

동적 웹 콘텐츠 변환을 위한 규칙기반 전송 환경 해석

장민수⁰ 김재홍 손주찬
한국전자통신연구원, 전자거래연구부
{minsu⁰, jhkim504, jcsohn}@etri.re.kr

Rule-Based Delivery Context Interpretation for Dynamic Web Content Transcoding

Minsu Jang⁰ Jaehong Kim Joo-chan Sohn
EC Dept., Electronics & Telecommunications Research Institute

요 약

인터넷에 접속하여 웹 콘텐츠를 이용할 수 있는 장치의 종류가 다양해짐에 따라 장치 독립적으로 웹 콘텐츠를 제공할 수 있는 기술의 필요성이 커지고 있다. 장치 독립적으로 웹 콘텐츠를 제공하기 위해서는 동적으로 전송 환경을 인지한 후 인지된 전송 환경에 적합한 형태로 콘텐츠를 조합 및 변환할 수 있어야 한다. 본 논문은 전송 환경을 인지하기 위해 필요한 전송 환경 해석 기능을 효과적으로 수행할 수 있는 규칙 기반 전송 환경 해석 기법을 설명한다. 본 기법을 활용하면 유연하고 동적인 변환 시스템 구축이 가능하므로 변환 기능의 손쉬운 변형 및 확장, 새로운 전송 환경에 대한 적응성 향상 등의 장점을 취할 수 있다.

1. 서론

인터넷에 접속하여 웹 콘텐츠를 이용할 수 있는 장치의 종류가 다양해짐에 따라 장치 독립성(Device Independence)[1] 지원은 웹 기술의 당면 과제로 떠오르고 있다. 장치 독립성 지원 기술은 크게 장치 독립적 저작(Device Independent Authoring)과 웹 콘텐츠 변환(Web Content Transcoding)이라는 두 분야로 나누어 볼 수 있다. 전자는 장치 독립적인 서비스가 가능하도록 웹 콘텐츠를 저작하는 기술이고, 후자는 작성된 웹 콘텐츠를 동적으로 전송 환경에 적합한 형태로 적응(Adaptation) 및 변환(Transcode)하는 기술이다.

상술한 장치 독립성 기술에 있어 중요한 요소는 전송 환경이다[2,3]. 전송 환경은 웹 클라이언트가 웹 자원에 접근하고 그 자원을 이용하는 과정에 개입하는 모든 환경적 요소들 - 사용자, 장치, 웹 브라우저, 네트워크 등 - 의 집합을 가리킨다. 장치 독립적인 서비스를 제공하기 위해 웹 콘텐츠 서비스 제공 서버는 전송 환경을 구성하는 요소들의 특성을 인지하고 그 환경에 적합한 응답을 보낼 수 있어야 한다. 전송 환경 인지는 전송 환경에 능동적으로 반응하기 위한 필수적 기능으로서, 장치 독립적 저작 시스템 및 웹 콘텐츠 변환 시스템 모두 공히 효과적이고 유연한 전송 환경 인지 기능을 갖추어야 한다.

기존의 장치 독립적 웹 출판 및 변환 시스템들[4,5,6,7]은 전송 환경을 구성하는 여러 요소들 중 장치 특성 정보만을 추출 및 참조하며, 이 정보를 활용하는데 있어서도 몇 개의 속성값을 바탕으로 변환 기능을 결정하는 단순한 수준에 머물러 있다.

본 논문에서는 전송 환경을 구성하는 복수개의 요소들

에 대한 특성 정보들을 정합하고 그로부터 변환 기능을 결정할 수 있는 기법을 소개하고자 한다.

2. 웹 콘텐츠 변환 시스템

2.1 전송 컨텍스트(Delivery Context)와 CC/PP

전송 환경을 해석하기 위해서는 전송 환경에 대한 정보를 형식적으로 기술한 데이터가 필요한데, 이를 전송 컨텍스트(Delivery Context)[2]라고 부른다. 전송 컨텍스트는 전송 환경을 구성하는 각 요소의 특성을 표현하는 속성들의 집합으로 구성된다. 각 속성은 속성이름과 속성값의 쌍으로 이루어진다. 한 개의 전송 환경 요소 - 예를 들면, 장치 - 에 대한 속성들의 집합을 프로파일(Profile)이라 부른다. 즉, 전송 컨텍스트는 여러 개의 프로파일로 이루어진다. W3C는 전송 컨텍스트를 기술할 수 있는 메타 정보 표현 모델로 CC/PP[8]를 제안하고 있다. CC/PP는 RDF를 기반으로 한 RDFS 확장으로서 전송 컨텍스트를 구성하는 속성들을 RDF 문장(Statement)으로 기술하도록 하고 있다. 그림 1은 CC/PP로 기술된 프로파일의 일부분이다.

```
<rdf:Description ID="MyProfile">  
  <prf:component>  
    <rdf:Description ID="HWPlatform">  
      <prf:ScreenSize>101x80</prf:ScreenSize>  
      <prf:Model>T68R1</prf:Model>  
      <prf:BitsPerPixel>8</prf:BitsPerPixel>  
      <prf:ColorCapable>Yes</prf:ColorCapable>
```

