

オブジェクト指向設計法によるチベット文字認識システム

- 認識前処理部について -

布宮千夏子^{†1}

小島正美^{†2}

川添良幸^{†3}

概要: 本論文では、チベット文献自動認識システム構築のための初期実験として、認識システムの前処理部について、オブジェクト指向設計技法であるUML(Unified Modeling Language)によるシステム設計を行うことにより、ユーザーの側に立ったシステム構築が可能であることを示す。また、本手法がチベット文字固有の特徴を活用した認識手法に有効であることを確認する。さらに、認識処理に加え、チベット文献に関する情報を容易に蓄積できるデータベース機能を兼ね備えた一貫したシステム作りを目指す。

Recognition System of Tibetan Characters by Using Object Oriented Design

NUNOMIYA Chikako^{†1}, KOJIMA Masami^{†2} and KAWAZOE Yoshiyuki^{†3}

Abstract: In this paper, we design a system of automatic character recognition for Tibetan characters by UML(Unified Modeling Language), which is a newly developed method of Object Oriented Design. The purpose of this study is to support the research on Tibetan literatures and to establish the effectiveness of character recognition method using characteristics of Tibetan characters. This study also tries to accumulate information on Tibetan characters as a Tibetan database.

^{†1} 山形県立山形職業能力開発専門学校能力開発支援課, 〒990 - 2473 山形県山形市松栄二丁目2番1号

2-2-1, Matuei Yamagata-shi Yamagata-ken 990-2473 Japan,

Yamagata Vocational Skills Development Institute Skills Development Support Division

^{†2} 東北工業大学, Department of Electrical Communication, Tohoku Institute of Technology

^{†3} 東北大学金属材料研究所, Institute for Materials Research, Tohoku

1. はじめに

仏教文化の受容とチベット仏教文化の形成・伝承を記す貴重な文献資料が膨大な量として我々に残されている。これら多くの財産は、チベット学研究者の手によって解読され、ローマ字化して再記述される。その方法として、研究者はチベット文字を1文字毎に認識し、表音文字に変換しなければならないため、これらの作業に多くの時間を費やしてきた。チベット文献をコンピュータで自動認識することができれば、インド原典、チベット訳文献、漢訳文献

などの研究者が本来の文献学に専念できる点において極めて有効であり、その実現がチベット学研究者から強く望まれている¹⁾。

本研究の最終目的はコンピュータを利用し、文献の読み込みから認識結果の出力までを一貫して行える「チベット文字認識システム」と文献情報を蓄積することができる「チベット文字文献データベースシステム」を構築することである。本システムを構築するにあたり初めにユーザであるチベット学研究者からチベット文献システムに対する要望を出していただい

た。その初期実験として認識を行う前段階である「認識前処理部」についてシステムの設計から実装までを行った。図1にそのシステムイメージを示す。

従来までの実験において、チベット文字の認識には、チベット文字固有の特徴を活かした認識手法として、「類似文字による認識」が認識率の向上につながることが分かっている。その

ため、本論文におけるシステムの設計にはチベット文字固有の特徴抽出に有効と考えられるオブジェクト指向設計手法であるUML (Unified Modeling Language)を用いて行うこととした。さらに、本設計手法を用いることにより、ユーザにとって利用し易いコンポーネントによるGUI (Graphical User Interface) の実現、及びデータベース機能の追加を試みた。

