
웹 페이지에서 참조무결성 제약조건에 관한 연구

유남현 · 손철수 · 김원중
순천대학교

A Study on Referential Integrity Constraint on Web Page

Nam-hyun, Yoo · Cheol-Soo · Won-Jung, Kim
Sunchon National University
E-mail : kwj@sunchon.ac.kr

본 논문은 정보통신부 정보통신연구진흥원에서 지원하고 있는 정보통신기초기술
연구지원사업의 연구결과입니다.

요 약

다수의 사람들이 관리·운영하는 많은 분량의 HTML 문서들로 이루어진 웹 사이트들의 경우에는 현수참조(Dangling Reference)와 오문참조(Inaccurate Contents Reference)와 같은 깨진 링크(Broken Link)를 관리하는데 많은 어려움이 따른다. 즉, HTML 페이지의 내용 중 연결구조에 관련된 부분을 변경하거나 삭제하는 경우 문서들간의 연결구조에 대한 무결성(Integrity)을 보장하기가 어렵다.

본 논문에서는 Parent 페이지와 Child 페이지에서의 무결성 제약조건(Integrity Constraint Condition)들을 조사하여 정의하고, 확장 UML로 표현하는 방법에 대해 연구하였다.

ABSTRACT

The Web site which many people manage and maintenance much HTML page is so difficult to solve broken link like as dangling reference and inaccurate contents reference. That is, in the case we update or delete connections structure between HTML documents, it is difficult to guarantee referential integrity of Web Pages.

In this chapter, we surveyed and defined integrity constraints condition between the parent page and child page, and studied about method to express by extension UML.

WWW, Constraints, UML, Integrity, Dangling Reference, Broken Link

1. 서 론

Tim Berners-Lee에 의해 개발된 WWW이 처음 소개되었을 때만 해도 그 누구도 WWW가 모든 컴퓨팅 환경의 기본이 될 것이라는 상상을 하지 못하였다. 그러나 WWW은 기존의 클라이언트 서버(C/S) 환경 기반의 모든 응용 프로그램들을 새로운 웹 환경으로 전이를 시켰으며, 그에 따른 관련된 개발 기술도 빠르게 발전하였다. 90년대 초창기 웹 환경에서는 단순하게 정보를 전달하기 위하여 HTML(Hyper Text Markup Language)만을 이용하여 단방향 정보 전달 수단으로 활용되었지만, 현재의 웹 환경은 HTML을 기반으로 사용자와 상호 대화 형태의 웹 환경을 구축하기 위한 서버 사이드 스크립트 언어, 플래시, 자바, Active-X 등을 사용하여 대규모로 구성되고 있다.

대규모로 구성된 웹 페이지들은 추가, 삭제, 갱

신 등이 끊임없이 일어나고 있으며, 계속되는 웹 페이지들의 수정으로 인하여 페이지 사이의 링크가 끊어지거나 고립된 페이지 등과 같은 문제들이 발생하게 된다. 이러한 링크나 페이지들은 발견하기가 매우 어려우며, 때에 따라서 기업에 큰 손실을 가져올 수 있다.

이에 본 논문에서는 웹 페이지들간의 링크가 끊어지거나 고립된 페이지들이 발생하지 않도록 하기 위하여 수 많은 웹페이지들간의 연결 상태를 분류하여 관계를 정의하였으며, 그 관계에 따른 무결성 제약 조건을 설정할 수 있도록 하였다.

II. 관련연구

2.1 참조 무결성

참조 무결성이라는 개념은 맨 처음 관계형 데이

터베이스에서 나타났다. 관계형 데이터베이스에서의 참조 무결성은 데이터의 정밀성(Accuracy)이나 정확성(Correctness)을 의미하는 것으로서 데이터 베이스에 저장된 데이터를 의미적으로 정확하게 유지하는 것을 말한다.

무결성 제약조건(Integrity Constraint Condition)은 데이터베이스에서 어떤 데이터 값이 허용되고 어떤 트랜잭션의 수행 결과가 인정되는지를 나타낸다. 예를 들어 나이는 1부터 150사이의 정보 값이어야 한다는 데이터 값에 대한 무결성 제약 조건이 정의될 수 있고, 나이를 감소시키는 트랜잭션의 결과는 데이터베이스에 반영될 수 없다는 것 등이다.