그림 1. CC/PP 프로파일의 예

CC/PP 프로파일은 세 계층으로 구성된다. 최상위에는 프로파일 인스턴스 자신을 가리키는 자기 참조가 위치하고, 그 아래에 여러 개의 컴포넌트(Component)가 기술된다. 컴포넌트는 프로파일을 구성하는 속성들을 기술 대상(Subject)이 동일한 것끼리 분류한 것이다. 마지막으로 각 컴포넌트 아래에 속성(Property)들이 기술된다. 그림 1을 보면, 아이디가 MyProfile인 리소스는 프로파일 인스턴스이고, 아이디가 HWPlatform인 리소스는 컴포넌트이며, 그 하위에 기술된 Model, ScreenSize 등은 HWPlatform의 특성을 기술하는 속성이다.

2.2. CATS

CATS(Content Adaptation & Transcoding System)[9]는 웹 콘텐츠를 전송 환경에 적합한 형태로 변환해 주는 기능을 수행하는 변환 시스템으로서 본 논문에서 소개하는 규칙 기반 전송 환경 해석 기법이 적용된 시스템이다. 그림 2는 CATS의 변환 기능 수행 과정을 전송 환경 해석 기능 중심으로 간단하게 표시한 것이다.

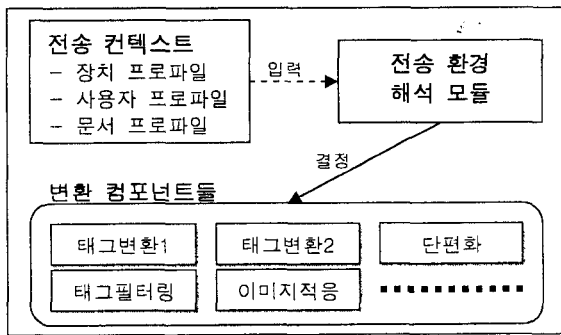


그림 2. CATS의 변환 기능 수행 기본 구조

3. 규칙 기반 전송 환경 해석

3.1 규칙 기반 전송 환경 해석의 필요성

그림 2의 변환 기능 수행 과정을 보면, 전송 환경 해석의 역할은 입력으로 접수된 전송 환경 정보를 해석하여 해당 전송 환경에 적합한 변환 기능을 수행하기 위해 필요한 변환 컴포넌트들을 선택하는 것임을 알 수 있다.

주목할 점은 전송 컨텍스트가 복수개의 프로파일로 구성되어 있다는 점이다. 서로 다른 프로파일은 상호 관련성을 가지거나 상호 상충되는 속성값을 포함할 수 있다. 예를 들어, 장치 프로파일에는 장치가 표시할 수 있는 최대 이미지 색상 수가 4096으로 기술되어 있고, 사용자 프로파일에는 표시 허용 색상수가 4로 기술되어 있다고 하자. 이런 경우 두 속성의 상호 관련성과 상대적 우선 순위 등을 고려하여 변환 기능을 결정해야 할 것이다. 기존의 변환 시스템들은 장치 프로파일에만 의존하므로 이러한 문제로부터 자유롭다. 그러나, 점차로 전송 컨텍스트를 표현하는 프로파일의 종류와 양은 증가할 것으로 보이므로 문제는 좀 더 복잡해질 것으로 예상된다.

한편, 변환 기능에 대한 요구 사항은 매우 다양하고 변화의 소지가 높다. 주어진 전송 환경이 동일하더라도 웹 콘텐츠의 부류, 웹 콘텐츠 제공자의 사업 전략, 소비자의 기호 등에 따라 실행할 변환 기능이 달라질 수 있다. 즉, 전송 환경 해석 기능은 서로 다른 요구 사항에 따라 빈번하게 고객화(Customization)될 수 있어야 한다.