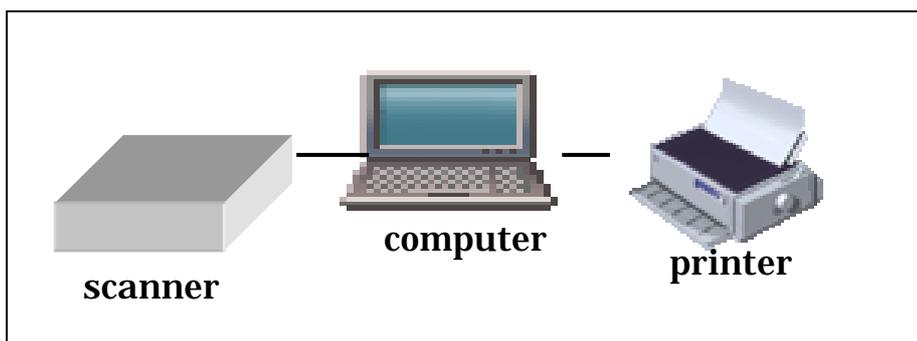


図1 チベット文字認識システムイメージ
Fig. 1 Recognition system of Tibetan characters

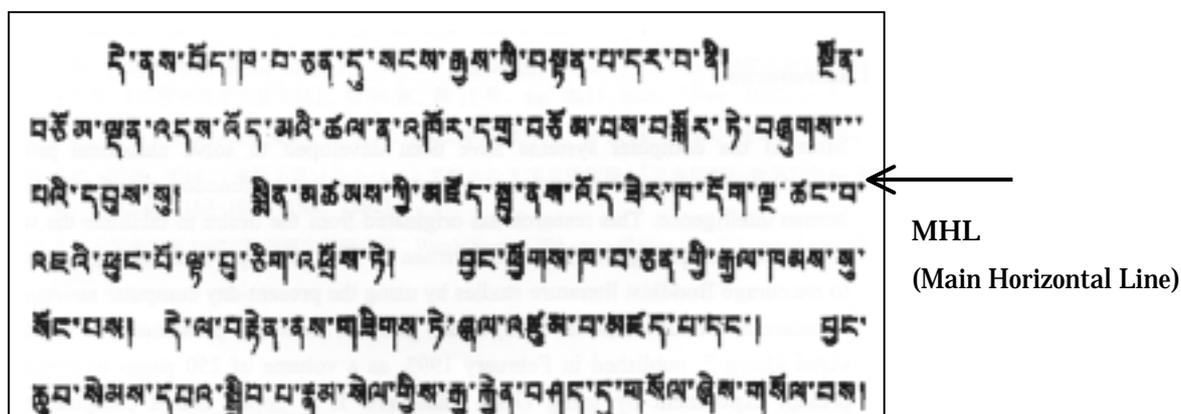


図2 チベット文字文献
Fig. 2 Literature written by Tibetan characters

2. チベット文字

実験に用いる文献は、図2に示すようなチベット文字で書かれた文献である。チベット文字は文法的にも文字構成的にも多くの特徴を有している。

1つの例として、チベット文字は、基本30子音と母音4文字により1音節構成されている。その1音節構成の最大要素は図3に示すように基字(基本文字)、付頭字、付足字、前接字、後接字、再後接字、母音の7種類から構成

され、基字と付頭字、基字と付足字の組み合わせは重層字と呼ばれる。さらに母音“ i ”、“ e ”、“ o ”に相当する文字は上部に付き、母音記号“ u ”に相当する文字は下部に付き、母音の“ a ”に相当する記号はチベット文字に内在している。これらの文字を総計すればその組み合わせは615文字程度となる^{2,3)}。

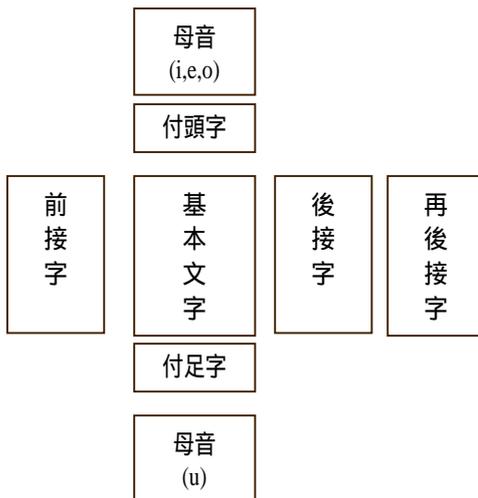


図3 チベット文字の1音節構成
Fig.3 Structure of Tibetan script

3. UML 設計

チベット文字の自動認識システムにおいて重要なのは、その固有の特徴を各プロセス毎に取り入れ、システムの拡張性や再利用性を実現することである。そのため、これらの要求を満たすためのUMLによる設計方法を採用した。また、ソフトウェアを部品化するコンポーネント指向的思考方を採用することで、チベット学研究者にとって煩わしい操作やわかりにくい認識ロジックをブラックボックス的に処理することで、実用的に利用可能なシステム実現を目指した。そのため、チベット文字認識システムをユーザーサイドおよびシステムサイドからアプローチし、分析することで、よりチベット学研究者が利用し易いシステム構築を目指した⁴⁻⁶⁾。

3.1. ユーザーサイドからのアプローチ

まずは、ユーザーの声を可能な限りシステムに取り入れるため、チベット学研究者がシステムを操作しているイメージでシナリオを作り、その内容に沿った GUI (Graphical User Interface) を設計した。そのデザイン設計画面を図4に示す。システムのメイン画面は、先のシナリオから問題点を抽出・分析し、ユーザーの利便性を重要視してデザインした。