또한 외래키(Foreign Key) 제약 조건은 외래키에 의해 참조되는 튜플(Tuple)은 관계를 맺고 있는 릴레이션에 반드시 존재하여야 함을 명시한다. 따라서 무결성 제약 조건을 만족시키지 못하는 갱신(Update)은 거절(Reject)되거나 트리거 연산을 호출하여 만족시키도록 하여야 한다[2].

표 1. 데이터베이스에서의 참조 무결성 형태2

	Delete	Insert	Update
	Restrict	Restrict	Restrict
Parent	Cascade	Cascade	Cascade
	None	None	None
	Restrict	Restrict	Restrict
Child	Cascade	Cascade	Cascade
	None	None	None

데이터베이스에서 참조 무결성에 적용되는 Parent와 Child에서 발생할 수 있는 여러 제약조건들을 표 1에서 보여주고 있다. 데이터베이스에서 참조 무결성의 종류로는 크게 Parent와 Child로 구별될 수 있는데 이들은 각각 삽입(Insert), 삭제(Delete), 갱신(Update) 등이 일어날 수 있으며, 이들 각각은 다시 Restrict, Cascade, None과 같은 세부 제약조건을 부여함으로써 약 18가지의 유형으로 구별할 수 있다.

2.2 확장 UML

UML(Unified Modeling Language)은 소프트웨어 개발 방법론을 적용할 때 사용되는 하나의 표준 표현 도구로서 객체지향 개념을 기반으로 한 응용 프로그램 개발에 적합한 모델링 언어이다[3]. 이러한 UML을 웹 환경으로 확장한 것이 바로 확장 UML(Extended UML)이다. 확장 UML은 기존의 UML에서 일련의 스테레오 타입(Stereo Type), 꼬리표 값(Tagged Value), 제약(Constraint), 그리고 노트(Note)를 정의하여 웹 어플리케이션을 모델링할 때 사용할 수 있게 하였다. 스테레오 타입과 제

약은 확장 UML에서만 볼 수 있으며, 특정 컴포넌트에 적용할 수 있는 주요한 요소로는 웹 페이지이며 웹 페이지에는 여러 가지 스테레오타입을 적용할 수 있고, 다른 스테레오 타입들도 다른 HTML 요소에 적용하여 Frame, Target, Form과 같은 시스템의 주요 아키텍처 컴포넌트를 표현할 수 있다 [4].

III. 웹 페이지에서의 참조 무결성 제약 조건

웹 페이지에서의 참조 무결성 제약 조건은 관계형 데이터베이스에서 사용되는 제약 조건을 응용한 것으로서 관계형 데이터베이스에서 참조 무결성 제약 조건은 약 18가지 정도이나 웹 페이지에서는 그보다 훨씬 적은 5가지 정도로 정의할 수 있다. 표 2는 웹 페이지들간의 정의가 가능한 참조 무결성 조건으로서 모두 Delete에서만 정의가 되었다. 웹 페이지가 관계형 데이터베이스와 차이점이 생기는 것은 웹 페이지 자체로서 하나의 속성이나 레코드가 아닌 여러 개의 속성이나 레코드가 구성되어 하나의 개체를 이루기 때문이다.

표 2. 웹 사이트에서의 참조 무결성 조건3

	Delete	Insert	Update
	Restrict	-	-
Parent	Cascade	-	-
	None	-	-
	Restrict	-	-
Child	Cascade	-	-
	-	-	-

Parent와 Child 페이지 사이에서의 중복되는 관계가 발생할 경우 Child 페이지가 하위 Child 노드를 가지지 않는 최하위 노드일 경우는 Child 페이지 관점에서 제약조건을 기술하였다.

Parent 페이지에 관한 제약조건의 정의를 보면 PDR(Parent Delete Restrict)은 Child 페이지에 중요한 내용이 있는 경우로 Child 페이지가 존재할 경우 Parent 페이지의 링크를 삭제하는 것을 제한하여야 하며, PDC(Parent Delete Cascade)의 경우, Child 페이지가 독립적으로는 의미가 없어 Parent 페이지의 링크 삭제시 Child 페이지를 동시에 삭제하도록 한다. 또한, PDN(Parent Delete None)은 두 개 이상의 Parent가 하나의 Child 페이지를 가르키고 그중 관계를 추후에 삭제해도 문제가 없을 경우에 적용하도록 하였다.