전송 환경 해석 기능을 프로그램 코드로 개발하면 속성들간의 상호 관련성으로부터 변환 기능을 도출하는 복잡한 조건 분기 흐름을 일일이 철차적으로 기술해야 하므로 구현 비용이 높을 뿐만 아니라, 고객화 및 요구 사항 변경 등에 대해 유연하게 대처할 수 없다는 문제점이 있다.

규칙 기반 시스템은 사실(Fact)들간의 복잡한 논리 및 인과 관계를 선언적으로 표현하므로, 이러한 장점을 잘 활용하면 전송 컨텍스트를 구성하는 여러 속성들간의 복잡한 관계를 효과적으로 기술하여 속성 간 상충 요소를 해소할 수 있으며, 수행해야 할 변환 기능 목록을 손쉽게 결론으로 도출해 낼 수 있다. 또한 규칙 베이스를 수정함으로써 동적으로 요구 사항에 맞게 해석 기능을 변경할 수 있으므로 고객화 및 기능 변경에 매우 유리하다.

3.2. 규칙 기반 전송 환경 해석 절차

그림 3은 규칙 기반 시스템을 이용하여 전송 환경을 해석하는 과정이다.

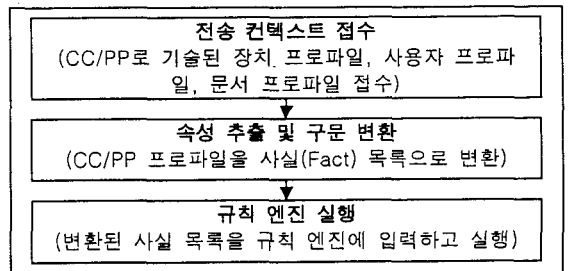


그림 3. 규칙 기반 전송 환경 해석 절차

각 단계에 대해 살펴보면 다음과 같다.

3.2.1. CC/PP 프로파일의 사실(Fact) 목록화

CC/PP에 기술된 정보를 규칙 기반 시스템의 입력으로 하여 규칙 처리를 하기 위해서는 CC/PP 프로파일을 규칙 엔진이 이해할 수 있는 형태로 변환하여야 한다. 앞서 설명한 바와 같이 CC/PP는 RDF를 기반으로 한다.

RDF는 웹 자원에 대한 메타 정보 기술 프레임워크로서, (predicate, subject, object)의 형태로 된 문장들의 집합이다. 이러한 문장 구조는 규칙 기반 시스템에서 사실(Fact)을 기술할 때 사용하는 구문 형태와 대응되어 있으므로 RDF 문장을 손쉽게 사실 구문으로 변환할 수 있다. 표 1은 그림 1에 예시한 CC/PP 프로파일을 JESS[10]의 구문으로 변환한 예이다. JESS는 전술한 바 있는 변환 시스템인 CATS에서 사용하는 규칙엔진이다.

표 1. 그림 1의 CC/PP 구문을 JESS 구문으로 변환

스타일 1	스타일 2
(ScreenSize 101x80) (Model T68R1) (BitsPerPixel 8) (ColorCapable Yes)	(ScreenSize HWPlatform 101x80) (Model HWPlatform T68R1) (BitsPerPixel HWPlatform 8) (ColorCapable HWPlatform Yes)

위 표에서 스타일 1은 RDF 문장의 *subject*부를 임의로 제거한 형태이다. 프로파일 및 컴포넌트 간에 동일한 이름을 갖는 속성이 없는 환경에서는 이러한 단순한 사실 구문으로도 충분하다. CATS에서는 스타일 1과 같은 구문을 사용한다. 스타일 2는 *subject*부를 보존한 형태로서 서로 다른 컴포넌트 간에 동일한 이름을 갖는 속성이 존재하는 경우 이러한 구문 형태를 사용하게 된다.

CC/PP 프로파일 안에는 전송 환경 특성을 표현하는 속성 기술 문장 외에 컴포넌트, 컴포넌트 타입 등 CC/PP 프로파일의 구조 정보를 기술하는 문장들도 존재한다. 전송 환경 해석 기능에는 속성 정보만 있으면 충분하므로 CC/PP 구조 정보를 기술하는 문장들은 모두 삭제하고 속성 기술 문장들만을 사실 목록화 대상으로 사용한다.

3.2.2. 규칙 기반 전송 환경 해석

규칙 기반 전송 환경 해석은 두 가지 기능으로 구성된다. 하나는 상호 관련된 속성들을 정합하는 기능이고, 둘째는 정합 내용을 바탕으로 필요한 변환 기능이 무엇인지 결정하는 기능이다.

