

Java と XML

飯島 正

慶應義塾大学理工学部

アブストラクト

拡張可能なマーク付け言語 XMLとプログラミング言語 Javaの組み合わせが注目されている。XMLによって提供される情報表現, XMLが影響を与える業務ワークフロー, XMLを利用したソフトウェアアーキテクチャという3点から, Javaを用いたソフトウェア開発との関連について論じる。

Java and XML

Tadashi Iijima

Faculty of Science and Technology, Keio University

Abstract

The combination of XML(eXtensible Markup Language) and programming language Java has attracted a great deal of public attention.. In this paper, relationship of Java and XML is discussed from the viewpoints of data representation, business workflow, and software architecture.

1. はじめに

XML[1]は, 拡張可能なマーク付け言語 (eXtensible Markup Language)であり, HTML に代わるものとして注目されている。こうしたマーク付け言語は, 情報中に混入させるタグによって情報(文書, データ)を「構造」化しそこに計算機が直接扱えるメタ情報として「意味」情報を付与していくことができる。その使われ方には情報(文書, データ)の「整形/出版」, ネットワークを介したデータ「交換書式」, 柔軟なスキーマ記述による情報の「蓄積(と検索)」がある[9]。そうした使われ方には, 既存の, 文書整形組版技術である TeX, 文書管理技術である SGML, WWW 上での文書表現技術である HTML からの影響が明らかに見受けられる。が, それらは同時に, ネットワークというインフラの上での「情報流通」における重要な3つの観点を提供しているともいえる。

一方, Java も, ネットワーク上での利用や流通を意識したプログラミング言語である。Java Applet は WWW サーバからダウンロードされて, WWW ブラウザ上で実行されるプログラム配布形態である。そうしたネットワークを介して流通するプログラム自体は「文書」の一種といえないことはない。また, 実行中のプログラム(オブジェクト, コンポーネント)の実行状態もデータでありネットワークを介して流通できる。Java はそうした観点からも XML 技術と関連が深い。

他に Java と XML との接点を求めると, XML 文書の処理に使われる DOM(文書オブジェクトモデル)は, その名の通り Java 言語が基礎とするプログラミングパラダイムであるオブジェクト指向モデルを採用しており, 両者の親和性は高い。そのため, Java は XML の処理記述言語としても広く使われている。XML のパーシングをはじめ, いろいろな XML に関連した処理

のための API が数多く提供されている[2-6]。

続く3章では, 順次, XMLによって提供される情報表現, XMLが影響を与える業務ワークフロー, XMLを利用したソフトウェアアーキテクチャという3点から, Javaを用いたソフトウェア開発との関連について論じる。

2. XMLと情報

2.1 XMLで扱う情報

XMLで扱う情報としては, いわゆる文書とデータがあるが, この両者の境目は厳密ではない。データといった場合には伝票やカタログなどのように「意味」を与える項目タグと「値」とが対応づけられることが多く, しばしば項目-値リストや表の形態をとる。そのように構造化されていても, その要素は原子的なものが基本で, 計算機での検索/操作に適した定型データを意図する。XMLではDTD(文書型定義)によって, そのデータスキーマ(メタデータ)を定義する。一方, 文書(テキスト)にしたところで, 構文的なタグによって与えられる木や網(ハイパーテキスト)のような文書構造上が基本になるが, 必ずしも構文だけでなく意味的なタグ付けもなされる。すなわち, 不定型データにすぎないともいえる。もともと, 使用言語の文法レベルのような詳細な構文はタグでは表現されないため, テキストデータの場合には, データを集めてXMLで構造化するというよりは, テキストの構造の一部をタグで明示化するという方向性の違いはある。

筆者も 1995 年終わりから 1996 年にかけて, その両者(データとテキスト)を統一的にみて, SGMLといった文法表現によって広く情報のスキーマを与えることを提唱している[8]。また, 一般に文書利用の面からみても, 論文等の文書中に表が含まれていることは多く, 文書とデータを区別して, あえて「文

書へのデータの混入」ととらえることに意味があるとおもえない。逆に、データベース研究における技術的な面からも、テキストを対象とする全文検索技術や、柔軟なスキーマ表現による半構造データベースが着目されるなど、その両者を区別することの意義が薄れつつある。

