

単一のドキュメントデータからの書籍、 e-Learning システム、 質問受付システムの生成

内田智史
神奈川大学工学部経営工学科

筆者は、2002年6月30日に電波新聞社より『Java プログラミング徹底入門～基礎編～』という書籍を出版した。この出版物は LaTeX によって記述されているが、そのソースファイルは、あるオリジナルの原稿ファイルを組版システムによって LaTeX に組み直したものである。そのオリジナルの原稿の中には、サンプルプログラムのソースファイルなども含まれており、このようなプログラミング書籍出版の際に問題となる書籍掲載のソースプログラムとオリジナルソースプログラムのバージョンの違いなどを無くしている。さらに、このオリジナル原稿の中に適当なタグを埋め込むことにより、書籍に載った原稿を利用した形で e-Learning システムや半自動的な質問応答システムを開発し、実際に Web 上で公開する予定である。今回は、そのシステムの概要について報告する。

System for generation of books, e-Learning system, and semi-automated answering system from original text data

Satoshi Uchida

Department of Industrial Engineering and Management, Kanagawa University

This paper presents a system for generating various contents using original text data, developed specifically for application to text books such as computer programming texts, e-learning content, and semi-automated answering system. The system has already been used successfully to published a book on Java programming; “An exhaustive introduction to Java programming – basic course –”, Dempa Publications, Inc., June 30, 2002. In the near future, the system will be used to produce contents for an e-Learning system and semi-automated answering system from the original data

E-mail: uchida@inf.ie.kanagawa-u.ac.jp

URL: <http://www.inf.ie.kanagawa-u.ac.jp/>

1. はじめに

プログラミング関係の書籍を出版する際にさまざまな点が問題になる。以前は、LaTeXによる出版も多かったが、最近はクオークのようなDTPソフトウェアが多く使われている。原稿を書く側からは、手元において何度でも校正が出来る環境の方が望ましいので、LaTeXによる出版が望ましいと思われるが、書籍としての見栄え・体裁などの点から最近では敬遠されているようである。また、書籍に掲載するソースプログラムの管理も問題となる。500ページ程度の書籍の場合、300個程度のソースファイルを管理することになるが、執筆しているうちにソースファイルの内容を書き直すことが頻繁に発生する。最新版のソースプログラムを書籍に含めたつもりでも、さまざまな人的ミスで古いプログラムが掲載されてしまうという事例は多くある。また、ソースファイルを取り込む際にカットアンドペーストの人為的ミスから最後の行を抜かしてしまったりなどというケースもよくみかける。また、最近ではほとんどの著者が、ワープロなどの手段を使って電子的に原稿を作成しているはずであるが、書籍発行後、せっかく作成したその原稿が利用されることは少ない。

以上の問題点を考え、筆者は次のようなシステムを開発した。まず、オリジナルの原稿を作成し、それをすべてのソースとする。このオリジナルソースに組版システムを通すと、LaTeXの原稿が生成され、それをLaTeXにかけると出版に耐え得る品質のDVIファイルあるいはPSファイルが生成される。また、そのオリジナルソースの中には、サンプルプログラムのソースプログラムも入っているので、そこからプログラムを抜き取り、書籍に必要なコメントなどを自動的に入れ込むプログラムを作り、ソースファイルのバージョンの問題を解決した。

また、現在、オリジナルの原稿の中に、質問回答用のタグを埋め込み、学習者の質問に応じて、オリジナルテキスト中のデータをベースにして、その質問に回答するシステムを開発中であり、近日中にWeb上に公開する予定である。さらに、この同一のテキストの中から、プログラムのトレースによる初心者向けのWeb教材などのe-Learningシステムを生成する予定である。

2. 組版システム

LaTeXは、理工系の論文作成などの世界ではポピュラーなツールである。すべてテキストデータで構成されているという点が大きな特徴である。筆者は、過去数冊のプログラミング関係の書籍を出版してきたが、1990年前半ではLaTeXを使ったものが多かった。LaTeXは、筆者自ら組版に携われるので、校正回数をあまり気にする必要がないという利点があった。しかし、近年、印刷業界の世界では、クオークによる組版が一般的となったようで、筆者の最近の書籍はクオークによるものになっている。しかし、クオークは専門のオペレータが操作するため、校正にはコストがかかり、校正回数が増えるとコストが増

