

ネットワークにおける XMLデータ通信

富士ゼロックス情報システム（株）

村田 真

 Fuji Xerox Information Systems

目次

- データ通信におけるXMLの利点
- 現時点で採用可能な方式
- 将来の方式

 Fuji Xerox Information Systems

データの送受信における XMLの利点

- きわめて簡単に使える
- 完成度の高いXMLパーサがどこにでもある
- 日本語が通る（徹底的な国際化）
- XMLが提供する階層構造でほぼ十分
- XMLはテキストとして扱える

現時点で可能な方式

プロトコル

- HTTPでPOST, GETを使う
- 受信側はCGI, Servlet

XMLの形式

- DTDを持たない整形XML文書
- DTDを持つ妥当なXML文書
 - DTDはどこから？
 - デフォルトと実体と記法は使用しないこと

符号化方式

- UnicodeとJIS X 0208との変換揺れは必ず問題になる

\\ - - ¥ @ £ ~

- UTF-8かUTF-16を薦める

送信側でのXMLデータ作成

- write文を繰り返してXML文書を作成
 - とっつきやすい
 - XML文法エラー / 予想外の結果
- DOMでXML文書をメモリ上に生成してから書き出す
 - DOMを覚える必要あり
 - 文法エラーはほとんど発生しない

送信側でのプログラミング言語

- write文を使うのならとくに選ばない
- DOMを使いたいのなら、DOMが使える言語(Java, C++, perl, JavaScript, VBScript, ...)

受信側でのXMLデータの解読

- SAXか他のイベントベースAPIを用いる
- DOMを用いる

受信側でのプログラミング言語

- DOMを使いたいのなら、DOMが使える言語DOMが使える言語(Java, C++, perl, JavaScript, VBScript, ...)
- ServletにするのならJava

将来の方式

動機

- SAXではデータがメモリに作られない
- DOMでは、メモリに作られるデータは、つねにElementクラスとTextクラスのインスタンス

目標

- アプリケーション固有のクラス定義に従うデータを、XML文書からメモリ上に自動的に生成する。
- アプリケーション固有のクラス定義に従うデータから、XML文書を自動的に生成する。

手段

- XMLのための強力なスキーマ言語(データ型...)
- XMLのスキーマとプログラミング言語のスキーマとの関連付け
- XML文書とデータとの間の自動変換

XMLスキーマ

- XML Schema(W3C)
- XML-Data Reduced(Microsoft)
- SOX
- DDML
-

DTDにあるデータ型(1)

- #PCDATA (要素)
- CDATA (属性)
- ID (属性)
- IDREF (属性)
- IDREFS (属性)

DTDにあるデータ型(2)

- ENTITY (属性)
- ENTITIES (属性)
- NMTOKEN (属性)
- NMTOKENS (属性)
- NOTATION (属性)

XML Schemaのデータ型(1)

- string
- boolean
- real
- timeInstant
- timeDuration
- recurringInstant
- binary
- uri
- language
- NMTOKEN
- NMTOKENS
- Name
- NCName



Fuji Xerox Information Systems

XML Schemaのデータ型(2)

- ID
- IDREF
- IDREFS
- ENTITY
- ENTITIES
- NOTATION
- decimal
- integer
- non-negative-integer
- positive-integer
- non-positive-integer
- negative-integer
- date
- time



Fuji Xerox Information Systems

マッピング

- XMLのスキーマからプログラミング言語のクラス定義を生成する
 - XML Data binding facility for Java(Sun)
 - XML/Value RFP(OMG)
 - ...
- XMLスキーマとプログラミング言語のクラス定義を関連付ける
 - Coins



新たなプロトコル

- SOAP(HTTPベース)
 - Internet Draft (Informational)
 - M-POST



まとめ

- XMLを用いたデータの送受信は容易
- HTTP
- SAXとDOM
- XML Schema?