBuffer::file
```

A fájl, melybe írunk.

Ez a dokumentáció a struktúráról a következő fájl alapján készült:

- src/fileBuffer.c

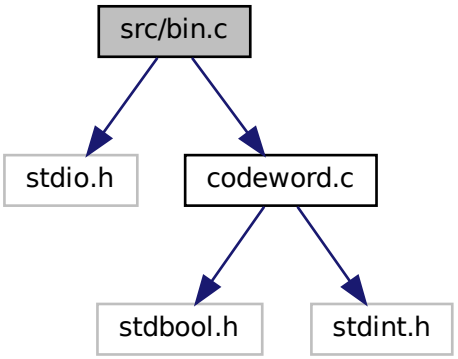
5. fejezet

Fájlok dokumentációja

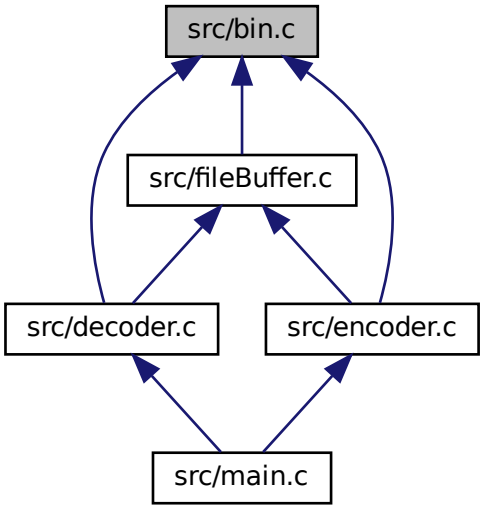
5.1. src/bin.c fájlreferencia

```
#include <stdio.h>
#include "codeword.c"
```

A bin.c definíciós fájl függési gráfja:



Ez az ábra azt mutatja, hogy mely fájlok ágyazzák be közvetve vagy közvetlenül ezt a fájlt:



Függvények

- Bits getBitFromRight (Bits bits, int n)

  Adott bitsorozatnak megadja a jobról számított `n` *-edik bitjét.*
- void bits_pushBit (Bits *bits, Bits append)

  Egy bitsorozatot bővít egy másik bitsorozattal jobb oldalról.
- void print_bits (Bits bits)

5.1.1. Függvények dokumentációja

5.1.1.1. bits_pushBit()

```
void bits_pushBit (
            Bits * bits,
            Bits append )
```

Egy bitsorozatot bővít egy másik bitsorozattal jobb oldalról.

Paraméterek

bits	a bővítendő bitsorozat
append	A hozzáfűzendő bitsorozat

5.1.1.2. getBitFromRight()

```
Bits getBitFromRight (
            Bits bits,
            int n )
```

Adott bitsorozatnak megadja a jobról számított `n` -edik bitjét.

Paraméterek

bits	A bitsorozat, melyből kiválasztjuk a bitet
n	Jobbról számítva hányadik bit

Visszatérési érték

A keresett bit

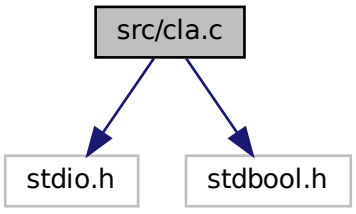
5.1.1.3. print_bits()

```
void print_bits (
            Bits bits )
```

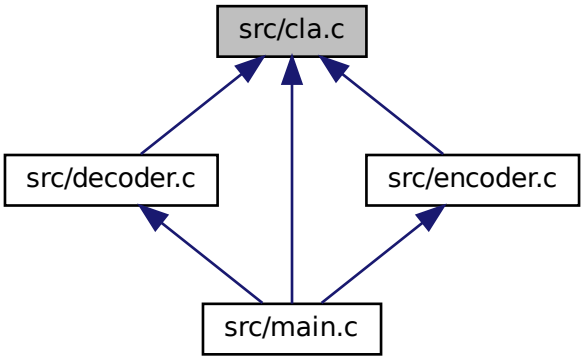
5.2. src/cla.c fájlreferencia

