

1. Function and Use.

This small program will convert Big 5+ or GBK encoded Chinese characters into a ‘preprocessed’ form. The need of this program arises from the fact that these two encodings use the characters ‘\’, ‘{’, and ‘}’ which have special meanings in T_EX.

Use this program as a filter:

```
extconv < input_file > output_file
```

2. The program.

The only function of this program is to replace all occurrences of Big 5+ and GBK encoded characters XY (X and Y are the first and the second byte of the character) with $\text{\textasciix}X\text{\textasciix}ZZZ\text{\textasciix}$, where ZZZ represents the second byte as a decimal number. $\text{\textasciix}0x7F$ is used as an active character and delimiter.

Additionally we define a \TeX macro at the very beginning to signal a preprocessed file.

The following code is very simple. No error detection is done because \TeX which will see the output of `extconv` complains loudly if something is wrong.

```
#define banner "extconv(CJKver.4.8.5)"
#include <stdio.h>
#include <stdlib.h>

int main(int argc, char *argv[])
{
  int ch;

  fprintf(stdout, "\\def\\CJKpreproc{%s}", banner);
  ch = fgetc(stdin);

  while (!feof(stdin))
  {
    if (ch >= #81 & ch <= #FE)
      {fprintf(stdout, "\\177%c\\177", ch);
        ch = fgetc(stdin);
        if (!feof(stdin))
          fprintf(stdout, "%d\\177", ch);
        }
    else
      fputc(ch, stdout);
    ch = fgetc(stdin);
  }
  exit(EXIT_SUCCESS);
  return 0;
}
/* never reached */
```