

kconfigの国際化



SIOS

サイオステクノロジー株式会社

佐藤嘉則

目次

- きっかけ
- 現状の問題点
- お試し
- いろいろ考える
- 考察
- 本題



きっかけ

■ とある書き込みを見かけた所から…

[Blurred text]

etchでカーネル2.6の再構築を行おうと思ってるのですが、
以前の2.4とかだとJFにあるConfigure.helpをとってきて、ソースディレクトリの配下にあるものと置き換え
は
make menuconfig でのヘルプを日本語化できたのですが、2.6からはどうすればよいのでしょうか？

1. そういえばgettextizeするパッチが流れてたなあ…
2. ログを探してみる
3. 見つけたので眺める
4. こんなコードを見つけてしまう。

```
:  
708         str_append(&help, _(menu_get_help(menu)));  
:
```

嫌な予感が…



きっかけ

- 深追いすることに…
 - パッチの出所を探す
 - <http://tlktp.sourceforge.net>
 - 駄目だこいつ…、早く何とかしないと(AA略)

- ダメな理由はこの後で。



目次

- きっかけ
- **現状の問題点**
- お試し
- いろいろ考える
- 考察
- 本題



gettext

- て何ぞやという人は
 - つ “man 3 gettext”
- と言ってもすぐに見れないと思うので
 - メッセージ(text)を渡すと、どこかから紐付いているメッセージをを引っ張って来て(get)、返してくれる。
 - 持ってくるメッセージを翻訳しておく事で、英語→日本語みたいな事が出来る。



gettext

- 引っ張ってくるメッセージはどこに?
 - ~.moというファイル。
 - 単純なテキストファイルではない。
 - 普通は"/usr/share/locale"あたりにいろいろ入ってます。
 - 言語毎に用意されていて、指定された言語の物を使う。



gettextの中身

- やってること
 - msgidをキーにして
 - 元のメッセージ
 - msgstrを検索する
 - 紐付いているメッセージ
- 極端に単純化するとこんなもの
 - それ以外にも色々やってくれますが、今回はあまり関係ないので省略



ダメな理由

- 検索処理の宿命
 - データ量に比例して遅くなる
 - データベースとかはがんばって遅くなりなくくする
- gettextはそこまで頑張っていない
 - まあDBエンジンじゃないし
 - 一応の努力はしている
 - キャッシュのようなものを持っていたり



ダメな理由

- gettextの検索アルゴリズム
 - ハッシュ
 - なのでうまく行けば $O(1)$
 - ただハッシュ値がそれなりに衝突する
 - なのでそうならないことが結構ありそう
 - オープンアドレスなので外れると線形探索
 - なので最悪 $O(n)$
 - データが少なければ別に問題は無いけれど…



ダメな理由

- 配布されている.poファイルを調べてみる

```
$ grep msgid it.po|wc -l  
10766
```

- ちなみに手元で一番大きかったのはgimpの約2,500
- いくら何でも多すぎ
- しかもmsgidが長いのでstrcmpが重くなりそう。



ダメな理由

- gettextの目的
 - メッセージをバイナリから分離して
 - 言語毎に個別のメッセージを持てるように
 - 元ネタはリソース文字列 (たぶん)
- kconfigの使い方は…
 - 外部ファイルのメッセージを食わせている
 - 元々分離しているメッセージにgettextは必要?



ダメな理由

やり方は一つじゃない
でも最適解はそんなに多くない。

使い捨てじゃなければ最適解を追求すべき

- gettextのカatalogは小さい方がいい
- 外部ファイルのメッセージ食わせるな

いまのはダメ



目次

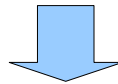
- きっかけ
- 現状の問題点
- お試し
- いろいろ考える
- 考察
- 本題



どうする？

■ gettextステ

- 全部捨てるんじゃなくて、外から読むメッセージは通さないようにする
- 外から読んでくるメッセージはあらかじめ訳しておけばいい



たぶん正解



本当に捨てられるか?

