도메인 모델을 이용한 온톨로지 모델로부터 시스템 모델 생성

남승환*, 임재현*, 김치수*
*공주대학교 컴퓨터공학과
e-mail:shnam@kongju.ac.kr

An System Model Construction from the Ontology Model Using the Domain Model

SwoongHwan Nam*, Jae-Hyun Lim*, Chi-Su Kim*
*Dept. of Computer Engineering, Kong-Ju National University

요 약

지식을 재사용하는 온톨로지 모델은 재사용 수준을 높여줄 수 있는 지식 모델이라 할 수 있다. 본 논문에서는 소프트웨어의 재사용 보다 지식을 재사용하기 위해 개발프로세스에서 지식과 소프트웨어 모델 사이에 관련성 있는 매핑을 만들고자 한다. 또한 UML을 온톨로지 모델 언어로 사용하여 UML 기반 온톨로지 모델로부터 시스템 모델을 추출하기위해 온톨로지 도메인 시스템 방법을 제안한다.

1. 서론

온톨로지와 문제 해결 방법은 재사용 컴포넌트로부터 지식 시스템을 구축하는데 있어서, 서로 보완적인 재사용 컴포넌트처럼 보여 질 수 있다. 또한 온톨로지를 사용함으로써 더 높은 상위단계의 재사용을 할 수 있다[1]. 이는 소프트웨어 재사용 대신 지식을 재사용하는 것이다. 현재, 지식기반시스템을 연구하는 모임에서는 온톨로지 표현과 획득에 도움이 되는 다양한 언어와 플랫폼을 연구하고 있다. 예를 들면 KIF기반 언어인 Ontolingua, KL_ONE기반 언어인 LOOM과 Protege 같은 지식 획득 도구들이 있다[6]. 또한 지식표현 파라다임이 강조된 다양한 언어와 도구들도 있다. 예를 들면 프레임기반, 서술로직, 객체지향 등이 있다. 그러나 이런 언어와 툴은 인공지능 연구자 이외에는 이용자가 적고, 온톨로지 모델과 지식기반시스템의 소프트웨어 모델을 결합하는 시스템적 방법이 부족하다[2]. 따라서 지식기반시스템의 설계를 위해 본 논문은 온톨로지 모델, 도메인 모델, 시스템 모델에 근거한 온톨로지 도메인 시스템(ODS) 방법을 제안한다. 그리고 예제 도메인으로서 재활용분리수거 시스템 기술 도메인을 사용한다. 재활용분리수거 시스템은 사용자가 입력한 자료를 기초로 재활용 가능한 선택방법을 제시함으로써 재활용 분리수거 방법의 결정을 제공하는 시스템이다[7].

2. 지식과 시스템 모델의 통합

본 논문에서의 지식 기반 시스템 구축은 2단계의 소프트웨어 개발 절차로 이루어진다. 첫 번째 단계는 도메인 지식과 사용자의 요구를[1] 결정하는 지식 중심 단계이고, 두 번째 단계는 사용자의 요구를 만족시키기 위해 컴퓨터 기반 해결책을 찾는 시스템 중심 단계이다.

사용자와 소프트웨어 개발자들이 온톨로지를 더 쉽게 이해할 수 있고, 더 많이 사용하게하려면 쉽게 이해할 수 있는 언어를 사용하여 표현해야만 한다. 온톨로지가 모든 지식을 포함할 경우 개발자들은 문제 해결을 위한 소프트웨어 시스템의 구축 시 온톨로지를 사용할 것이다[4].

이런 관점으로 볼 때, 온톨로지 모델은 주로 첫 번째 단계와 관련된다. 반면에 소프트웨어 모델은 두 단계 모두 관련되지만, 두 번째 단계를 더 강조하고 있다. 현재 온톨로지 모델은 좋은 소프트웨어를 생성하지 못하고 있다. 그 이유는 소프트웨어를 중심으로한 지원이 부족하기 때문이다. 그래서 온톨로지 모델과 지식 기반 시스템 사이에서의 모델링 시 발생하는 갭을 극복할 수 있는 연결 고리가 필요하다.

