【问题描述】

编写一程序实现一个利用 Huffman 编码对一个文件进行压缩和解压的工具 hzip.exe。

Huffman 压缩文件原理如下：

1. 对正文文件中字符按出现次数（即频率）进行统计
2. 依据字符频率生成相应的 Huffman 树（未出现的字符不生成）
3. 依据 Huffman 树生成相应字符的 Huffman 编码
4. 依据字符 Huffman 编码压缩文件（即按照 Huffman 编码依次输出源文件字符）。

说明：

1. 只对文件中出现的字符生成 Huffman 码。
2. 采用 ASCII 码值为 0 的字符作为压缩文件的结束符（即可将其出现次数设为 1 来参与编码）.
3. 在生成 Huffman 树时，初始在对字符频率权重进行（由小至大）排序时，频率相同的字符 ASCII 编码值小的在前；新生成的权重节点插入到有序权重序列中时，出现相同权重时，插入到其后（采用稳定排序）。
4. 遍历 Huffman 树生成字符 Huffman 码时，左边为 0 右边为 1。
5. 源文件是文本文件，字符采用 ASCII 编码，每个字符占 8 位；而采用 Huffman 编码后，最后输出时需要使用 C 语言中的位运算将字符 Huffman 码依次输出到每个字节中。

【输入形式】

基于 Huffman 的文件压缩解压工具 hzip.exe 命令行使用形式如下：

hzip [-u] <filename.xxx>

（注：xxx 是文件扩展名）

当-u 参数缺省时，对当前目录下文本文件<filename.txt>采用 Huffman 编码方式将文件压缩到文件<filename.hzip>中，被压缩的文件必须以.txt 为扩展名的文本文件，生成的压缩文件名同文本文件，但扩展名为.hzip。例如在命令行执行如下命令：

>hzip myfile.txt

采用 Huffman 编码方式压缩文件 myfile.txt，并将压缩后结果存到文件 myfile.hzip 中。

当命令行有-u 参数时，对当前目录下压缩文件<filename.hzip>进行解压，被解压的文件必须以.hzip 为扩展名的文件，解压后的文本文件名同压缩文件，但扩展名为.txt。例如在命令行执行如下命令:

>hzip -u myfile.hzip

myfile.hzip 必须是用 hzip.exe 压缩工具生成的压缩文件，解压后文件名为 myfile.txt。

命令 hzip.exe 应具有一定的对命令行参数错误处理能力，如参数个数不对、参数格式不正确等。下面为命令错误使用及提示信息（错误提示信息显示在屏幕上）：

>hzip

Usage: hzip.exe [-u] <filename>

>hzip srcfile.txt objfile.hzip
Usage: hzip.exe [-u] <filename>

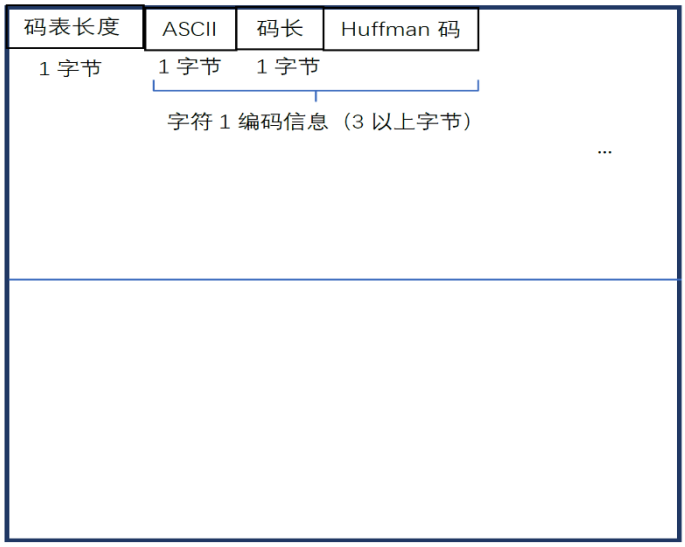
>hzip -h srcfile.txt
Usage: hzip.exe [-u] <filename>

>hzip -u myfile.c
File extension error!

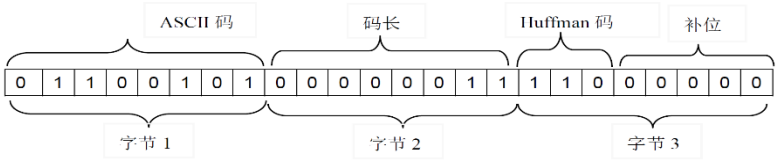
>hzip myfile.c
File extension error!