その結果、主な考察ポイントとして、「文献の読み込みから認識結果の出力までを1画面で行うこと」、「文献情報、認識結果情報、文献情報を1画面で行うこと」、「操作はすべてメニュー表示すること」などに重点をおくこととした。その後、3.2、3.3項でシステムサイドからのアプローチを行っている。

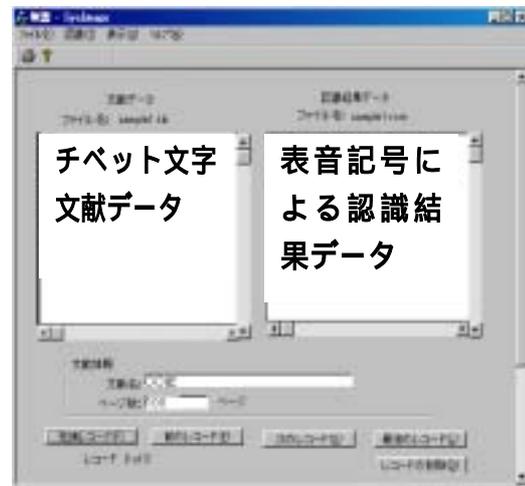


図4 システム画面イメージ
Fig. 4 Example of display system

3.2. ユースケースからクラス

本論文は、更に詳細なシステム機能を分析・設計するため、UMLによるシステム設計を行った。その例として図5に問題記述からユースケースの抽出までの流れを示す。問題記述からユースケース抽出は、チベット学研究者からの

要求によるシステムの制約事項や機能、効率、信頼性、使い易さ、拡張性などを記述している。図5では機能のみについて記述している。

抽出されたユースケースによりシステム設計の展開として以下の作業を行っている。その例を示す。

1) ユースケース図の作成

ユースケース図は、図6に示すようにユーザから見たシステムの機能をモデル化しており、システムの利用者の役割を示す「アクタ」とシステムの機能を示す「ユースケース」との関連を示している。ユースケース図を作成することにより、よりユーザの立場で利用可能なシステムを実現することが可能となる。例えば、図6の例では、アクタであるチベット学研究者は、文献読み込み、認識、データ削除などといった様々な機能を要求していることを意味する。これにより、チベット学研究者(アクタ)が必要とするシステム機能がダイアグラムによって表現され、機能の過不足や設計時におけるシステムに対する拡張が効率よく適切に行われることとなる。

2) アクティビティ図の作成

アクティビティ図は、図7に示すようにチベット学研究者であるアクタの立場で、システム中で行われる複数のユースケースを表現しており、ユースケースのメソッド(振る舞い)についてモデル化している。図7では、チベット学研究者(アクタ)がスキャナに文献をセットしてから文献イメージデータの読み込み処理が終了するまでの一連のロジックを表しており、従来のフローチャートにアクタの要素を付加した形で表現している。従ってチベット学研究者の立場を保ちつつメソッドを設計することが可能となっている。

上記1)や2)の外にも、幾つかのUML設計プロセスを実行し、それらの結果を踏まえな

がら、以下では、クラス設計プロセスまでを行う。

例として、図8にクラス関連図の一部を示す。クラス図では、クラス図を先の設計プロセスで記述された問題記述等より「名詞」だけに注目して抽出する。そして、抽出されたクラス間について、互いの関係をモデル化している。図8の例は、「チベット文字クラス」には「文献クラス」、「表音記号クラス」、「文献情報クラス」が1対多の関係で関連していることを意味している。個々の機能(ユースケース)についてモデリングし、クラスを抽出する設計プロセスは、すべてアクタであるチベット学研究者の立場で行わなければならない。すなわち、アクタにとって有効なロジック、ユースケース(機能)にとって必要なデータの発見に繋がる設計手法となっている。

3.3. データベース

チベット学研究者にとってチベット文字の自動認識システムのほか、文献データ、文献情報、認識結果データを蓄積・検索・活用できるデータベースシステムが必要である。そこで、本論文では、チベット文字認識システムに認識機能とデータベース機能の2つの機能を併せ持つシステムの構築を試みた。チベット学研究者がシステムに文献を認識させると同時に、「必要な情報をデータベースとして蓄積する」、「必要な時に必要な情報を即座に検索できる」チベット文字認識システムの設計を行った。この2つの機能を併せ持つことは、チベット学研究者が本来の研究を効率よく進めるうえで大切な要件といえる。

