

워크플로우 시스템에서 HTML 문서를 이용한 데이터 접근성의 조정

김홍숙, 한동수

한국정보통신대학원대학교 시스템 소프트웨어 연구실

Data accessibility control for Workflow System using HTML document

Hong-Soog Kim(kimk@icu.ac.kr), Dong-Soo Han(dshan@icu.ac.kr)
System Software Lab. of Information and Communications University

요 약

기존의 워크플로우 시스템의 단위 업무간의 흐름 조절이 제어의 흐름에 중점을 두고 연구되었으나 실제 단위 업무간에 흐름 조절 시에 데이터도 같이 흘러가므로 이에 대한 연구가 필요하다. 본 논문에서는 워크플로우 시스템에서 단위 업무간의 흐름 조절 시에 제어의 흐름에 따라 데이터에 대한 접근성을 HTML문서를 이용하여 표현하는 방법을 제안한다. 단위 업무간의 흐름 조절의 형태에 따른 데이터의 접근성에 대한 요구사항을 고찰하고 이를 제안된 데이터 접근성 표현 방법은 사용하여 구현하는 방법에 대하여 제시한다.

1. 서 론

1.1. 워크플로우 관리 시스템

회사의 업무 프로세스는 수많은 단위 업무와 담당자 그리고 다양한 정보들을 사용하여 여러 부서를 경유하며 이루어지고 때로는 다른 회사가 일련의 업무 프로세스에 참여하여 이루어질 수도 있다. 그러나 각 부서별로 이질적인 전신 자원과는 별개로 업무 프로세스는 정해진 목적에 따라 일정한 절차와 규칙에 따라 진행되는 일련의 관련된 단위 업무간의 제어와 데이터의 흐름이라고 볼 수 있다. 이러한 이기종 분산 환경하의 일련의 단위 업무들 사이를 연결하고 관리할 수 있는 즉, 업무 프로세스의 측면에서 여러 애플리케이션을 관리할 수 있는 시스템의 필요성이 대두되었다. 이러한 요구사항을 충족시키기 위하여 설계된 시스템이 워크플로우 관리 시스템 (workflow management system)이며, 워크플로우 관리 시스템은 프로세스의 규칙, 절차, 업무 담당자 및 애플리케이션 프로그램, 데이터 및 업무 조직 구조 등의 비즈니스 핵심 요소들을 통합시켜 적시에 정확한 데이터를 업무 담당자에게 제공하여 최선의 의사 결정을 내리도록 지원한다. 더 나아가 업무 프로세스의 사후 평가를 통하여 지속적인 향상과 혁신을 가능케하여 효율적인 업무 프로세스를 구축하게 해준다[1].

1.2. 관련 연구

워크플로우 표준의 개발과 보급을 위하여 1993년 8월에 결성된 비영리 표준화 단체인 WfMC(Workflow Management Coalition)는 다양한 표준 사양(specification)의 개발을 지원하기 위하여 용어를 정리한 "The WfMC Glossary"와 참조 모델인 "The Workflow Reference Model"을 제시하였다[2]. WfMC의 Working Group 1에서는 Workflow Reference Model에 기반하여 워크플로우 프로세스를 정의하는 언어인 "Workflow Process Definition Language WPDL" 표준안을 발표하였다. WPDL은 이기종 워크플로우 계층간의 워크플로우 프로세스 정의의 상호교환을 위한 중간표현(intermediate format)으로 사용하기 위하여 제안된 것이지만, 워크플로우 프로세스 정의를 위한 최소한의 구문을 정의하고 그 범위를 넘는 것은 각각의 제품 공급자들이 구현하도록 하였다[3].

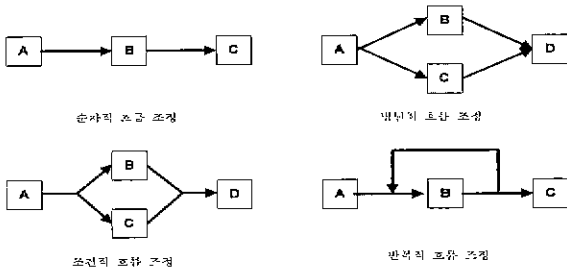
W.M.P. van de Aalst는 Petri net에 기반한 워크플로우의 모델

링과 분석에 대한 문제에서 기존의 Petri net을 워크플로우 모델링에 사용하기 위하여 token이 attribute를 가질 수 있는 colored Petri net, 시간 개념을 도입한 timed Petri net과 대규모의 워크플로우 모델링을 위한 hierarchy를 추가하여 high level Petri net에 대하여 설명하고 있다. high level petri net을 사용함으로써 기존의 Petri net이 가지는 시간과 데이터의 표현상의 계약을 극복하였으나, high level Petri net으로 워크플로우 관리 시스템에서 다루어야 하는 Organization Unit을 다룰 수가 없으므로 별도의 Organization Management System으로 처리해야 한다는 점을 지적하고 있다[4].

