

## 5 講義・実習記録

### <使用ソフト>

使用 OS Windows XP Home Edition (プレインストール)  
使用ソフト EmEditor (<http://www.emeditor.com/jp/index.htm>)  
M's TeX Helper 2 (<http://ich.sakura.ne.jp/manabu/>)  
T<sub>E</sub>X 一式 (L<sup>A</sup>T<sub>E</sub>X2<sub>ε</sub> 美文書作成入門に付属 CD-ROM から)  
ヒラギノ OpenType 基本6 書体パック  
<oXygen/> XML Editor (<http://www.oxygenxml.com/>)

使用 OS Linux (KNOPPIX 守岡版 DVD)

使用ソフト XEmacs  
AUCTeX  
CHISE  
OTF パッケージ  
Perl

使用テキスト セミナー教材用テキスト

L<sup>A</sup>T<sub>E</sub>X2<sub>ε</sub> 美文書作成入門 (奥村 晴彦、2004 年、技術評論社)  
入門 XML (Erik T.Ray、2004 年、オライリー・ジャパン)

### ※使用ソフトの概要

**EmEditor** <http://www.emeditor.com/jp/index.htm>

EmEditor はアカデミック使用に限り、無償で利用できる Windows 用多機能テキストエディタである。特徴としては、プラグインを使って様々な機能を搭載する事ができることである。更に特筆すべき点は Unicode に完全に対応していることである。

**M's TeX Helper 2** <http://ich.sakura.ne.jp/manabu/>

M's TeX Helper 2 は EmEditor の無償の拡張プラグインであり、T<sub>E</sub>X を文章を作成するにあたり、各種 T<sub>E</sub>X コマンドなどの入力やコンパイル、プレビュー、印刷、PDF 出力等を支援してくれる強力なプラグインである。

### **T<sub>E</sub>X**

T<sub>E</sub>X とは、印刷物を作る工程の一つである組版という作業をコンピュータ上で実現するために開発されたソフトウェアのことである。本セミナーでは、「L<sup>A</sup>T<sub>E</sub>X2<sub>ε</sub> 美文書作成入門」に付属の CD-ROM を使って必要なもの一式をインストールした。T<sub>E</sub>X 自体は無償なソフトであって、Unix、Mac OS X、Windows 等で利用できる。

<oXygen/> **XML Editor** <http://www.oxygenxml.com/>

<oXygen/> は有償ソフトであるが、30 日間は試用できる。このソフトは XML 関連する XSLT、XML スキーマ、Relax NG、DTD などが扱える XML に特化したエディタである。Windows、Mac OS X、Linux、Solaris で利用できる。

## KNOPPIX

KNOPPIX (クノーピクス) とは、CD のみでブート可能な Linux ディストリビューションである。ドイツの Knopper 氏が Debian パッケージを元に開発している。独立行政法人産業技術総合研究所 (産総研) が日本語対応版に改良したものを公開している。

オリジナル <http://www.knopper.net/knoppix/>

日本語版 <http://unit.aist.go.jp/itri/knoppix/>

KNOPPIX は CD のみで動作するため、既存のハードディスクインストールすることなくアプリケーションを試すことができるため、安心して気軽に Linux 環境を体験することができる。Linux は、どのパソコンでも動くとは限らない (特に最新の機種) ので Linux をインストールしてみたいと思ったときに、自分の持っている機種で動作するのかを確認するためなどにも KNOPPIX は有効である。CD から起動するため動作はあまり速くないのは仕方がないが、ハードディスクにもインストール可能で WindowsXP とのデュアルブート環境を構築することも可能である。本セミナーでは受講生全員がハードディスクにインストールを行なった。

## XEmacs

XEmacs は UNIX 系の OS で非常によく利用されているエディタである Emacs の一種である。ユーザーからみて特徴的なことは、マウスでの操作を想定せずに設計されている点である。キーボードによる入力から操作することが基本となっているため、使いこなすには数多くのキーバインドを覚えなければならず、初心者を使い易いエディタとは言い難いが、マウスによるカーソル移動やメニュー等を選択する事もできる。また、様々な形式のテキストファイルをより効率良く編集するために、それに特化したエディタに変身することも最大の特徴である。特に本講義で使用する XEmacs とは、Emacs をより視覚的に操作しやすいように改良したものであるが、Emacs とは内部構造がかなり異なっている。現在、講師の守岡知彦氏が中心となって開発が行なわれている CHISE が動くのは XEmacs だけである。

## AUCTeX

AUCTeX とは GNU Emacs / XEmacs で TeX 文章を書くときに生ずる手間を軽減してくれる、TeX 入力支援環境であり、ショートカットキーで TeX のコマンドや環境の挿入、さらにコンパイル、プレビュー、プリントアウトなどが行なえる。「.tex」という拡張子が付いたファイルを XEmacs で開くと自動的に AUCTeX の機能が使用できる状態になる。

## CHISE

CHISE (CHaracter Information Service Environment) は本セミナーの講師である守岡知彦氏が中心となって開発が行なわれている、世界のさまざまな文字に関する知識をデータベース化するとともに、そのデータベースをさまざまな形で利用する新しい文字処理アーキテクチャを開発するものである。

XEmacs CHISE は、文字を文字属性の集合で表現することにより文字符号に依存せずに文字を処理することを目指した「Chaon モデル」の XEmacs に基づく実装である。

