発表概要

ストリームベース XML 処理を支援する Java クラス生成器

杵淵哲也 岩崎英哉

XML 処理において,タグごとにイベント処理を記述する SAX や StAX といったストリームベース API を用いた記述は,木構造を直接扱う DOM などのツリーベース API を用いた場合に比べてメモリ効率が良いという利点を持つ一方,パーサの処理が XML 文書のどの部分まで進んだのかを追跡する必要があるため,記述が煩雑になりがちであり,プログラマへの負担が大きいという欠点がある.本発表では,ユーザにあらかじめアクセスする要素のパターンを記述してもらい,そのパターンに適合する要素を取り出すアクセサを自動生成する機構を提案する.自動生成されたアクセサは Java のクラスとして提供され,主要なストリームベース API である StAX を用いて処理を行う.煩雑な記述はアクセサによって隠蔽されるので,ユーザは生成されたクラスを用いて簡潔な記述により XML 処理を行うことができるようになる.

A Generator of Java Class Files for Stream-based Processing of XML Documents

Tetsuya Kinefuchi† and Hideya Iwasaki†

For processing XML documents, stream-based API such as SAX and StAX handles tagoriented event-driven programs and has the advantage of memory efficiency, compared with tree-based API such as DOM that handles tree structure directly. However, programs using stream-based API are complex because they are necessary to track how far processing has proceed within the XML document, and thus impose large load on the programmer. To resolve this problem, we propose a mechanism that automatically generates accessors to the elements that are specified by the user. Each accessor is implemented as a method that uses StAX API. The proposed mechanism enables the user to write more concise XML processing programs, because complex descriptions are encapsulated within the accessors.

(平成 18年 10月 12日発表)