

## SUMMARY OF Ph.D. DISSERTATION

School Open and Environmental Systems	Student Identification Number	SURNAME, First name JIN Zhe
<b>Title</b>  A Study on XML Publishing Language PPX		
<b>Abstract</b> <p>Since XML is widely used for data exchange, data description and data processing, a tool for formatting XML data is an important research issue. There are widely used tools for converting XML code into HTML code such as XSLT and XQuery. These tools however are procedural and cumbersome. Thus, there is a significant need for a handy tool to convert XML data into HTML data.</p> <p>This research proposes a new language called Pretty Printer for XML (PPX). This language is based on the extended TFE of SuperSQL, which converts RDB into HTML. PPX has to overcome two difficulties. First, XML is less structured than RDB. Specifically, sibling nodes may have many different data structures. Second, the order of siblings may have a significant semantic meaning in XML. PPX solves the first problem by introducing a conditional branch statement (IF-THEN-ELSE). The second problem is handled by introducing a repeat operator, which guarantees to retain the original order of data even after sorting, grouping and eliminating duplications.</p> <p>PPX provides two modes of layout planning: manual and automatic modes. With the automatic mode, PPX invokes a layout decision algorithm which automatically formats data in HTML. This is a convenient tool for users since they do not have to investigate the structure of the XML code nor design the HTML layout.</p> <p>An experiment of applying PPX to the W3C queries is conducted. The results show that PPX can successfully express 61 out of 76 queries in total. In addition, it is confirmed that PPX can publish a large scale XML data set in a neat manner suitable to its purpose, and that PPX is better than XSLT and XQuery as a tool to generate HTML tables of various structures.</p>		

# 主 論 文 要 旨

報告番号	甲 乙 第 号	氏 名	金 哲
主論文題目：  XML 整形出力言語 PPX に関する研究			
<p>(内容の要旨)</p> <p>汎用マークアップ言語 XML はデータの交換，情報の記述・処理など様々に利用されている．その変換と出版は格納や検索とともに重要な研究課題である．XML の内容を任意にレイアウトしてから書式情報を持たせ，HTML のタグを与えることによって Web ブラウザに表示することができる．このような目的に XSLT や XQuery などの利用が広く普及しているが，そのコーディングは複雑で，手軽に HTML 化を行うための言語のニーズは高い．</p> <p>本研究では XML から HTML への変換を行なう PPX (Pretty Printer for XML) 言語を提案する．PPX は関係データベースに対して同様のことを行なう SuperSQL に基づいているが，イレギュラーな XML ノードを持つ異なるデータ構造に対して異なるレイアウト方法を指定するために条件分岐構文を導入した．</p> <p>また，文書中心の XML で重要な意味を持つ兄弟ノードの出現順序を保存するため，元の出現順にソート，グルーピング，重複削除などの処理を行う順序保存型の反復演算子を追加した．</p> <p>更に，部分 XML を元の構造を反映して自動的に HTML 化する自動レイアウト演算子を導入し，XML インスタンスの統計情報に基づいたレイアウト自動決定アルゴリズムを開発した．</p> <p>評価実験では W3C の Query Use Cases に PPX を適用し，76 のサンプル XQuery 質問のうち，61 について PPX によって等価な結果データを得る表現が可能であることを確認した．また，XML データの HTML への変換に必要なコーディングについて，XSLT や XQuery よりも生産性が高いことを評価実験によって示した．</p>			