2.2 データの共有性/独立性/共有性

一般にデータは、それを取り扱うアプリケーションよりも寿命が長く、かつ、複数のアプリケーションから共有されるため、特定のアプリケーションに依存しない独立性/中立性が要求される。Javaがプラットフォーム中立なプログラミング言語であるのと同じように、XMLはいろいろなアプリケーションにとって中立な情報表現となる。Javaのスローガンの一つであるWORA™(“Write Once, Run Anywhere!”)がある。XMLデータにしても、これと同じように、WOUA (“Write Once, Use Anywhere!”)というスローガンを作ることができる。一回、記述したデータ(ならびにメタデータ)は、いろいろなところで使うことができる。たとえば、EC(電子商取引)におけるカタログデータなどは、加工を含めて広く再利用することができる。

また、特にXMLにおいては、DTDといったメタデータを用いてデータをパーズングすることで、経年的なデータ書式の改良(バージョンアップ)や、異種システムとの連携において柔軟性を発揮することができる。しかし、これにはトレードオフがある。たとえば、基幹システムのような堅固に結合されたシステム間の交換データは固定的な書式にしたがえばよく(データ書式の揺らぎに対応する柔軟性は必要ない)、それをタグ付きのテキストに変換して送信し、通信の度にパーズングするのは、かえって無駄である。これは丁度、プログラムのインタープリタ(逐次解釈系)の実行とコンパイラでの実行の相違に相当する。そこで、メタデータをコンパイルしてアダプタに変換するようなツールの利用が要求される(このアダプタは、ある程度、JavaにおけるVM(仮想機械)の位置づけに対応する)。

2.3 プロセスの表現

プログラムそのもの、そしてそれだけでなく実行中のプロセスの状態(直列化データ)もネットワーク上で流通するXMLデータとして取り扱うことができる[9]。特にJavaをベースにそうした技術が蓄積されている[2,5]。いわゆる移動エージェントは、その上に構築することができる。ネットワークを介して相互接続された環境においてワークフロー管理/制御を行うには、こうした技術が有効である[10]。

3. XMLと業務ワークフロー

XMLによって伝票(や、いわゆるトランザクション)のような業務間でやりとりされるデータの情報を定義することは、プロトコル(たとえばECのためのOTP(Open trading Protocol)[7])などや、更には、その情報を取り扱うプロセスさえも規定してい

く(ヒントになる)ことがある。そうしたプロセスは、業務をモデル化する上で重要である。これは粗く「縦と横」に分けることができる。横のプロセスは、蓄積/利用を意図したXMLデータを作成していくプロセスとして、縦のプロセスは業務間でのデータ交換を通してデータを加工していくプロセスとして認識される。たとえば、縦のプロセスを追跡することは、たとえば、製造から販売までの流通パイプライン上のデータとプロダクトのフローをリエンジニアリングするといったSCM(サプライチェーンマネジメント)等の業務プロセス改善(BPR)につなげることができると期待できる。

4. XMLとソフトウェアアーキテクチャ

J2EE (Java2 platform enterprise edition)においては、3層C/Sモデル(3-Tier Client/Server model)をベースに各種のJava関連技術を利用した幾つかのシステムアーキテクチャモデルが提示されている。そうしたモデルで注目されているのが、WWWベースの軽量なクライアント(Thin Client)と、中間層としてビジネス論理を持つEJB(Enterprise JavaBeans)ならびにJSP(Java Server Pages)とServletをおき、データ層として、DBを配置するようなアーキテクチャを示している[11]。この中でもXMLの利用が期待されている。

謝辞

INS エンジニアリングの大野邦夫氏、東芝の河込和宏氏他の皆様の御教示ならびに議論に感謝します。

参考文献

- [1] XML/SGML サロン:「標準 XML 完全解説」, 技術評論社(1998)
- [2] 特集 Java プログラマのための XML 大研究, Java World, No.6, 1999
- [3] 青柳龍也, Java 徒然草(9), XML for Java, 同上, 6, 1999
- [4] 青柳龍也, Java 徒然草(10), XSL, 同上, 7, 1999
- [5] マークジョンソン, XML&JavaBeans, 同上, 10, 1999
- [6] 吉川和己, IDE と XML のコンビネーション, 同上, 11, 1999
- [7] Sean McGrath (株式会社クイック訳):XML開発事例 エレクトリックコマース, アスキー出版局, 1999
- [8] 飯島正: 文法に基づくデータモデル, 情報処理学会第 52 回(平成 8 年前期) 全国大会, pp.261-262, 1996.
- [9] 飯島正: XML で文法を与えるプログラミング言語, ソフトウェア工学研究会, SIGSE-120-9, pp.53-60, 1998.
- [10] 飯島, 山本, 土居: 共生・寄生エージェント・モデルのワークフロー管理への応用, 信学技報, AI98-60, 1998
- [11] Anne Thomas: Java™2 Platform, Enterprise Edition - Ensuring Consistency, Portability, and Interoperability, Patricia Seybold Group, 1999