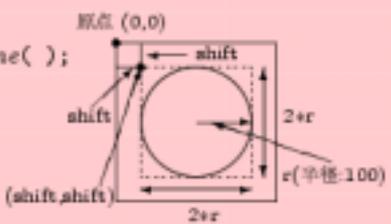
すという欠点がある。

しかし、LaTeX は、数学論文を主体とする体裁の原稿には向いているが、最近の丁寧な書籍作りにあるような詳細でグラフィカルな紙面構成には向いていない。そこで、筆者は単一のオリジナルテキストデータを用意し、そのデータを加工することで LaTeX のソースファイルを生成するようにした。図 1 は、それによって生成された書籍の一例である(以下の図は、出版前の PS ファイルを表示したもので実際のものとは多少異なる。実際の書籍は青いインクを使っている)。

9.1 フレームクラスとメインメソッドの分離

7章までは、基本的に 1つのプログラムファイルで 1つのフレームを表示するプログラムを作成してきました。たとえば、次のプログラム 9-1 は、赤い円を表示する *RedCircleFrame* というクラスを作り、メインメソッド (14 行目~21 行目) の中で、*RedCircleFrame* クラスの参照型変数 *rcf* を作り (16 行目)、それを描画しています (20 行目)。

```
★ program list 9-1 ★ 赤い円を表示するフレームとメインメソッド (RedCircleFrame.java)
1  import java.awt.*;
2
3  class RedCircleFrame extends JFrame
4  {
5      public void paint( Graphics g )
6      {
7          int r = 100;    // 円の半径
8          int shift = 50; // 円に外接する正方形の左上隅の座標 (shift, shift)
9
10         g.setColor( Color.red );
11         g.fillOval( shift, shift, 2 * r, 2 * r ); // 中心の円を描く
12     }
13
14     public static void main( String [] args )
15     {
16         RedCircleFrame rcf = new RedCircleFrame( );
17
18         rcf.setSize( 300, 300 );
19         rcf.setTitle( "赤い円" );
20         rcf.setVisible( true );
21     }
22 }
```



このプログラムは、同じフォルダに *ExitFrame.class* と *ClosingListener.class* を置いて実行させてください
このプログラムは、終了ボタンによって終了させることができます

図 1 生成された LaTeX テキストによって得られたページ(内田智史著『Java プログラミング徹底入門、基礎編』、電波新聞社)

もちろん、このような加工は LaTeX あるいは TeX のマクロを使えば十分にできるのであるが、オリジナルデータの他の利用目的も考慮し、このような形式を採用した。

3. 質問自動受付システム

書籍を読めば、それに対して質問が生まれるのは当然のことである。最近、E-mailなどで直接筆者に問い合わせることもあるようであるが、多くの読者にとって直接あるいは出版社を通して筆者に問い合わせることは躊躇することが多いであろう。読者からの質問は何パターンかに分類できる。一つは、書籍に書いてあるのであるが、それを見落としている場合の質問で、これは「書籍の何ページを読んでもください」で済む。次に、書籍の記述がよく分からない場合の質問がある。これは、書籍のページ数の制限から来ることが多く、説明が不十分なために起こる。そこで、書籍には一部の説明しか載せていなくても、オリジナルデータには十分な説明を記載しておき、読者からの質問に対して十分な分量の記述で回答できる仕組みを開発している。基本的なアイデアは、オリジナルテキスト中に質問回答用のタグを入れておき、読者からの質問に対してそのタグを検索し、それを回答するというものである。その際に、書籍の原稿を引用ページを提示しながら、さらには追加部分の説明で補足するという形態である。もちろん、この追加部分の原稿は、本の出版時には用意できるはずがないので、この作業は書籍の出版後になる。Web上に置いて公開することで、このタイムラグを解決できる。

現在は、タグの構成の原案とオリジナルテキストへの追加の原稿を作成している段階である。

4. e-Learning システム教材の抽出

オリジナルテキスト中には、ソースプログラムを始め、さまざまなコンテンツデータが格納されている。これらのデータを抽出し、e-Learning システムのコンテンツとして提供する仕組みを開発する。現在、予定していることは、オリジナルデータ中に埋め込まれたソースプログラムの自動トレース機能の提供や学習用のコンテンツのオリジナルテキストの再利用である。

5. 終わりに

以上の試みを近日中に Web 上で一般の方々を対象とした実験システムを構築し、出来ることから公開していきたいと考えている。将来的には、Web だけでなく書籍に含める CD-ROM の中にこのようなシステムを配置することにより、より多次元的な書籍の出版を考えていきたい。