Ω/CHISE は、文字データベースより得られる文字オブジェクトの情報に基づいた、柔軟かつ拡張性の高い組版システム的设计と実装を行うものである。

### **OTF パッケージ**

OTF パッケージとは、OpenType フォントに含まれている文字を L<sup>A</sup>T<sub>E</sub>X で使えるようにしたパッケージである。通常日本語 T<sub>E</sub>X で使えるのは、JISX208 仕様のフォントだけで JIS 第一水準 (2,965 文字)、第二水準 (3,388 文字) の計 6,353 文字しか使えない。しかし、OpenType のフォントは JISX213 仕様に準拠し、例えば Mac OS X に標準で装備されているヒラギノフォントの場合、Adobe-Japan1-5 に対応しているため 20,317 のグリフを T<sub>E</sub>X で使えるようにできる。

### **Perl**

Perl とはプログラミング言語の一つで、テキストの検索や抽出に向けた言語である。表記法は C 言語に似ていて、インタプリタ型であるため、プログラムを作成したら、コンパイルなどの処理を行なうことなく、すぐに実行することができるスクリプト言語である。Perl は今までのテキスト処理ツールの Grep、Awk、Sed、C などの便利な部分だけを引き継いだもので、特に正規表現が強力であることが特徴である。また、インターネットのウェブサーバーにおいてユーザーのリクエストを処理する CGI (Common Gateway Interface) と呼ばれるプログラムを書くのに向いていることも特徴である。バージョン 5.8 より Unicode に完全に対応している。

## 講義・実習内容

### 1 日目

#### 第 1 講 漢字情報学序説「入力, 出力, そして検索」

安岡 孝一

コンピュータで漢字を入力、出力するための仕組みや、技術的な歴史についての解説。コンピュータ上での文字、特に漢字とはどのようなものなのか。そして、文字コードの成立と Unicode に至るまでの史的展開及び UTF-8 の特徴についての講義。



#### 第 2 講 TeX 入門「人文系研究者のための文書整形術」

真下 裕之

OS : Windows XP

ソフト : EmEditor、M's TeX Helper 2、TeX

人文系研究者にとってコンピュータを効果的に活用するためにはどうするか。TeX を使うことによって得られるメリットについての概説。Windows に TeX をインストールし、EmEditor と M's TeX Helper 2 を使って TeX の基礎を実習。



## 2 日目

### 第 3 講 $\text{\LaTeX}$ 応用「効率的な文書作成のために」

守岡 知彦

OS : Linux (KNOPPIX)

ソフト : XEmacs、AUCTeX

KNOPPIX を使用するにあたっての設定等の説明および Linux の操作の説明。エディタ XEmacs の基本操作の実習。そして AUCTeX を使った  $\text{\TeX}$  のコマンドの入力の仕方の講義と演習。



### 第 4 講 多言語化概論「 $\Omega$ , $\text{\TeX}$ と OTF パッケージ」

宮崎 泉

OS : Linux (KNOPPIX)

ソフト : XEmacs、CHISE、 $\text{\TeX}$ 、

OTF パッケージ

XEmacs CHISE の使用方法の説明と部品から漢字生成の演習と  $\Omega$ /CHISE を使った多言語環境における  $\text{\TeX}$  文章作成の実習。OTF パッケージを使った Unicode や CID 番号を使って標準の日本語  $\text{\TeX}$  で扱えない漢字を出力する方法の実習。



### 3 日目

#### 第 5 講 マークアップ概論「マークアップ概念の導入 – L<sup>A</sup>T<sub>E</sub>X を例に」

守岡 知彦

OS : Linux (KNOPPIX)

ソフト : XEmacs、AUCTeX

テキストのマークアップとしての T<sub>E</sub>X コマンドの有効性の概説。T<sub>E</sub>X コマンドでマークアップすることによるテキストの検索や抽出の具体的な例を用いた講義と実習。



#### 第 6 講 XML 入門「The Basics – Understanding and Using XML」

クリスティアン・ウィテルン

OS : Windows XP

ソフト : <oXygen/> XML Editor

XML および TEI とは、どのようなものなのかの講義。<oXygen/> を使った XML でマークアップされたテキストを組版し PDF に出力する仕方の演習。





4 日目

**第 7 講 XML 応用「Applications of XML – XML for Text Encoding」**

クリスティアン・ウィテルン

OS : Windows XP

ソフト : <oXygen/> XML Editor

TEI によるマークアップ構想を概観しながら、XML 文章の特徴を具体的に説明及び実習。

**第 8 講 XML プログラミング入門「正規表現と XML 文書の操作」**

師 茂樹

OS : Linux (KNOPPIX)

ソフト : XEmacs、Perl

正規表現とは何か。正規表現を使うと、いったいどのような検索ができるのかを具体例を用いた解説。Perl の DOM モジュールを使った XML テキストの検索や文字抽出の講義と実習。



## 5日目

### 第9講 XSLT 入門「XML で XML を処理する？」

岩井 茂樹

OS : Windows XP

ソフト : <oXygen/> XML Editor

文献テキストを XML でマークアップすることによってどのようなことができるのか、共同研究班の会読テキストの『元典章』の実例を紹介。XSLT で XML テキストの処理の仕組みを Perl の場合と比較した説明と演習。



### 第10講 多言語情報処理概論「漢字政策と漢字情報処理」

安岡 孝一

日本・中国・台湾・韓国の常用漢字の歴史と各国独自の文字コードの仕組みについての概説。Unicode、UTF-8 や Adobe の CID フォントのコード番号についての講義。