【输出形式】

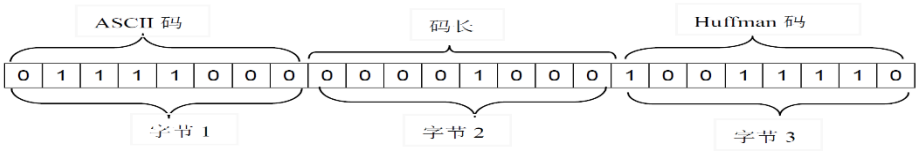
hzip.exe 工具压缩后<filename.hzip>文件格式如下：压缩文件由 2 部分组成，第一部分为 ASCII 码与 Huffman 码对照码表，第 2 部分为以 Huffman 编码形式的压缩后的文件。如下图所示。ASCII 码与 huffman 码对照码表格式如下：



从文件头开始第 1 个字节为码表长度，即压缩文件时 ASCII 码与 Huffman 码对照表中实际字符个数（即字符频率为非 0 的字符个数，包括文件结束符）。
从文件头开始第 2 个字节为相应码表内容，每个码表项依次为每个字符的 ASCII 码与相应的 Huffman 码信息，码表项按 ASCII 码由小至大排列。每个码表项由 3 部分组成，第 1 部分为字符 ASCII 码（占一个字节）、第 2 部分为其对应 Huffman 码（二进制）长度（占一个字节）、第 3 部分为其对应 Huffman 码（占不定长位），每个码表项要占完整字节（为 3 个以上字节），多余位要补 0。
假设，字符 e 的 Huffman 编码为 110，由于 e 的 ASCII 码为十进制 101（二进制 01100101）、Huffman 编码长度为 3 位（对应一字节 8 位二进制值为 00000011），则其码表项占 3 个字节（多余 5 位补 0），内容为：



若字符 x 的 Huffman 编码为 10011110，由于 x 的 ASCII 码为十进制 120（二进制 01111000）、Huffman 编码长度为 8 位（对应一字节 8 位二进制值为 00001000），则其相应码表项正好占 3 个字节，内容为：



对 hzip.exe 工具压缩后的.hzip 文件执行相应的解压命令将得到压缩前文件。
若命令行参数错误，则错误提示信息显示在屏幕上。

【样例】

若当前目录下 myfile.txt 中内容如下：

I will give you some advice about life. Eat more roughage; Do more than others expect you to do and do it pains; Remember what life tells you; Do not take to heart every thing you hear. Do not spend all that you have. Do not sleep as long as you want.

在命令行执行如下命令：

>hzip myfile.txt

将会在当前目录下生成一个 myfile.hzip 压缩文件，该文件若用二进制文件查看器查看，其内容如下：

myfile.hzip

Offset(h)	00	01	02	03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F	对应文本
00000000	1F	00	08	B6	20	02	00	2E	06	90	3B	06	60	44	06	94	...¶ .....;.`D.”
00000010	45	08	B7	49	08	D0	52	08	D1	61	04	A0	62	07	B0	63	E.·I.ÐR.Ña. b.°c
00000020	07	B2	64	06	D4	65	04	F0	66	07	B4	67	06	E8	68	05	.²d.Ôe.ðf.´g.èh.
00000030	B8	69	05	70	6B	08	D2	6C	05	D8	6D	06	EC	6E	05	E0	¸i.pk.Òl.Øm.ìn.à
00000040	6F	03	40	70	06	98	72	05	78	73	05	80	74	04	C0	75	o.@p.˜r.xs.€t.Àu
00000050	05	88	76	06	9C	77	06	64	78	08	D3	79	05	68	D0	19	.ˆv.œw.dx.Óy.hÐ.
00000060	76	F6	74	E9	FC	6A	89	05	DF	95	AC	EE	B3	E5	58	51	vötéüj‰.ß•¬î³åXQ
00000070	C3	6E	B5	F2	16	F5	87	69	FE	3D	47	AB	D7	5E	C1	2A	Ãnµò.õ‡iþ=G«×^Á*
00000080	3B	4F	F3	2F	5C	16	5F	DF	03	F4	E6	FB	38	35	44	C4	;Oó/\._ß.ôæû85DÄ
00000090	6A	8A	E6	A6	A8	76	13	53	B9	06	0D	1F	EF	FB	B1	EF	jŠæ¦¨v.S¹...ïû±ï
000000A0	19	BD	61	B7	5A	F3	3F	7B	80	D5	16	09	51	C5	86	56	.½a·Zó?{€Õ..QÅ†V
000000B0	97	98	8B	FD	3F	0F	9F	DE	D3	2E	EE	74	35	44	BF	D3	—˜‹ý?.ŸÞÓ.it5D¿Ó
000000C0	E4	25	47	16	10	9B	F9	A9	5B	D9	97	AC	1A	A2	5E	A7	ä%G..›ù©[Ù—¬.¢^§
000000D0	F9	09	51	C5	84	37	FF	31	50	36	B9	D1	50	1A	A2	33	ù.QÅ„7ÿ1P6¹ÑP.¢3
000000E0	5C	C9	2D	80													\É-€

在命令行对所生成的压缩文件 myfile.hzip 执行如下命令：

>hzip -u myfile.hzip

将在当前目录下得到压缩前文件 myfile.txt。

【评分标准】

该题要求实现文件的压缩和解压，**本题只测试压缩功能**，提交程序名为 hzip.c。