■ とりあえず試してみる

```
ファイル(E) 編集(E) 表示(V) ターミナル(I) 移動(G) ヘルプ(H)
.config - Linux Kernel v2.6.27-rc3 Configuration

Linux Kernel Configuration
Arrow keys navigate the menu. <Enter> selects submenus --->.
Highlighted letters are hotkeys. Pressing <Y> includes, <N> excludes,
<M> modularizes features. Press <Esc><Esc> to exit, <?> for Help, </>
for Search. Legend: [*] built-in [ ] excluded <M> module < >

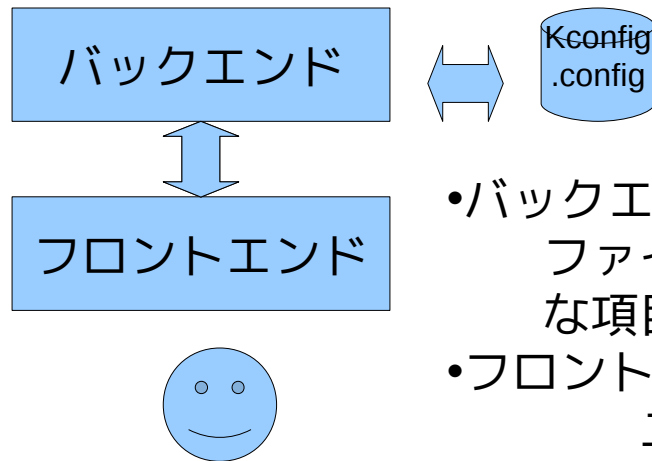
[*] 読める!読めるぞお! (NEW)
  General setup --->
  [ ] Enable loadable module support --->
  [*] Enable the block layer --->
  Processor type and features --->
  Executable file formats --->
  [ ] Networking support --->
  Generic Driver Options --->
  [*] Memory Technology Device (MTD) support --->
  [ ] Block devices --->
v(*)

<Select> < Exit > < Help >
```

注：画面ははめ込み合成ではありません

本当に捨てられるか?

- kconfigの中を一応調べてみる。
 - たまたまうまく行ったとかだと困るので。
- おおざっぱな内部構造



- バックエンド
ファイルの読み込み、依存関係を整理して設定可能な項目をフロントエンドに提供。
- フロントエンド
ユーザーインターフェースを提供

本当に捨てられるか?

- バックエンド
 - flex+bisonで生成された構文解析器なので大丈夫 (gcc3で問題ないし)
- フロントエンド
 - バックエンドからもらったメッセージをそのまま表示ライブラリに渡すので問題無し。
- 結論 捨てても大丈夫(ということにする)



目次

- きっかけ
- 現状の問題点
- お試し
- いろいろ考える
- 考察
- 本題



仕様を決めます

- 翻訳したメッセージをどうやって持つ?
 - 単純に置き換えるのはまずい
 - Kconfigファイルにはconfigに必要な情報も入っている
 - それを置き換えてしまうと色々不都合が…
 - 変な.configが出来たり
 - 新しい項目が設定できなかったり
 - 元情報と同居するのが前提
 - ついでに複数の言語が同居出来た方が良さそう



仕様を決めます

- gettextの場合
 - Kconfigはさわらない。
 - 複数言語を同梱できる。
 - /usr/share/localeにカタログを入れないと使えない。
 - なので不便。
 - まだ取り込まれていない修正で改善される模様。
- このあたりは考慮すべき。



仕様を決めます

- 各言語を並べて書く。

```
config FOO
  bool "Foo"
  bool-ja "ふー"
```

- 邪魔 or 鬱陶しいという声が聞こえるような気がする。
- 意図せずこわされて、泣きながら直す姿が見えてしまう。
- 言語毎にコマンドを追加しないといけない。



仕様を決めます

- 翻訳したKconfigを別に用意する。
 - 言語毎に別ファイルに出来るので、先ほどのような問題は発生しない。
 - 既存ファイルに手を加えないので、配布が楽
 - 単純に上書きすればいい。
 - 文法を合わせてしまえば構文解析器をそのまま使える。
- よさそう



仕様を決めます

■ 罨

- 単純に翻訳Kconfigを読むようにすると、
 - Kconfigファイルにはconfigに必要な情報も入っている
 - それを置き換えてしまうと色々不都合が…
- という問題が。

■ どうする？

- 答えは次のページ



仕様を決めます

- どうする？
 - 元のKconfigも読む。
 - こっちを基準にする。
 - 翻訳版はメッセージのみ使う。
 - default / depends onとかは書いてあっても無視、
- これなら
 - 悲しいことはおきない。
 - 新規項目は英語で表示。
- よさそう