Child 페이지에서 제약조건을 보면, CDR(Child Delete Restrict)이 가장 일반적인 경우로 Child 페이지를 지울 경우 Parent 페이지의 링크가 DLP가 발생하게 됨에 따라 참조되고 있는 Child 페이지의 삭제를 제한하며, CDC(Child Delete Cascade)는 Child 페이지가 삭제될 경우 Parent 페이지의

링크도 같이 삭제하도록 정의하였다.

3.1 PDR

PDR(Parent Delete Restrict)이란 Parent 페이지의 링크를 삭제하면 Child 페이지가 고아 페이지가 되는 경우로, Child 페이지만으로는 의미가 없는 제약조건을 말한다. 따라서 Child 페이지에 중요한 내용이 있는 경우, Child 페이지가 존재할 경우 Parent 페이지의 링크를 삭제하는 것을 제한하여야 한다는 조건을 말한다.

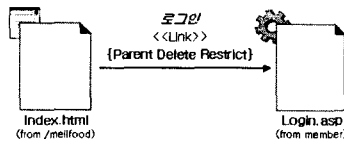


그림 6. 로그인 관련 사이트 링크에서 PDR 제약조건

M사 웹 사이트의 쇼핑몰 부분의 UML 웹 확장 구조도로서 로그인 관련 사이트와의 PDR 제약조건과 관련된 내용을 그림 1에서 보여주고 있다. Index.html은 Member 폴더의 Login.asp 와 <Link> 되어 있다. Index.html 페이지와 로그인 페이지 사이에는 자식노드가 중요하므로, Parent Delete Restrict 의 제약 조건을 줄으로써 부모노드의 삭제를 못하게 하는 제약조건을 주고 있다.

3.2 PDC

PDC(Parent Delete Cascade) 제약조건이란 Child 페이지가 독립적으로는 의미가 없어 Parent 페이지의 링크 삭제시 Child 페이지를 동시에 삭제하도록 하도록 하는 제약조건을 말한다.

관련사이트 링크에서 PDC 제약조건에 관하여 그림2에서 보여주고 있다. index.html은 Good_Site.html과 관련사이트라는 <Link> 명으로 연결되어 있으며, Parent Delete Cascade 라는 제약조건을 부여했다. 이는 Parent 페이지의 링크가 삭제될 경우, 독립적으로는 의미가 없는 Child 페이지가 동시에 삭제하도록 한 것이다.

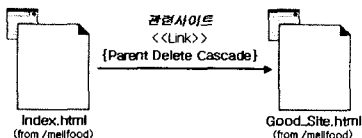


그림 7. 관련 사이트 링크에서 PDC 제약조건

3.3 PDN

PDN(Parent Delete None) 제약조건이란 PDN으로 관련된 제약으로서 두 개 이상의 Parent가 하나의 Child 페이지를 가르키고 그중 관계를 추후에 삭제해도 문제가 없을 경우에 적용하도록 하는

제약조건을 말한다.

도움말 관련 사이트 링크에서 PDN 제약조건에 관련된 내용을 그림 3에서 보여주고 있다.

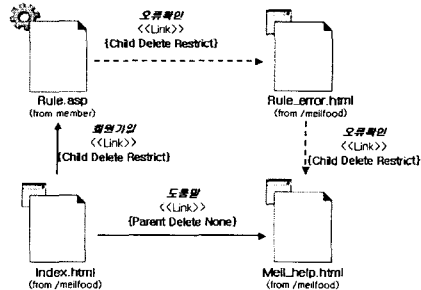


그림 8. 도움말 관련 사이트 링크에서 PDN 제약조건

index.html 페이지는 Meil_help.html 페이지와 '도움말'이라는 관계 식별명으로 PDN(Parent delete None)의 제약조건으로 <Link> 되어 있고, 또한 Rule.asp 페이지는 오류확인(Rule-error.html) 절차 과정을 거쳐 도움말 기능을 하는 Meil_help.html 페이지와 CDR(Child Delete Restrict) 제약조건으로 <Link> 되어 있다.

이와 같은 경우 오류확인이라는 링크 식별명으로 CDR의 제약조건으로 Rule_error.html과 <Link>로 연결되어 있는 Meil_help.html 페이지는 중요 오류와 관련되어 도움말을 제공하는 중요한 역할을 하므로 삭제를 해서는 안되지만, index.html 페이지에서 도움말이라는 이름으로 Meil_help.html로 링크된 페이지의 경우는 Parent를 삭제시켜도 의미상으로는 크게 변화가 없으므로 삭제를 하여도 되는 경우이다.