Alonso et al.은 고급 트랜잭션 모델(advanced transaction model)로서 워크플로우 프로세스를 정의하는 문제에 대하여 고찰하였다. 이들은 트랜잭션 개념이 내포하고 있는 ACID property가 회복(recovery)나 병행성(concurrency)문제를 처리하는데 있어서 이론적으로 잘 정리되어 있으나 현재의 워크플로우 시스템들이 회복에 대한 처리를 하지 않고 있음을 지적하고 있다. 또한 워크플로우 시스템에서 처리할 수 있는 애플리케이션의 관리 기능이나 선택적인 단위 업무의 라우팅 기능과 같은 부분은 트랜잭션의 개념으로 처리할 수가 없고 이를 위하여 ACID property를 약화하여 다양한 현실적인 환경에 적용하기 위한 고급 트랜잭션 모델이 이론적으로 제시되었으나, 상임적으로 구현되지 못하고 있음을 지적하고 있다. 또한 고급 트랜잭션 모델은 워크플로우 프로세스 모델로서 표현이 가능한 예시를 보이고 워크플로우 프로세스의 개념이 트랜잭션보다 광범위한 개념임을 확인하였다[5].

1.3. 기존 연구의 문제점

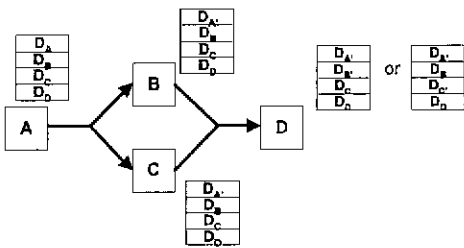
WfMC의 워크플로우 프로세스 정의 언어인 WPDL이나 Petri Net을 이용한 연구에서는 데이터와 제어의 흐름 중 제어의 흐름을 조정(routing)하는 것에 치우쳐 있다는 점이다. 즉 단위 업무간에 이동하는 대상은 제어뿐만 아니라 데이터도 이동이 되는데, 데이터의 이동에 대하여 WPDL에서는 단위업무의 정의시 입력 파라미터와 출력 파라미터로 단순히 표현하고 있다. Petri net에서는 단위업무에 주어지는 데이터를 단위업무(transition)의 input place에 주어지는 token의 형태로 모델링하고 있다. 두가지 모델에서의 흐름의 조정(routing)은 [그림 1]과 같은 기본적인 4가지로 구성되고 있다.



[그림1] 기본적인 흐름 조정 구조

단위 업무간의 흐름 조정에서 단위 업무간에 흐르는 데이터에 대하여 다음과 같은 점이 고려되어야 한다. 워크플로우 프로세스 내에서 단위업무 간에 이동이 되는 데이터는 처리되는 시점에서의 단위업무에 해당하는 데이터만이 접근 가능하여야 하고, 분기되는 시점에서는 전달될 데이터를 다음 단위업무에 맞게 접근성(accessibility)이 제한되어야 하고, 결합되는 시점에서 각각의 경로를 통해 변경된 데이터들간의 병합이 되도록 지원이 되어야 한다.

조건적 흐름 조정의 경우를 예를 들어 설명하면, 워크플로우 프로세스를 흐르는 데이터 전체를 D라 하고, 단위업무 A가 처리할 데이터를 D_A, 단위업무 B가 처리할 데이터 부분을 D_B, 단위업무 C가 처리해야 할 데이터 부분을 D_C, 단위업무 D가 처리해야 할 데이터 부분을 D_D라고 할 때 [그림 2]와 같은 형태로 데이터에 대한 접근이 이루어져야 한다.