<問題記述>

- ・チベット学研究者は、本システムによりチベット文字文献を表音記号文字に翻訳する。
- ・スキャナにチベット文字文献をセットして本システムに読みこむ。
- ・スキャナから読みこんだチベット文字文献を記録する。
- ・記録された文献は、文献のイメージそのものである。
- ・
- ・
- ・一度記録した文献のイメージファイルと表音記号のテキストファイルの両方を削除する。
- ・ファイルの記録は自動的に行われ、記録ミスを防ぐ。
- ・記録されたファイルの検索は、簡単な操作で行う。

<アクタの抽出>

アクタ名	概要
チベット学研究者	チベット文字文献の記録、チベット文字文献を表音記号に変換、表音記号ファイルの記録、ファイルの管理を行う人

<ユースケースの抽出>

ユースケース名	概要
文献を読みこむ	新規文献を読み込む
文献を保存・表示・印刷する	読みこんだ文献を保存・表示・印刷する
認識する	読みこんだ文献を認識して表音記号にする
表音記号を保存・表示・印刷する	認識後の表音記号データ保存・表示・印刷する
文献情報を保存・表示・印刷する	文献情報を保存・表示・印刷する
データを削除する	文献・表音記号データのファイルを削除する
データを検索する	保存したデータを検索する

図5 問題記述からユースケース抽出

Fig. 5 Extraction of Use Case from description of problems in Tibetan character

具体的なデータベース機能については、前述した「文献データ」、「文献情報」、「認識結果データ」のデータを各コンポーネント（部品）として扱うことにより行う。この「コンポーネント指向」を用いることにより、ユーザーであるチベット学研究者が各データをブロックイメージで取り扱える。

例えば、各データブロックをGUIのイメージそのままに見せることが出来る。また、必要最小限のデータ保存情報により検索することができる、といった利点も生まれる。図9にコンポーネントイメージを示す。図9では、「文献データ」、「文献情報」、「認識結果データ」と

いった小さなコンポーネント(データのコンポーネント)が集まり、更に大きな集合となり、単位として文献1頁分の情報(ページのコンポーネント)となる。また、グループ化したい1頁分の情報を集めれば1つのファイル(ファイルのコンポーネント)として扱える。