3.2.2.1. 속성 정합

복수개의 프로파일 간 충돌을 해소하고 적절한 변환 기능 선택에 참조할 수 있는 파생 지식을 생성하는 기능이다. 상호 연관성을 지니는 속성들 간의 관계와 정합 방법을 규칙으로 기술함으로써 이러한 기능을 구현할 수 있다.

3.1절의 예와 그에 대한 정합 규칙을 JESS 구문으로 기술해 보면 다음과 같다.

```
(NoOfColors 4096) ; 장치 지원 색상 수
(PreferredNoOfColors 4) ; 사용자 선호 색상 수
```

```
(defrule resolve-no-of-colors
(NoOfColors ?cno)
(PreferredNoOfColors ?pno&(< ?pno ?cno))
=>
(assert (UseNoOfColors ?pno))
)
```

상기한 규칙은 '사용자 선호 색상 수가 장치 지원 색상 수보다 작으면 사용자 선호 색상 수를 사용하라'는 내용을 담고 있다. 이는 사용자 선호 정보에 더 높은 우선 순위를 두고 속성 정합을 수행한 예이다.

후에 요구 사항의 변화로 인해 장치 지원 색상 수에 더 높은 우선 순위를 부여해야 한다면 상기한 규칙을 변경함으로써 손쉽게 요구 사항을 만족시킬 수 있다.

3.2.2.2. 변환 기능 결정

일련의 규칙에 의해 다양한 속성들이 정합되면 정합 결과를 바탕으로 필요한 변환 기능을 결정해야 한다. 다음은 이러한 기능을 수행하는 규칙의 예이다.

```
(defrule is-image-transcoding-needed
(UseNoOfColors ?no1)
(HasNoOfColors ?no2&(< ?no1 ?no2))
=>
(printout t "<Transcoder><name>ImgTrans</name>
<prop><name>colorNo</name><value>
?no1
"</value></prop></Transcoder> crlf" )
)
```

*HasNoOfColors*는 문서 프로파일에서 추출된 속성으로서 서비스 대상 이미지의 색상 수를 값으로 갖는다. 따라서 상기한 규칙은 '현재 허용된 색상 수가 이미지의 색상 수보다 작으면 이미지의 색상 수를 줄여야 하므로 이미지 변환기를 실행하라'는 의미를 갖는다.

후에 이미지 변환 기능을 수행하는 변환기의 사양이나 이름이 바뀌면 상기한 규칙을 변경함으로써 손쉽게 변화된 환경에 대응할 수 있다.

4. 결론

장치 독립성을 지원하기 위해서는 전송 컨텍스트를 해석할 수 있어야 한다. 본 논문에서는 CC/PP로 기술된 전송 컨텍스트를 사실(Fact) 목록으로 변환한 뒤 규칙 기반 시스템을 통해 해석하는 기법을 소개하였다.

본 논문에서 소개된 기법을 사용함으로써 복수개의 프로파일 사이에 존재하는 복잡한 논리적 상호 연관성을 효과적으로 기술할 수 있었다. 또한, 전송 환경 속성의 변경, 변환 기능의 확장 및 변경, 요구 사항의 변화 등에 유연하게 대처할 수 있다.

참고문헌

- [1] W3C Device Independence Activity (<http://www.w3.org/2001/di/>)
- [2] Roger Gimson et al, "Device Independence Principles", W3C Working Draft, 18 September 2001.
- [3] Roger Gimson et al, "Delivery Context Overview for Device Independence", W3C Working Draft, 13 December 2002.
- [4] Apache Cocoon (<http://xml.apache.org/cocoon/>)
- [5] IBM Web Sphere Transcoding Publisher (<http://www-3.ibm.com/software/webservers/transcoding/>)
- [6] Brience Mobile Processing Server (<http://www.brience.com/>)
- [7] Cysive Cymbio Interaction Server (http://www.cysive.com/solutions/device_support.html)
- [8] Graham Klyne et al, "Composite Capabilities/Preference Profile (CC/PP): Structure and Vocabularies", W3C Working Draft, 08 November 2002.
- [9] Minsu J., Jaehong K., Joo-chan S., "CATS - Content Adaptation & Transcoding System", Proceedings of International Conference on Advanced Communication Tech., pp322~pp326, January 2003.
- [10] JESS: Java Expert System Shell (<http://herzberg.ca.sandia.gov/jess/>)