```
#include <stdio.h>
#include <stdbool.h>
```

A cla.c definíciós fájl függési gráfja:



Ez az ábra azt mutatja, hogy mely fájlok ágyazzák be közvetve vagy közvetlenül ezt a fájlt:



Adatszerkezetek

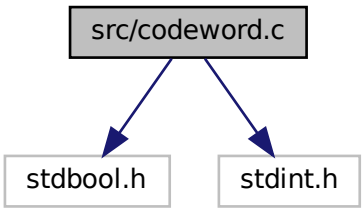
- struct commandLineArguments

  A program parancssori argumentumait rendező struktúra.

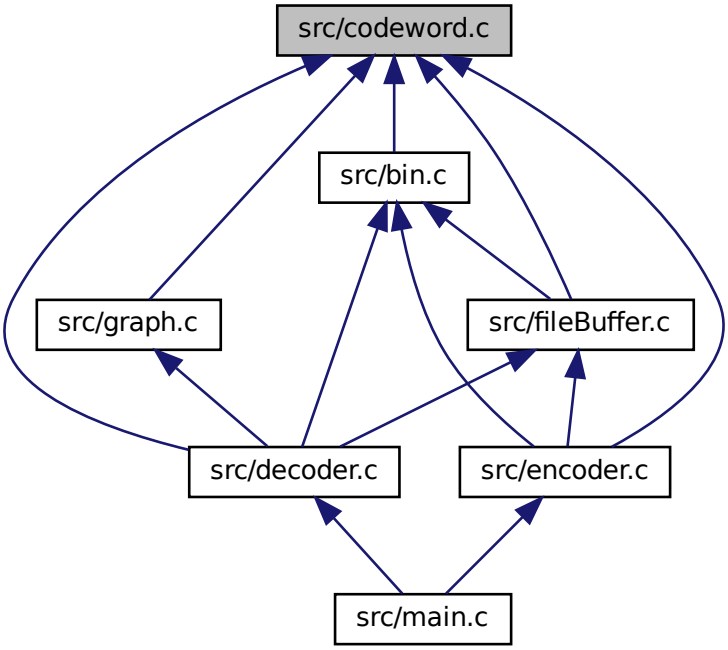
5.3. src/codeword.c fájlreferencia

```
#include <stdbool.h>
#include <stdint.h>
```

A codeword.c definíciós fájl függési gráfja:



Ez az ábra azt mutatja, hogy mely fájlok ágyazzák be közvetve vagy közvetlenül ezt a fájlt:



Adatszerkezetek

- struct Bits

  Tetszőleges hosszú bitsorozat eltárolására alkalmas struktúra.
- struct CodeWord

  Karakter, és az azt kódoló bitsorozat.

Típusdefiníciók

- typedef unsigned char uchar

  Előjel nélküli 8 bites karakter.

Függvények

- bool bits_equ (Bits b1, Bits b2)

  Összehasonlít két bitsorozatot.
- bool isNullbit (Bits b)

  Megmondja, hogy egy adott bitsorozat értelmes-e.

Változók

- const Bits NULLBIT

  Hibás kimenetet jelentő bitsorozat, melynek hossza 0.

5.3.1. Típusdefiníciók dokumentációja

5.3.1.1. uchar

```
typedef unsigned char uchar
```

Előjel nélküli 8 bites karakter.

5.3.2. Függvények dokumentációja

5.3.2.1. bits_equ()

```
bool bits_equ (
            Bits b1,
            Bits b2 )
```

Összehasonlít két bitsorozatot.

Paraméterek

b1	Az összehasonlítandó bitsorozat
b2	Az összehasonlítandó bitsorozat

Visszatérési érték

igaz, hogyha a bitsorozatok hossza és bitjei megegyeznek, különben hamis

5.3.2.2. isNullbit()

```
bool isNullbit (
            Bits b )
```

Megmondja, hogy egy adott bitsorozat értelmes-e.

Paraméterek

b	A vizsgálandó bitsorozat

Visszatérési érték

igaz, hogyha a bitsorozat hossza 0, különben hamis

5.3.3. Változók dokumentációja

5.3.3.1. NULLBIT