재활용분리수거 시스템 도메인을 대상으로 온톨로지를 생성 할 경우, 관련된 지식은 분류 계층 형태로 분석되고, E-R 다이어그램을 사용하여 명확히 표현된다[2]. 명확하게 표현된 결과는 온톨로지 에디터를 사용하여 구현 온톨로지로 변환된다.

Protege는 명확한 도메인 모델을 생성하고 유지하는데 소프트웨어 개발자들에게 도움을 준다. 그러나 protege는 제약 사항과 관계에 대해서는 명확하게 표현하기 어렵다. 따라서 protege는 높은 수준의 재사용을 필요로 하는 온

이 논문은 2006년도 정부(과학기술부)의 재원으로 한국과학재단의 지원을 받아 수행된 연구임(NO. R01-2006-000-10555-0)

톨로지 기반 소프트웨어 개발 프로세스에는 적절치 못하다. 또한 protege를 사용하여 구현한 온톨로지는 시스템 모델로 변경하기가 매우 어렵고 개발자기 이해하기가 어렵다. 따라서 본 논문에서는 이러한 단점을 보완한 방법을 제안한다.

3. 온톨로지 도메인 시스템 방법

이장에서는 지식기반 시스템 개발을 위한 온톨로지 도메인 시스템(ODS)을 제안한다. [그림 1]에서 보는 것처럼 ODS 방법은 온톨로지 모델과 시스템 모델 사이의 갭을 해결하기위한 방법으로써 도메인 모델을 사용한다[7]. 여기서 온톨로지 모델은 특정한 도메인에 관한 전체적인 지식과 관련이 있고, 도메인 모델은 특정한 도메인을 대상으로 시스템을 구축할 때 관련된 정보를 기술한다. 시스템 모델은 특정한 시스템을 위한 도메인 모델을 구체화한다.

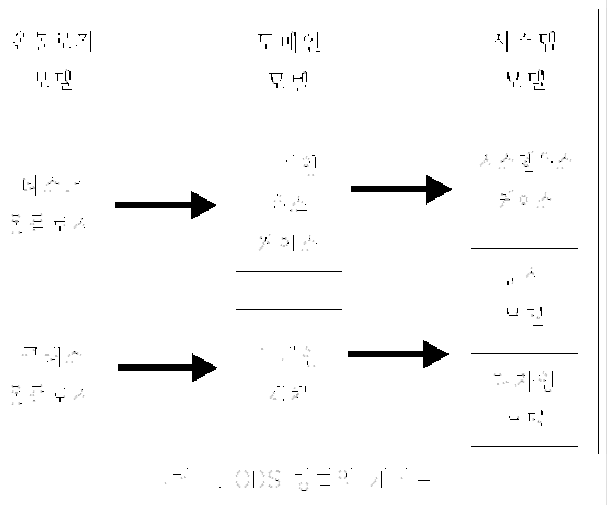
온톨로지는 도메인 지식, 데이터베이스 생성을 위한 기본적 지식획득 도구로 사용된다. 지식획득 과정 시 지식분석을 위해 추론모델기술(IMT)을 이용한다. 추론모델기술은 지식추출, 지식분석, 지식표현의 3개의 과정으로 이루어진다[7].

온톨로지 모델을 정의하기 위하여 지식추출 단계는 다음과 같은 과정을 거친다.