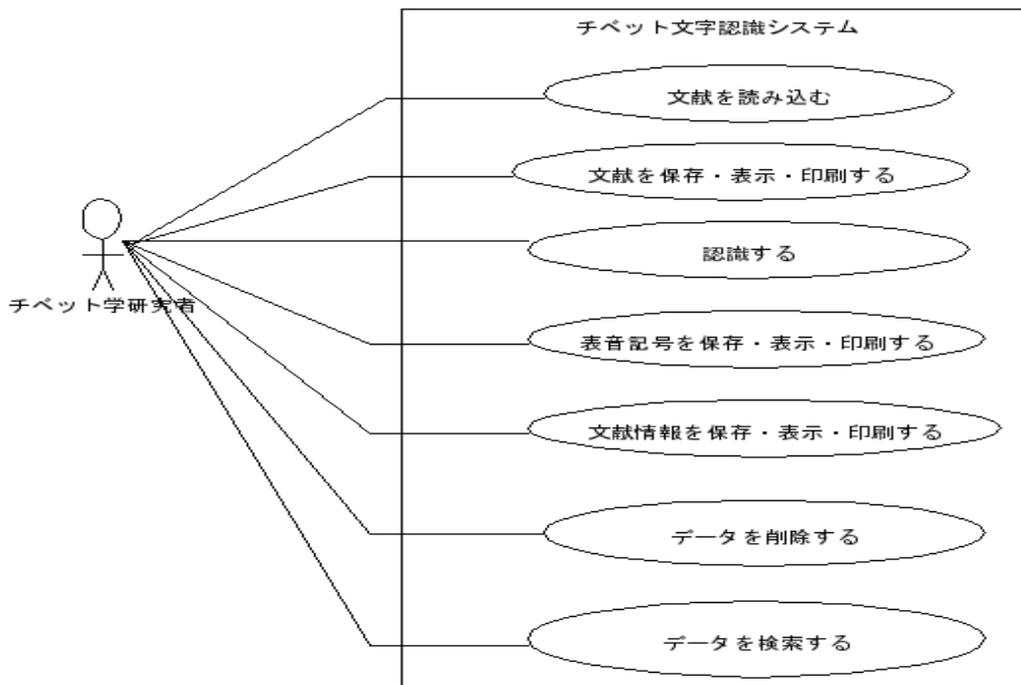


図6 ユースケース図
Fig. 6 Use Case Diagram

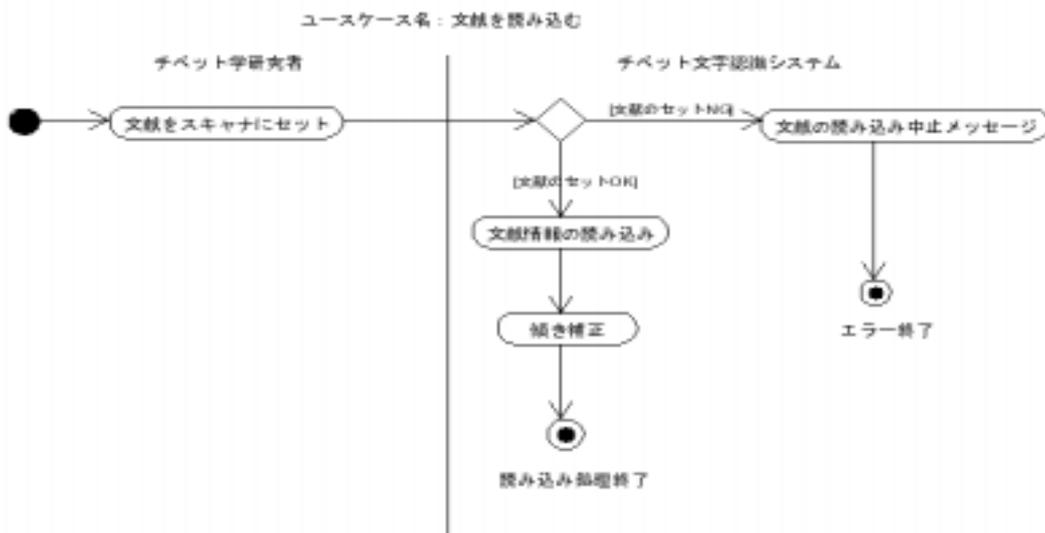


図7 アクティビティ図
Fig.7 Activity Diagram

クラス図

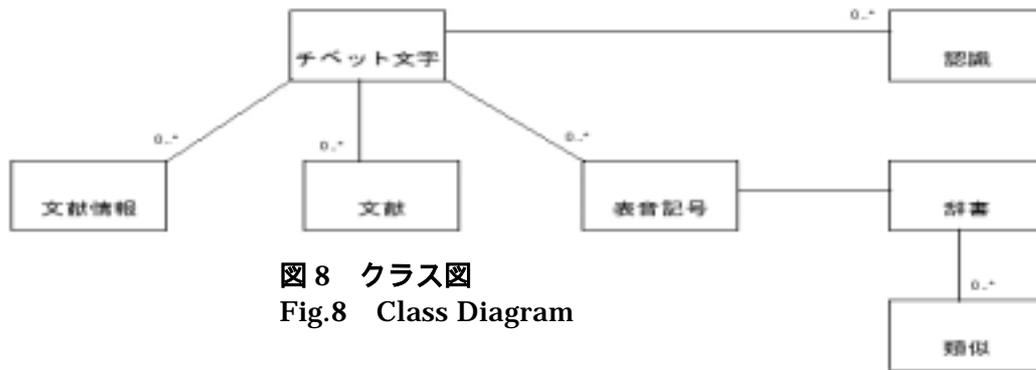


図8 クラス図
Fig.8 Class Diagram

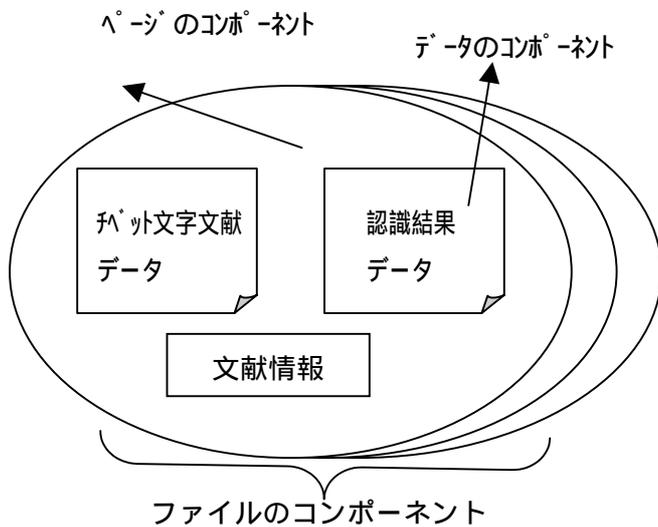


図9 データベース構造
Fig.9 Structure of the DataBase

4. 前処理部

チベット文字文献をスキャナーからシステムに読み込んだ後、認識前処理部として「傾き補正(LPP法:Local Projection Profile)」、「行切り出し」、「1文節切り出し」、「1音節切り出し」、「1文字切り出し」、「文字の正規化」を行う必要がある。本論文では、「行切り出し」までのプロセスについてシステム作成のための初期実験を行っている。

ここでは、チベット文字の特徴である情報を

システム自身が自動的に検出・取得しながら認識前処理部を行い、これらの情報を認識前処理部における重要な属性として認識部に反映させることを試みた。例えば、図2に示す主要な水平方向射影ヒストグラム(MHL:Main Horizontal Line)である。この情報を用いて認識率を向上させようとするものである。切り出しの際、様々な重要な情報が生まれ、認識されずに捨てられているのかもしれない。例えば、文節中には認識させたい文字位置情報がある。文字の中には文字の大きさ情報やホール情報などがある。これらの情報を可能な限り取り上げ、行の属性、音節の属性、文字の属性として活用する。

5. まとめ

UML設計により、可能な限りユーザである「チベット学研究者」の側に立ったシステム構築を目指して開発することができた。例えば、システム構成においては、チベット学研究者らがスキャナに文献をセットして、行の切り出し、文字切り出し、認識、データの保存・蓄積といった一連の操作について1画面中でマウスをクリックすることが容易に実現できるシステムが構築可能であることが分かった。