仕様を決めます

- その他(ほぼ確定)
 - 翻訳ファイルは"Kconfig.{言語コード}"
 - 元ファイルと同じディレクトリに置く
 - 場所変えると面倒なので
- その他(未確定)
 - 国コードは?
 - 翻訳ファイルのエンコーディング



こんな風になります

■ 取り合えず日本語の場合

■ Kconfig

```
config FOO
    bool "use FOO"
    default y
    help
        FOO is ...
```

■ Kconfig.ja

```
config FOO
    bool "〇〇を使用する"
    default y
    help
        〇〇は...
```



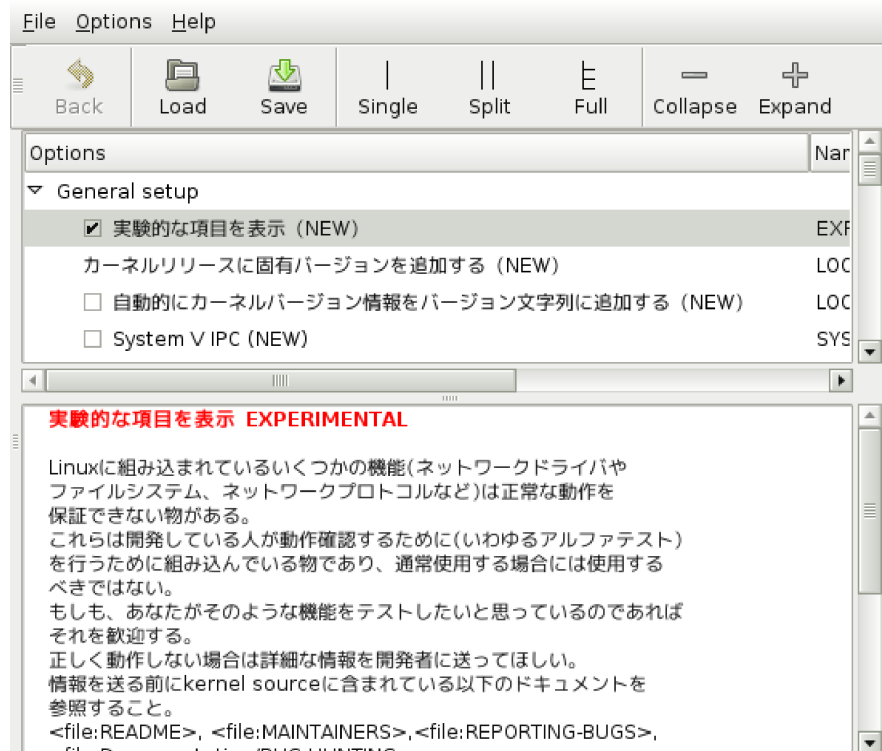
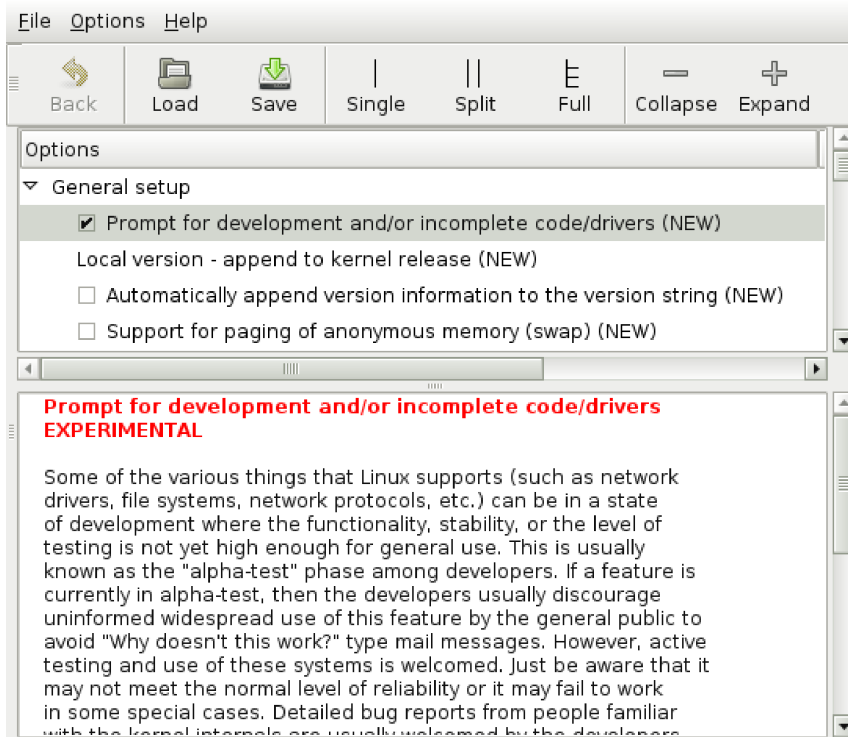
目次

