

I117 (13) CSV (その3) 補足

知念

北陸先端科学技術大学院大学 情報科学研究科
School of Information Science,
Japan Advanced Institute of Science and Technology

2辞書出納プログラム — 作業確認

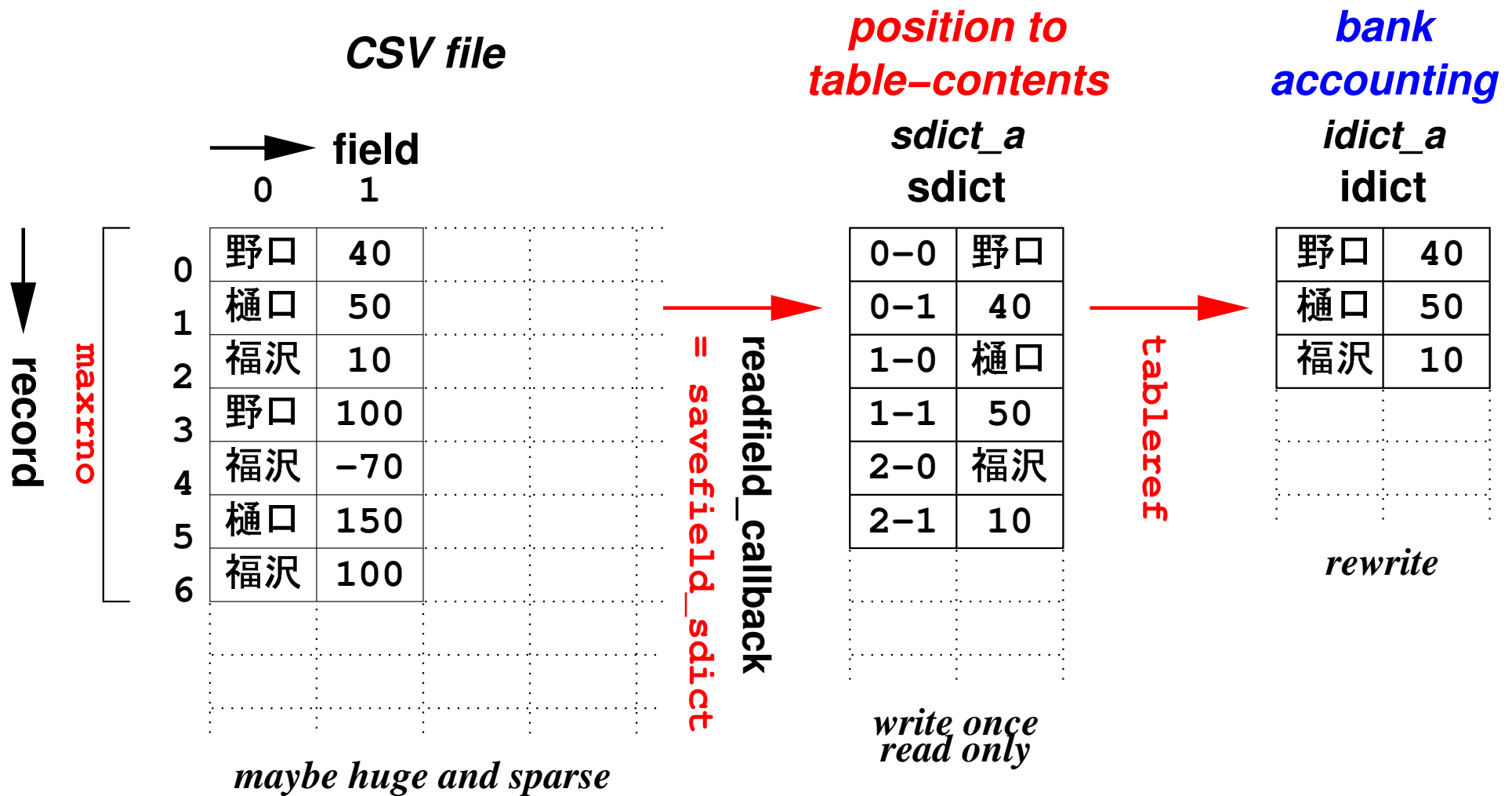
- 1) WWW ページから 2dict.c を取得 — 2dict.c
- 2) CSV 操作関数群を作成 — libcsv.c
- 3) idict 操作関数群を作成 — libidict.c
- 4) sdict 操作関数群を作成 — libsdict.c
- 5) コンパイル — a.out
- 6) データファイル作成 — bamount.csv