```
const Bits NULLBIT
```

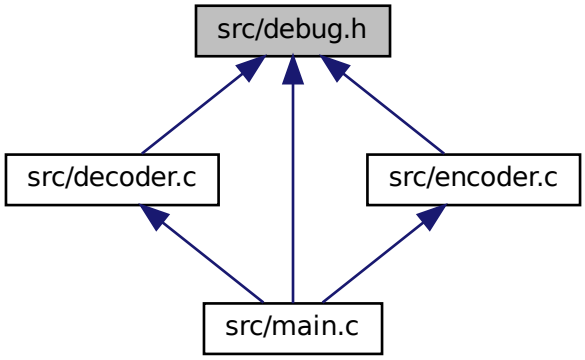
Kezdő érték:

```
= {
    .b = 0,
    .length = 0
}
```

Hibás kimenetet jelentő bitsorozat, melynek hossza 0.

5.4. src/debug.h fájlreferencia

Ez az ábra azt mutatja, hogy mely fájlok ágyazzák be közvetve vagy közvetlenül ezt a fájlt:



Makródefiníciók

- #define PRINTDEBUG_MALLOCNULL() ;;

  Kiírja hogy egy memóriafoglalás sikertelen volt.
- #define PRINTDEBUG_FILEERR() ;;

  Kiírja, hogy a fájlművelet sikertelen volt.
- #define PRINTDEBUG_CORRUPTEDFILE() ;;

  Kiírja, hogy dekódolás közben nem várt karakterrel találkoztnk.
- #define PRINTDEBUG_CUSTOM(str, ...) ;;

  Általános hibakeresésre használható, konzolra való kiírásra alkalmas.

5.4.1. Makródefiníciók dokumentációja

5.4.1.1. PRINTDEBUG_CORRUPTEDFILE

```
#define PRINTDEBUG_CORRUPTEDFILE( ) ;;
```

Kiírja, hogy dekódolás közben nem várt karakterrel találkoztnk.

5.4.1.2. PRINTDEBUG_CUSTOM

```
#define PRINTDEBUG_CUSTOM(
              str,
              ...  ) ;;
```

Általános hibakeresésre használható, konzolra való kiírásra alkalmas.

5.4.1.3. PRINTDEBUG_FILEERR

```
#define PRINTDEBUG_FILEERR( ) ;;
```

Kiírja, hogy a fájlművelet sikertelen volt.

5.4.1.4. PRINTDEBUG_MALLOCNULL

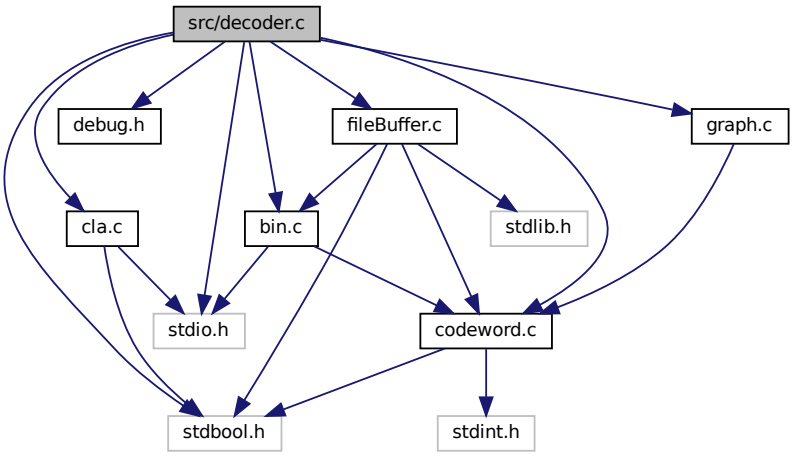
```
#define PRINTDEBUG_MALLOCNULL( ) ;;
```

Kiírja hogy egy memóriafoglalás sikertelen volt.

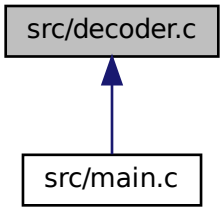
5.5. src/decoder.c fájlreferencia