[그림2] 조건적 흐름 조정시의 데이터 접근성

기존의 모델링에서는 제어의 흐름의 분기와 결합에 대하여 연구하였으나 제어의 흐름에 따른 데이터의 분할, 접근성의 조정, 병합에 대한 문제를 다루지 않고 있다. 이러한 문제점에 대하여 본 연구에서는 HTML문서를 데이터 전달의 매개체로 사용할 수 있는 웹 기반의 워크플로우 관리 시스템[6]에서 사용할 수 있는 접근성의 표현, 데이터의 분할, 데이터의 병합에 대한 해결책을 제시하고자 한다

2. HTML문서의 주석을 이용한 데이터 접근성의 표현

HTML문서에는 기본적으로 각각의 데이터 부분에 대한 접근성을 표현하는 방법이 없으므로 이에 대한 표현을 추가하여야 한다. 먼저 고려할 수 있는 방법은 접근성을 표현하는 새로운 태그를 추가하는 것이다. 그러나 이 방법은 HTML의 사양이 변경되어야 하고 기존의 브라우저에서는 해석이 불가능하다는 문제점이 있다. 그러므로 접근성을 표현하는 방법은 기존의 HTML문서의 사양에 대한 변경이 없도록 확장되어야 한다. 본 연구에서는 이러한 문제에 대한 해결책으로 HTML문서의 주석에 접근성에 대한 표현을 추가함으로써 기존의 HTML문서에 영향을 미치지 않도록 한다. 주석에 표현된 데이터에 대한 접근성은 기존의 웹 브라우저나 웹 서버의 해석에 있어서 무시되므로 웹 브라우저나 웹 서버의 변경이 필요 없다는 장점이 있다

데이터 접근성이 포함된 HTML문서는 워크플로우 시스템에서 접근성을 해당 데이터에 적용하여 표준 HTML문서로 변환함으로써 웹 시스템에 대한 투명성을 가진다.

2.1. 데이터 접근성의 표현

HTML문서에서의 주석은 다음과 같은 형태로 기술이 되고 이는 웹 브라우저에서는 무시된다.

```
<!--Comments here -->
```

주석부분에 다음과 같은 형태의 표현을 추가하여 접근성을 표현한다.

```
<!-- <<ACCESS_PERMIT =
      ALL "Activity_name", R|RW
      [, ALL "Activity_name", R|RW] >>
-->
```

ACCESS_PERMIT : 접근성을 표시하는 구문의 태그

ALL : 누구나 접근가능한 데이터 부분임을 표시

ACTIVITY_NAME : 데이터에 접근할 수 있는 단위 업무를 표시, 인용부호로 둘러싸인 문자열을 값으로 가짐

R | RW : "Activity_name"으로 주어진 단위업무가 데이터에 접근할 수 있는 모드를 지정

[그림2]에 대한 HTML 문서는 [그림 3]과 같이 형태로 표현이 가능하다

```
<HTML>
<!-- <<ACCESS_PERMIT = ALL R --> -->
<HEAD>
<TITLE>조건적 흐름 제어시 화면 예제</TITLE>
</HEAD>
<!-- <<ACCESS_PERMIT = ALL R --> -->
<BODY>
<FORM METHOD=POST
ACTION=/http://www.wf_server.tu.ac.kr/cgi-bin/condition_routing>
<!-- 단위업무 A 데이터 -->
<!-- <<ACCESS_PERMIT = A RW --> -->
<INPUT NAME="단위업무 A" SIZE=40>

<!-- 단위업무 B 데이터 -->
<!-- <<ACCESS_PERMIT = B RW --> -->
<INPUT NAME="단위업무 B" SIZE=40>

<!-- 단위업무 C 데이터 -->
<!-- <<ACCESS_PERMIT = C RW --> -->
<INPUT NAME="단위업무 C" SIZE=40>

<!-- 단위업무 D 데이터 -->
<!-- <<ACCESS_PERMIT = D RW --> -->
<INPUT NAME="단위업무 D" SIZE=40>

<!--
<INPUT TYPE="submit" <INPUT TYPE="reset"
</FORM>
</BODY>
</HTML>
```

[그림 3] 접근성이 표현된 HTML문서

2.2. 데이터 접근성이 표현된 HTML문서의 데이터 분할

접근성이 표현된 HTML문서에 표현된 데이터에 대한 접근성은 기존의 웹 브라우저에서는 의미가 없다. 그러므로 이러한 부분이 웹 브라우저에서 의미를 가지도록 접근성이 표현된 HTML문서를 표준 HTML문서로 변경하는 것은 워크플로우 관리 시스템에서 처리하는 것으로 가정한다. 이러한 변경작업은 CGI(Common Gateway Interface)등의 모듈을 워크플로우 시스템의 모듈로써 추가하는 것으로 가능하다. 워크플로우 관리 시스템에서는 접근성이 표현된 문서를 해석하여 접근성에 맞추도록 데이터 부분에 대한 변경을 취한다. 즉, 워크플로우 시스템은 현재 단위업무에 대한 정보를 이용하여 그 단위업무에서 사용되는 데이터에 대한 접근을 가능하게 하고 해당 단위업무가 접근할 수 없는 데이터는 접근을 무효화하여야 한다. 이를 위하여 HTML의 INPUT 태그를 동적으로 변경한다. 즉 현재 단위업무에서 접근가능하고 변경할 수 있는 데이터에 대하여 INPUT 태그를 활성화하고 그 이외에는 INPUT 태그를 비활성화 한다. [그림 2]와 같은 예에서 단위업무 B로 전달될 HTML문서는 단위업무 B가 접근할 부분만이 입력이 가능하고 다른 단위 업무에서 접근하는 데이터