チベット文字の文章構成や文字構成、文字パ

ターン、MHL情報、ホール情報など、認識結果が得られるまでには、多くのチベット文字の固有情報が存在する。これらのチベット文字情報を取得し、チベット文字固有の属性として認識処理に用いることは、認識率向上に有効であることがわかった。今後、前処理部においても切り出し時のチベット文字固有の属性情報を積極的に用いることによって、切り出し率向上を図りたい。

また、データイメージを損なうことなく、複雑でないデータベース機能を付加したチベット文字認識システムの実現が可能であることがわかった。本論文では、認識前処理部のみでの実験を行ったが、今後、チベット文字の特徴情報を抽出しながら更に有効な認識ロジックを構築する予定である。

謝辞

本研究を進めるにあたり、貴重なご意見をいただいています北陸先端科学技術大学院大学名誉教授木村正行先生、東北大学名誉教授塚本啓祥先生、磯田熙文先生に感謝いたします。また、チベット活字文献の資料提供ならびにご助言を頂きました大谷大学兵藤一夫教授に感謝いたします。

文献

- 1) 塚本啓祥：インド文学の形成と展開、「サンスクリット・チベット語のコンピュータによる総合研究」、東北大学特定領域研究組織 TURNS017 - 報告書(Feb. 1989); 磯田熙文：チベット文字の特色とコンピュータ利用について、ibid..
- 2) 小島正美、布宮千夏子、川村隆庸、秋山庸子、川添良幸：オブジェクト指向設計によるチベット活字辞書を用いた類似文字認識、情報処理学会論文誌、Vol.36, No.11, pp.2611-2621,(Nov. 1995).
- 3) 小島正美、布宮千夏子、川村隆庸、秋山庸子、川添良幸、木村正行：オブジェクト指向設計によるチベット文字認識について、情報処理学会第23回人文科学とコンピュータ研究会 (Sep. 1994).
- 4) ジョゼフ・シュムラー著、長瀬嘉秀監訳、独習UML改訂版、SHOEISH.
- 5) 鈴木重徳、倉骨彰、佐野元之、垣花一成著、IIOSS ~ UML に基づく設計 / 開発環境のすべて ~、ASCII.
- 6) イヴァー・ヤコブソン、グラディ・ブーチ、ジェームズ・ランボー著、日本ラショナルソフトウェア株式会社訳、藤井拓監修、UMLによる統一ソフトウェア開発プロセス ~ オブジェクト指向開発方法論 ~、SHOEISHA.