```
#include <stdio.h>
#include <stdbool.h>
#include "debug.h"
#include "cla.c"
#include "bin.c"
#include "fileBuffer.c"
#include "codeword.c"
#include "graph.c"
```

A decoder.c definíciós fájl függési gráfja:



Ez az ábra azt mutatja, hogy mely fájlok ágyazzák be közvetve vagy közvetlenül ezt a fájlt:



Függvények

- void appendCodeword (Node *root, Bits codeword, char set)

  Egy fagráfhoz, `codeword` bitjeinek bejárása alapján beállítja egy elem kódolt karakterét. Ha az adott elérés nem létezik, a függvény létrehozza azt.

- Node * createNodeIfNotexists (Node *parent, int dir)

  Egy fagráf adott eleméből megpróbál `dir` által meghatároott úton továbbhaladni. Ha az nem létezik, létrehozza azt.
- void freeTree (Node *root)

  Rekurzívan felszabadít egy fagráfot.
- int decode (commandLineArguments args)

  Dekódol egy Shanon-Fano algoritmussal kódolt fájlt.

5.5.1. Függvények dokumentációja

5.5.1.1. appendCodeword()

```
void appendCodeword (
            Node * root,
            Bits codeword,
            char set )
```

Egy fagráfhoz, `codeword` bitjeinek bejárása alapján beállítja egy elem kódolt karakterét. Ha az adott elérés nem létezik, a függvény létrehozza azt.

Paraméterek

root	A fagráf gyökere
codeword	A bejárás bitjei: 0 = bal, 1 (minden más) = jobb
set	A beállítandó karakter, melyhez elérkeztünk a bejárás végén

5.5.1.2. createNodeIfNotexists()

```
Node * createNodeIfNotexists (
            Node * parent,
            int dir )
```

Egy fagráf adott eleméből megpróbál `dir` által meghatároott úton továbbhaladni. Ha az nem létezik, létrehozza azt.

Paraméterek

parent	Az elem, melyből kiindulunk
dir	Az irány: 0 = bal, 1 (minden más) = jobb

Visszatérési érték

A gráf azon eleme, mely `parent` -től `dir` irányba helyezkedik el

5.5.1.3. decode()

```
int decode (
            commandLineArguments args )
```

Dekódol egy Shanon-Fano algoritmussal kódolt fájlt.

Paraméterek

args	Parancssori bemenet, amely a dekódolás folyamatát módosítja

Visszatérési érték

0, ha a dekódolás sikeres volt. Minden más érték sikertelen

5.5.1.4. freeTree()

```
void freeTree (
            Node * root )
```

Rekurzívan felszabadít egy fagráfot.

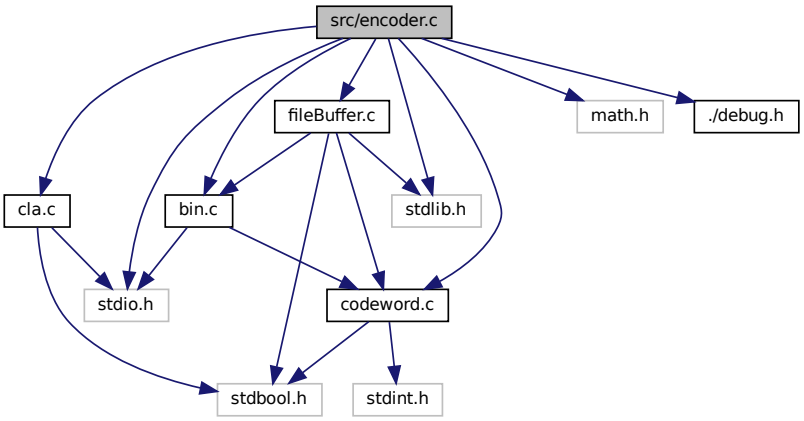
Paraméterek

root	a felszabadítandó gráf gyökere

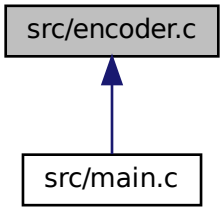
5.6. src/encoder.c fájlreferencia