1. 도메인에서 물리적 객체 정의
2. 1에서 확인된 개체의 특성 정의
3. 2에서 확인된 특성의 값을 정의
4. 함수 또는 기능으로 특성 정의
5. 1과 2에서 구별된 특성, 객체와 연관된 관계 정의
6. 강요소와 중요성 측면 고려하여 5에서 구별된 관계의 부분순서 정의
7. 1과 2에서 구별된 특성과 객체로부터 파생된 추론관계 정의
8. 7에서 구별된 추론관계의부분순서 정의
9. 문제의 타스크를 정의
10. 9에서 구별된 타스크를 추론관계로 분해
11. 10에서 구별된 추론의 부분순서 정의
12. 도메인에서 전략적 지식을 정의
13. 12에서 구별된 어떻게 전략적 지식이 타스크와 관계가 있는지 정의
14. 지식형태의 명세서가 전문가와 지식 엔지니어가 만족할 때까지 1부터 반복

이 과정을 거친 지식은 타스크온톨로지와 클래스온톨로지를 세분화 되어 도메인 모델을 위한 자료로 활용한다.

온톨로지 모델은 도메인과 관계된 지식에 대해 특정한 어플리케이션에 비의존적이고 독립적인 모델이라 할 수 있다[1]. 온톨로지 모델은 [그림 1]에서처럼 타스크 온톨로지 모델과 클래스 온톨로지 모델로 구성된다. 타스크 온톨로지 모델은 특정한 도메인에서 타스크와 관련된 문제를 풀기위해 사용된 단어의 의미적 어휘를 제공 한다. 클래스 온톨로지 모델은 특정한 도메인에서 클래스, 속성, 관계에 대해 기술한다. 도메인 모델은 온톨로지와 관계된 부분을 선택하고, 시스템을 위한 요구사항들과 조합되도록 결정하는 어플리케이션 의존적 모델이다.



도메인 모델은 도메인 유스케이스 모델과 도메인 객체 모델로 구성된다. 도메인 유스케이스 모델은 추상적인 시스템의 형태를 구체적으로 표현한다. 도메인 객체 모델은 클래스와 그들의 관계에 관한 정보를 포함하는 어플리케이션의 정적 구조를 나타낸다. 여기서 도메인 유스케이스 모델과 도메인 객체 모델은 타스크 온톨로지 모델과 클래스 온톨로지 모델로부터 각각 파생 된다.

시스템 모델은 생성된 시스템의 어플리케이션 의존적 모델이다. 이 시스템 모델은 시스템 유스케이스 모델, 분석 모델, 디자인 모델로 구성된다. 따라서 전체의 어플리케이션을 여러 관점으로 설명할 수 있다. 이 시스템 모델은 도메인모델로부터 생성된다.

3.1 UML-OCL을 사용한 온톨로지 모델 특성

온톨로지 모델은 UML과 OCL를 결합하여 사용한다[4]. UML은 소프트웨어를 중심으로한 시스템의 가공물로서 시각화, 특성화, 생성과 문서로 사용된다.

UML 다이어그램은 명세서의 모든 관련된 측면들을 표현하기에는 충분히 섬세하지 못하다. 모델링시 객체에 대한 제약사항을 표현하는 형식적 언어가 필요하다. 이 형식언어가 발전하여 OCL언어가 탄생하였고 이는 비즈니스 언어로 발전되었고 OCL로 표현은 부작용을 발생시키지 않는다. OCL은 제약사항을 명확하게 표현하기 위해 사용되고 다른 표현은 UML 모델을 따른다. OCL은 첫 번째 순서 논리 표현을 지닌 형식 언어이지만 읽고 쓰기 쉽다[3]. UML-OCL로 표현된 모델은 다른 사람이나 도구에 의하여 모호하지 않게 해석될 수 있다.

재활용분리수거 시스템에서에서 분리방법의 도메인을 위한 온톨로지가 그림 2에 UML-OCL로 표현되어있다.

온톨로지의 기본요소는 클래스들, 관계, 제약과 프로세스

들이다. 그림 2에서 온톨로지 개념은 클래스 표기를 이용해서 표시하고, 클래스들간의 관계는 클래스 다이어그램을 표시하고, 타스크와 도메인 프로세스는 액티비티 뷰 또는 작업 뷰로부터 엑티비티 다이어그램을 사용해서 표시하고, 제약은 OCL을 이용해 표현한다.