- きっかけ
- 現状の問題点
- お試し
- いろいろ考える
- **考察**
- 本題



試してみる

- がんばって考えた仕様で動かしてみました



注：くどいようですがはめ込み合成ではありません

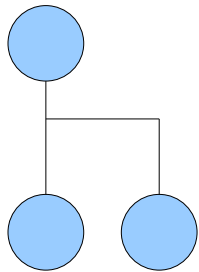
どうやったの？

- かなり力技
 1. 普通に"Kconfig"を読み込む。
 2. 翻訳版"Kconfig"を読み込む。
 3. 翻訳版で読んだメッセージを、普通版の所に上書きしてしまう。
 4. 翻訳版を隠す。
- 詳しく知りたい人は論文読んでください。

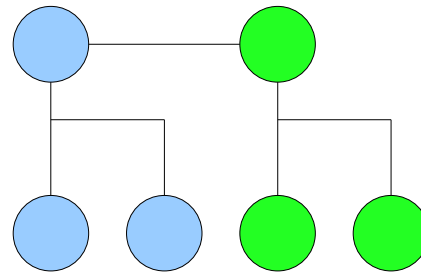


どうやったの？

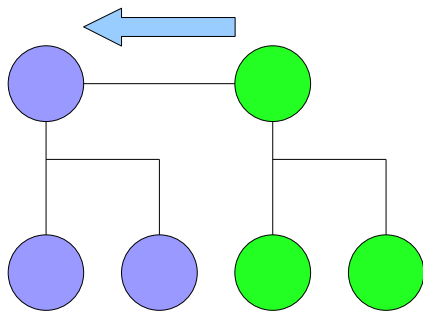
■ というのも何なので図解



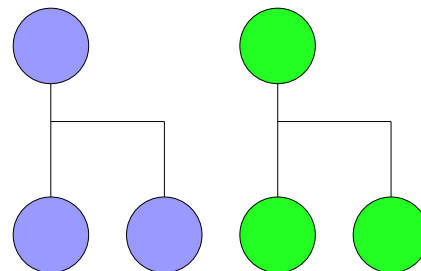
1. 読み込んだKconfigは
木構造になる



2. 翻訳を読み込むと
隣に生える



3. メッセージをコピー



4. 翻訳を切り離して
見えないようにする



どうやったの？

- こんな感じで考えた仕様を実装出来ました。
- だいたい予定通りかなあと考えています。
- ただ、今の実装はかなり手抜きなのでもっときれいに直す予定。
- <http://uclinux-h8.sourceforge.jp/kconfig-i18n/kconfig-i18n-test.tar.gz>



目次

- きっかけ
- 現状の問題点
- お試し
- いろいろ考える
- 考察
- **本題**



メッセージの翻訳

- ここから本題。
- Kconfigファイルはいくつありますか？
 - 約500個@2.6.26。
- 全部訳さないとダメ？
 - 訳が終わった所から使えます。



メッセージの翻訳

- どうやってそれだけの量を翻訳するつもりですか?
 - どうしましょう？
 - というかその辺は何も考えてません。



メッセージの翻訳(妄想)

- 個々のKconfigファイルに担当を割り当ててがんばってもらえばいいかなとか…
 - でかい.poをバラすより楽。
- 数が多いのはarchとdriversなので需要の多いところから訳せば何とかかなあとか…



ということで

- 面倒見てくれる人を募集

質問をどうぞ

即答できないものは

<http://sourceforge.jp/projects/uclinux-h8/wiki/kconfig-i18n>
に書きます。

Wikiなので適当に追加してもらっても構いません。