부분은 입력이 불가능하도록 처리가 되어야 한다.

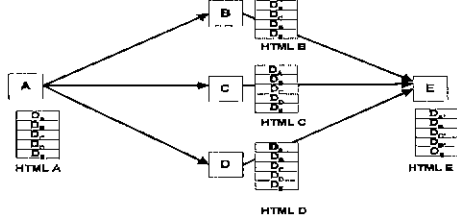
INPUT 태그의 활성화/비활성화는 주석 태그의 마지막 태그(-- >)의 위치를 변경하여 처리한다. 즉 비활성화 시에는 주석 태그의 마지막 태그를 INPUT태그의 뒤에 위치하고 데이터 있을 경우에는 주석 태그의 마지막 태그에 두도록 하여 처리된 데이터를 볼 수는 있지만 변경 불가능하도록 한다 이와 같이 주석 태그의 마지막 태그의 위치를 변경함으로써 원래의 접근성에 대한 표현 정보를 보존할 수가 있으므로 나중에 데이터에 대한 접근성을 활성화할 때 이용이 가능하다는 장점이 있다.

2.3. 데이터 접근성이 표현된 HTML문서들의 병합

접근성이 표현된 HTML문서들이 병합되는 시점은 두 가지의 형태를 가진다. 즉 조건적 흐름 조정에 의하여 분할된 데이터를 포함한 HTML문서의 병합과 병렬적 흐름 조정에 의하여 분할된 데이터를 포함한 HTML문서의 병합의 두 가지 경우가 있다.

조건적 흐름 조정에 의하여 분할된 데이터의 경우에는 제어의 흐름이 결합되는 시점에서 활성화된 경로는 분기의 시점에서 하나로 정해지므로 활성화된 경로상의 결합 직전의 단위업무부터 전달된 HTML문서를 그대로 사용하면 된다.

병렬적 흐름 조정에 의하여 분할된 데이터의 경우에는 제어의 흐름이 결합되는 시점에서 활성화된 경로는 분기의 시점에서 분기 가능한 모든 경로가 되므로, 결합 직전위치의 모든 단위업무에서 전달된 HTML문서를 병합하는 작업이 필요하다. 예를 들어 [그림 4]과 같은 경우에 단위업무 E에서는 직전의 단위업무 B,C,D에서 작업한 HTML B, HTML C, HTML D를 병합하여야 한다.



[그림 4] 병렬적 흐름 조정시의 결합 시점에서의 데이터 병합

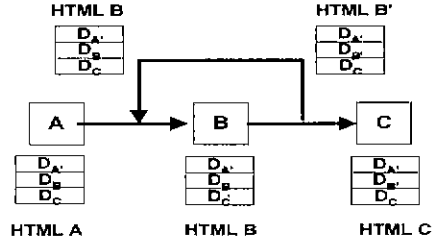
이 경우 단위업무 A에서 분기하는 병렬 업무 B,C,D에서 접근하는 데이터는 서로 독립적인 것으로 가정한다 선택적 흐름 조정의 경우에는 분기되는 경로가 배타적이므로 각 경로상에 단위업무가 접근하는 데이터가 서로 독립적일 필요는 없지만, 병렬적 흐름 조정 시에는 분기 가능한 모든 경로상에 있는 단위업무가 실현이 되므로 접근하는 데이터는 서로 배타적이지 않으면 서로 다른 경로상에 있는 단위업무에 의하여 동일한 데이터에 대한 변경이 발생 가능하므로 결합 시에 데이터의 병합에서 의미적인 혼란을 가져오게 되므로 병렬 흐름 조정 시에 서로 다른 경로에 있는 단위 업무들 간에 접근하는 데이터가 서로 배타적이 되도록 가경하는 것이 바람직하다.