3.4 CDR

CDR(Child Delete Restrict) 제약조건이란 가장 일반적인 경우로 Child 페이지를 삭제할 경우 Parent 페이지의 링크가 깨지는 DLP가 발생하게 됨에 따라 참조되고 있는 Child 페이지의 삭제를 제한하도록 하는 제약조건을 말한다.

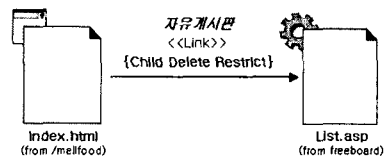


그림 9. 자유게시판 관련 사이트 링크에서 CDR 제약조건

자유게시판 관련 사이트 링크에서 CDR 제약조건을 지정한 내용을 그림 4에서 보여주고 있다. 자유게시판이라는 링크 식별명으로 Parent 페이지와 Child 페이지는 CDR(Child Delete Restrict) 제약조건으로 서로 링크 되어 있다.

Child 페이지 List.asp가 삭제될 경우, Parent 페

이지는 DLP가 발생되므로 웹 사이트의 무결성이 파괴되기 때문에 Child 페이지를 삭제할 수 없도록 제약조건을 부여하고 있다.

3.5 CDC

CDC(Child Delete Cascade) 제약조건이란 Child 페이지가 삭제될 경우 Parent 페이지의 링크도 같이 삭제하도록 한다.

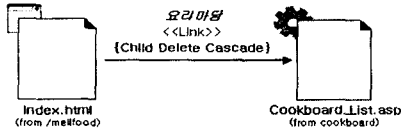


그림 10. 상품목록 관련 사이트 링크에서 CDC 제약조건

상품목록 관련 사이트 링크에서 CDC(Child Delete Cascade) 제약조건에 관한 사항을 그림 5.에서 보여주고 있다. 요리마당 이라는 링크 식별명으로 Index.html 페이지와 Shop_List.asp 가 CDC(Child Delete Cascade)라는 제약조건으로 <Link> 되어 있다. 이와 같은 경우는 Parent 페이지가 큰 의미를 부여하고 있지 않기 때문에 Child 페이지 Shop_List.asp가 삭제될 경우, Parent 페이지도 동시에 삭제가 이루어져야 한다는 제약조건에 관하여 나타내고 있다.

V. 결 론

본 연구에서 정의한 웹 페이지간의 관계 및 무결성 제약조건을 웹 사이트 개발 및 유지보수에 곧바로 적용할 수는 없다. 그러나 아무런 관계가 없

어 보이는 웹 페이지들 사이에서 Parent와 Child를 설정하고 그 사이에서 관계를 정의하면, 그 관계에 따른 제약조건들을 설정할 수 있다. 웹 페이지들간의 관계 및 제약조건을 정의하고 관리할 수 있는 응용프로그램이나 서버 프로그램을 개발하여 운용한다면 기존의 웹 개발자들이 항상 고민해 왔던 링크가 잘못된 페이지를 만들거나, 끊어진 페이지에서 발생할 수 있는 보안상의 취약점등을 미연에 방지할 수 있게 되어 웹 기반의 개발 환경의 개발 생산성을 향상시킬 수 있을 것이다.

참고문헌

- [1] Helen Ashman, Hugh Davis, "Missing the 4004: Link integrity on the World Wide Web", Seventh International World Wide Web Conference, 1998.8.
- [2] Stefano Ceri, G. Pelagatti, "Distributed Database Principles & Systems", McGraw-Hill, pp.61~62, 1984
- [3] Alan Dennis, Barbara Haley Wixon, David Tegarden, "Systems Analysis and Design With UML: An Object-Oriented Approach With UML", WILEY, pp.215~326, 2001.
- [4] Jim Conallen, "Building Web Applications with UML", 1999.
- [5] <http://www.w3c.org/>
- [6] 김원중, 이성재, "웹 사이트에서 하이퍼 링크의 무결성 유지를 위한 WPMS에 관한 연구", 한국정보처리학회추계학술논문집, 제10권 2호, pp.1089~1092, 2003.