2dict.c 取得

I117 2008年度のWWWページから、プログラム本体
2dict.c を入手
中身を読むと、3つファイルが必要なのが見えるだろ
う

```
#include "libcsv.c"  
#include "libdict.c"  
#include "libsdict.c"
```

用意しなければならない関数も分かるはず



libcsv.c 作成

- csvparse.c を libcsv.c という名前に変える
csvparse.c も WWWページにおいてある
- EUC 対応にする
 - ◇ ITEXT の変更
 - ◇ readch() で $ch \leq 0x7f$ を $ch \leq 0xff$ に
- コールバック (callback) を呼び出す形式にする
 - ◇ readfield で readfield_callback
 - ◇ (余裕があれば) readrecord で readrecord_callback
- main() を取り除く

libidict.c 作成

以下の型を用意する

```
typedef struct {
    char *key;
    int   value;
} idict_c;
typedef struct {
    idict_c *slot;
    int      len;
    int      use;
} idict_a;
```

libidict.c 作成 (*cont.*)

以下の関数を用意する

```
int idict_keycmp(const void *ap, const void *bp);
int idict_valuecmp(const void *ap, const void *bp);
int idict_init(idict_a *dict);
idict_a *idict_new();
int idict_expand(idict_a *dict);
int idict_print(idict_a *dict);
idict_c *idict_findpos(idict_a *dict, char *xkey);
int idict_add(idict_a*dict,char*xkey,int xvalue);
```

libidict.c 作成 (cont.)

これまで断片的に説明したことを整理
たとえば、l117-6.pdf p.43 はメンバ名が違うが内容はほぼ同じ

- idict_keycmp, valuecmp() は key や value でソートするための比較関数
- idict_new() は idict_a を malloc で確保
- idict_init() は idict_a のメンバ slot を malloc で確保、len や use を初期化
- idict_expand() は領域不足の際の領域拡張
l117-10 p.10 が初出

libidict.c 作成 (*cont.*)

- `idict_print()` は辞書内容表示
 - `idict_findpos()` は辞書内検索
 - `idict_add()` は辞書へ項目追加
- l117-13.pdf p.6 が `findpos` や `add` に近い

libsdict.c 作成

以下の型を用意する

```
typedef struct {
    char *key;
    char *value;
} sdict_c;
typedef struct {
    sdict_c *slot;
    int      len;
    int      use;
} sdict_a;
```

libsdict.c 作成 (cont.)

以下の関数を用意する

add の xvalue が char* になっている以外、idict と同じ

```
int sdictionary_keycmp(const void *ap, const void *bp);
int sdictionary_valuecmp(const void *ap, const void *bp);
int sdictionary_init(sdictionary_a *dict);
sdictionary_a *sdictionary_new();
int sdictionary_expand(sdictionary_a *dict);
int sdictionary_print(sdictionary_a *dict);
sdictionary_c *sdictionary_findpos(sdictionary_a *dict, char *xkey);
int sdictionary_add(sdictionary_a*dict,char*xkey,char*xvalue);
```

コンパイル

```
% cc 2dict.c
```

文法ミス、スペルミスに注意

データファイル作成

以下のような内容のファイルを作る、名前は amount.csv とする

最終行に LF があっても構わない

```
野口英世, 40 LF
樋口一葉, 50 LF
福沢諭吉, 10 LF
野口英世, 100 LF
福沢諭吉, -70 LF
樋口一葉, 150 LF
福沢諭吉, 100
```

実行

```
% nkf -e bamount.csv | ./a.out
#sdict expand 10 -> 20
no LF (record separator)
0 樋口一葉 200
1 福沢諭吉 40
2 野口英世 140
```