결합 시점의 단위업무 E에게 전달되는 HTML문서는 워크플로우 시스템에 의하여 HTML B, HTML C, HTML D를 병합한 결과이다. 병합하는 과정은 HTML B, HTML C, HTML D에서 공통적으로 포함된 D_A'과 D_C를 기본으로 하여 HTML B에서 D_B', HTML C에서 D_C', HTML D에서 D_D'을 추가하여 생성된 HTML 문서를 단위업무 E에 대한 접근성을 해석하여 표준 HTML문서로 변환함으로써 HTML E가 생성이 된다

2.4. 반복적 흐름 조정시의 데이터 분할 및 병합

반복적 흐름 조정시의 데이터의 분할 및 병합은 선택적 흐름 조정의 특수한 경우로 생각할 수 있다. [그림 5] 과 같은 반복적 흐름 조정의 경우에 단위 업무 B는 단위 업무 C 또는 다시 B로 선택적으로 업무의 흐름이 조정 이 될 수 있다 단위업무 B로 업무 흐름이 조정 이 되는 경우에는 HTML B'에서 HTML B로 변환이 되어야 한다. 이 과정은 접근성이 표현된 HTML문서에서 표준 HTML문서로의

변환의 역 변환으로 볼 수 있다. 역 변환시에는 표준 HTML문서 변환 시에 보존된 데이터의 접근성에 대한 정보를 이용한다.



[그림 5] 반복적 업무 흐름 조정시의 데이터의 분할 및 병합

3. 결론 및 토론

기존의 워크플로우 시스템의 단위 업무간의 흐름 조정이 제어의 흐름에 중점을 두고 연구되었으나 실제 단위 업무간에 흐름 조정 시에 데이터도 같이 흘러가므로 이에 대한 연구가 필요함을 본 연구에서 지적하였다. 본 연구에서는 워크플로우 시스템에서 단위 업무간의 흐름 조정 시에 제어(control)의 흐름에 따라 데이터에 대한 접근성을 HTML문서를 이용하여 표현하는 방법을 제안하였다. 단위 업무간의 흐름 조정의 형태에 따른 데이터의 접근성과 분기와 결합시의 데이터 접근성에 대한 요구사항을 고찰하고 이를 제한된 데이터 접근성 표현 방법에 의하여 구현 시에 고려하여야 할 사항을 고찰하였다.

제한된 데이터 접근성의 표현 방법은 기존의 웹 브라우저나 웹 서버에 대한 추가적인 변경을 요구하지 않으므로 HTML문서와 HTTP 프로토콜을 사용하는 웹 기반의 워크플로우 시스템에 직접 적용이 가능하다는 장점을 가진다 추가적인 연구가 필요한 부분은 보안의 문제이다 현재의 데이터 접근성에 대한 표현은 악의의 사용자가 변경이 가능하다 즉 HTML문서의 데이터 접근성을 사용자가 변경하여 동작시킬 수가 있다는 점이다. 그러므로 워크플로우 시스템은 클라이언트로부터 웹 서버에 전달된 HTML문서에 대하여 불법적인 변경이 있었는지를 검사하는 방법이 요구된다 이러한 검사는 웹 브라우저에게 전달되는 HTML문서를 캐칭하고 있다가 응답(response)으로 온 문서를 비교하여 해당 단위업무가 접근할 수 있는 데이터 이외의 영역에 변화가 있었는지를 검사하는 방법을 생각할 수 있다

참고 문헌

- [1] 김동업, "IBM FlowMark : 전문 워크플로우 시스템", The Magazine of Sisa, pp. 19-26, 1997 겨울호
- [2] D Hollingworth, "The workflow reference model", Technical Report TC00-1003, Workflow Management Coalition, December 1994
- [3] WfMC Work Group 1, "Workflow Management Coalition Interface 1 : Process Definition Interchange", Document number TC-1016, WfMC. May 1996
- [4] WMP van der Aalst "The Application of Petri Nets to Workflow Management ", The Journal of Circuits, Systems and Computers, pp. ??, 1998.
- [5] G. Alonso, D Agrawal, A. El Abbadi, M Kamath, R. Gunthor and C. Mohan, "Advanced Transaction Models in Workflow Contexts", Proceedings of the Twelfth International Conference on Data Engineering (ICDE'96), New Orleans, Louisiana, USA, February 1996.
- [6] 한동수, 박향재, "비즈니스 프로세스의 자동화를 지원하는 웹기반 HiFlow시스템의 설계 및 구현", 정보과학회논문지(C), 제4권, 1호. pp.50-62, 1998.