```
#include <stdio.h>
#include <stdlib.h>
#include <math.h>
#include "cla.c"
#include "fileBuffer.c"
#include "codeword.c"
#include "bin.c"
#include "./debug.h"
```

Az encoder.c definíciós fájl függési gráfja:



Ez az ábra azt mutatja, hogy mely fájlok ágyazzák be közvetve vagy közvetlenül ezt a fájlt:



Adatszerkezetek

- struct codewordFrequency

Függvények

- void setCodeWord (codewordFrequency codes[ ], int i, int j)

  Rekurzívan beállítja egy kód tömbön az adott karakter Shanon-Fano algoritmus szerinti kódját.
- Bits codewordToBits (codewordFrequency code[ ], int codesLength, uchar find)

  Megkeresi `code` tömbben, `find` karaktert.
- int compare_by_freq (const void ∗a, const void ∗b)

  Frekvenciájuk alapján összehasonlít 2 frekvenciával rendelkező karakterkódolást.
- int compare_by_bitlength (const void ∗a, const void ∗b)

  Kódolásuk hossza alapján összehasonlít 2 rendelkező karakterkódolást.
- int encode (commandLineArguments args)

  Kódol Shanon-Fano algoritmus alkalmazásával egy fájlt.

5.6.1. Függvények dokumentációja

5.6.1.1. codewordToBits()

```
Bits codewordToBits (
            codewordFrequency code[],
            int codesLength,
            uchar find )
```

Megkeresi `code` tömbben, `find` karaktert.

Paraméterek

code	A kódtömb
codesLength	A kódtömb hossza
find	A keresett karakter

Visszatérési érték

A keresett kódolás vagy NULLBIT

5.6.1.2. compare_by_bitlength()

```
int compare_by_bitlength (
            const void * a,
            const void * b )
```

Kódolásuk hossza alapján összehasonlít 2 rendelkező karakterkódolást.

Paraméterek

a	Az összehasonlítandó karakterkódolás
b	Az összehasonlítandó karakterkódolás

Visszatérési érték

0 = egyeznek, >0 = `a` kódja hosszabb, <0 `b` kódja hosszabb

5.6.1.3. compare_by_freq()

```
int compare_by_freq (
            const void * a,
            const void * b )
```

Frekvenciájuk alapján összehasonlít 2 frekvenciával rendelkező karakterkódolást.

Paraméterek

a	Az összehasonlítandó karakterkódolás
b	Az összehasonlítandó karakterkódolás

Visszatérési érték

0 = egyeznek, >0 = b frekvenciája nagyobb, <0 a frekvenciája nagyobb

5.6.1.4. encode()

```
int encode (
            commandLineArguments args )
```

Kódol Shanon-Fano algoritmus alkalmazásával egy fájlt.

Paraméterek

args	Parancssori bemenet, amely a dekódolás folyamatát módosítja

Visszatérési érték

0, ha a dekódolás sikeres volt. Minden más érték sikertelen

5.6.1.5. setCodeWord()

```
void setCodeWord (
            codewordFrequency codes[],
            int i,
            int j )
```

Rekurzívan beállítja egy kód tömbön az adott karakter Shanon-Fano algoritmus szerinti kódját.

Paraméterek

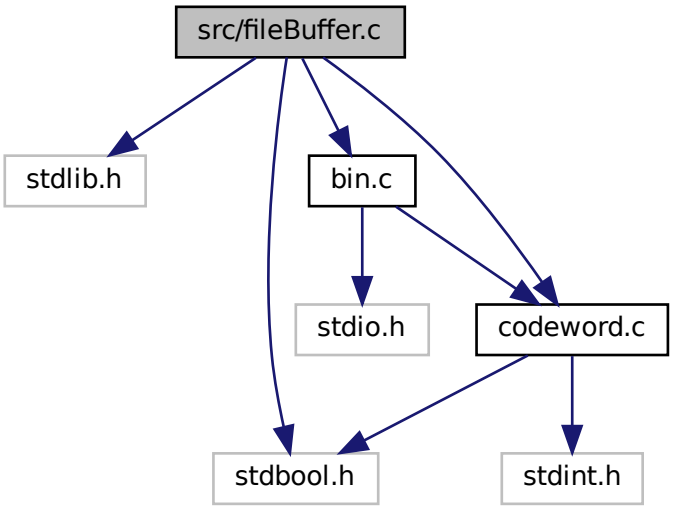
codes	A kódtömb
i	A tömb kezdeti indexe (inkluzív)
j	A tömb vegső indexe (inkluzív)

5.7. src/fileBuffer.c fájlreferencia

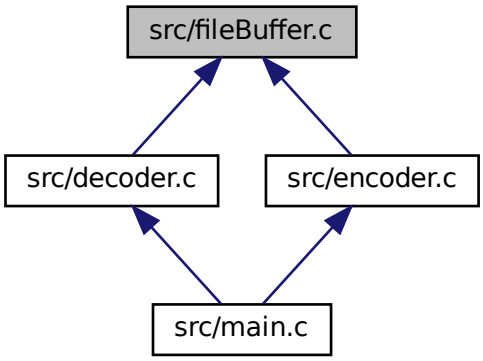
```
#include <stdlib.h>
```

```
#include <stdbool.h>
#include "bin.c"
#include "codeword.c"
```

A fileBuffer.c definíciós fájl függési gráfja:



Ez az ábra azt mutatja, hogy mely fájlok ágyazzák be közvetve vagy közvetlenül ezt a fájlt:



Adatszerkezetek

- struct InputFileBuffer

  Struktúra, mely lehetővé teszi a bitenkénti olvasást egy fájlból.
- struct OutputFileBuffer

  Struktúra, mely lehetővé teszi a bitenkénti írást egy fájlba.

Függvények

- InputFileBuffer buff_createInputFileBuffer (FILE ∗file)
- OutputFileBuffer buff_createOutputFileBuffer (FILE ∗file)
- void buff_destroyInputFileBuffer (InputFileBuffer buffer)
- void buff_destroyOutputFileBuffer (OutputFileBuffer buffer)
- void buff_rewind (InputFileBuffer buffer)
- bool buff_writeBits (OutputFileBuffer buff, Bits bit)
- bool buff_writeBit (OutputFileBuffer buff, Bits bit)
- bool buff_writeChar (OutputFileBuffer buff, uchar val)
- bool buff_writeInt (OutputFileBuffer buff, int val)
- bool buff_flush (OutputFileBuffer buff)
- Bits buff_readBit (InputFileBuffer buff)
- Bits buff_readBits (InputFileBuffer buff, int bitCount)
- Bits buff_readChar (InputFileBuffer buff)
- Bits buff_readInt (InputFileBuffer buff)

5.7.1. Függvények dokumentációja

5.7.1.1. buff_createInputFileBuffer()

```
InputFileBuffer buff_createInputFileBuffer (
            FILE * file )
```

5.7.1.2. buff_createOutputFileBuffer()

```
OutputFileBuffer buff_createOutputFileBuffer (
            FILE * file )
```

5.7.1.3. buff_destroyInputFileBuffer()

```
void buff_destroyInputFileBuffer (
            InputFileBuffer buffer )
```

5.7.1.4. buff_destroyOutputFileBuffer()

```
void buff_destroyOutputFileBuffer (
            OutputFileBuffer buffer )
```

5.7.1.5. buff_flush()

```
bool buff_flush (
            OutputFileBuffer buff )
```

5.7.1.6. buff_readBit()

```
Bits buff_readBit (
            InputFileBuffer buff )
```

5.7.1.7. buff_readBits()

```
Bits buff_readBits (
            InputFileBuffer buff,
            int bitCount )
```

5.7.1.8. buff_readChar()

```
Bits buff_readChar (
            InputFileBuffer buff )
```

5.7.1.9. buff_readInt()

```
Bits buff_readInt (
            InputFileBuffer buff )
```

5.7.1.10. buff_rewind()

```
void buff_rewind (
            InputFileBuffer buffer )
```

5.7.1.11. buff_writeBit()

```
bool buff_writeBit (
            OutputFileBuffer buff,
            Bits bit )
```

5.7.1.12. buff_writeBits()

```
bool buff_writeBits (
            OutputFileBuffer buff,
            Bits bit )
```

5.7.1.13. buff_writeChar()

```
bool buff_writeChar (
            OutputFileBuffer buff,
            uchar val )
```

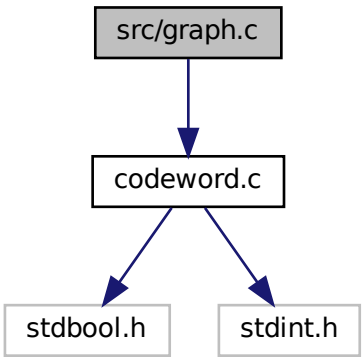
5.7.1.14. buff_writeInt()

```
bool buff_writeInt (
            OutputFileBuffer buff,
            int val )
```

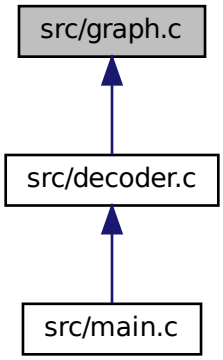
5.8. src/graph.c fájlreferencia

```
#include "codeword.c"
```

A graph.c definíciós fájl függési gráfja:



Ez az ábra azt mutatja, hogy mely fájlok ágyazzák be közvetve vagy közvetlenül ezt a fájlt:



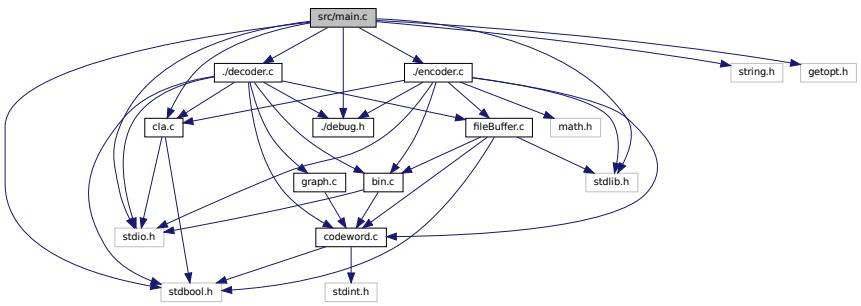
Adatszerkezetek

- struct Node

5.9. src/main.c fájlreferencia

```
#include <stdio.h>
#include <stdlib.h>
#include <string.h>
#include <getopt.h>
#include <stdbool.h>
#include "./debug.h"
#include "cla.c"
#include "./encoder.c"
#include "./decoder.c"
```

A main.c definíciós fájl függési gráfja:



Enumerációk

- enum MODE { ENCODE = 0 , DECODE = 1 , UNSET = -1 }

  Megadja, hogy a program a 'kodol' vagy 'dekodol' paraméterrel lett meghívva.

Függvények

- void printHelp ()
- int parseCLA (int argc, char **argv, commandLineArguments *args, enum MODE *mode)

  A bemeneti paraméterek feldolgozására szolgáló függvény.
- int main (int argc, char **argv)

5.9.1. Enumerációk dokumentációja

5.9.1.1. MODE

```
enum MODE
```

Megadja, hogy a program a 'kodol' vagy 'dekodol' paraméterrel lett meghívva.

Enumeráció-értékek

ENCODE	
DECODE	
UNSET	

5.9.2. Függvények dokumentációja

5.9.2.1. main()

```
int main (
            int argc,
            char ** argv )
```

5.9.2.2. parseCLA()

```
int parseCLA (
            int argc,
            char ** argv,
            commandLineArguments * args,
            enum MODE * mode )
```

A bemeneti paraméterek feldolgozására szolgáló függvény.

Paraméterek

argc	'argv' hossza
argv	parancssori argumentumok
args	

Visszatérési érték

0, ha sikeres volt az argumentumok elemzése, különben ettől eltérő

5.9.2.3. printHelp()

```
void printHelp ( )
```

Meta

5.10. Források, felhasznált irodalom

- Wayback Machine: C. E. Shannon, „A Mathematical Theory of Communication”, 1948 (https://web.archive.org/web/19980715014
  labs.com/cm/ms/what/shannonday/shannon1948.pdf)
- Halley's Comet software: Robert M. Fano, „The Transmittion of Information”, 1949 (https://hcs64.com/files/fano-tr65-ocr-only.pdf)
- Linux man pages online: (https://man7.org/linux/man-pages/index.html)
- BME InfoC: (https://infoc.eet.bme.hu)

5.11. Felhasznált segédprogramok

- Fejlesztői környezet: Visual Studio Code (https://code.visualstudio.com/)
- C compiler: GNU Compiler Collection (GCC) (https://gcc.gnu.org/)
- Projekt fordítása: Make (https://www.gnu.org/software/make/)
- Dokumentáció: Doxygen (https://www.doxygen.nl/)