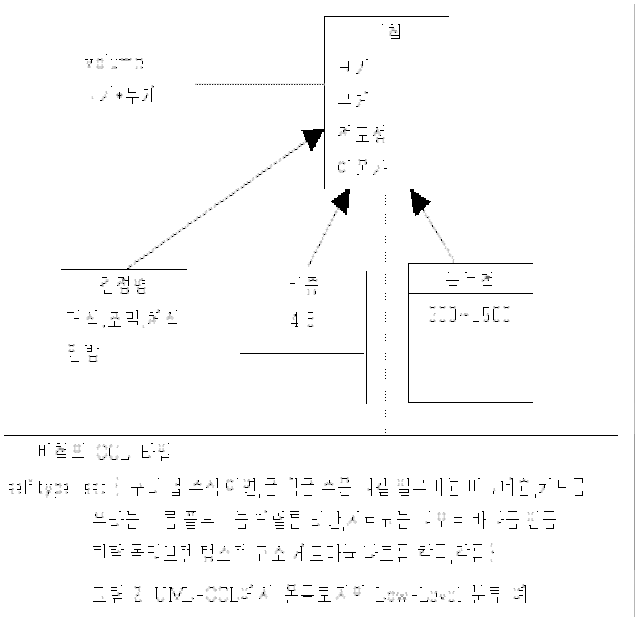


그림 2 UML-OCL을 이용 온톨로지의 Low-Level 뷰의 예

예를 들어 재활용분리수거 도메인에서 온톨로지를 개발하기 위하여 결정형, 비중, 녹는점 등의 다른 미디어에 대한 지식을 얻어 표시한다. [그림 2]는 재활용분리수거 도메인을 위한 클래스 온톨로지의 일부이며, 비철이라 불리는 클래스의 하위 레벨의 분류 다이어그램이다. 클래스들은 결정형, 비중, 녹는점, 이온가를 포함한 클래스 표기로 표현된다. [그림 2]에서 비철 클래스를 위한 OCL 표현은 비철의 모든 가능한 타입으로 표현하였다[7]. 한 가지 제약사항은 비철 클래스에 점선으로 접속되어있는데 이는 비철 클래스의 볼륨을 계산하기위한 식을 제공한다.

3.2 온톨로지 모델을 도메인 모델로 매핑

도메인 모델의 생성은 온톨로지 모델과 시스템 모델을 정의하는 중간단계이다. 지식기반 시스템의 어플리케이션을 개발할 때 지식의 일부는 온톨로지에서 사용된다. 따라서 특정한 도메인을 위한 온톨로지 기술은 매우 중요하다. 왜냐하면 온톨로지는 쉽게 적용되지 않고, 재사용 시 이들은 도메인의 일반적 특성으로 인해 부하나 부작용을 발생시킬 수 있기 때문이다.

도메인 모델을 위한 지식 표현은 객체, 속성, 값, 타스크, 객체와의 관계로 구별할 수 있다. 재활용부리수거 시스템에서 각각의 요소 표현 시 아래의 과정을 거쳐 명확해진다.

① 객체(O) : 비철 상태의 주위환경
② 속성(A) : 크기, 무게, 전도성, 이온가
③ 값(V) : 볼륨
④ 타스크(T) : 결정형, 비중 , 녹는점
⑤ 관계(R) : 결정형은 비철의 부클래스이다

위의 과정을 거친 표현은 클래스 프로세스 관계의 과정을 거쳐 관련된 지식들은 온톨로지 구조를 위한계층 클래스로 이용된다.

도메인 모델은 새로운 시스템을 위해 요구된 온톨로지로부터 관련된 지식을 조합하고 구성한다. 도메인 모델에서 표기는 표준 UML 표기를 따른다[3].

도메인 모델을 생성하기 위해 도메인 유스케이스 모델과 도메인 객체 모델은 타스크 온톨로지 모델과 클래스 온톨로지 모델로부터 파생된다.

첫째 타스크 온톨로지는 도메인 유스케이스 모델로 매핑된다. 도메인 유스케이스 모델은 시스템의 목적과 그 목적을 달성하기위한 단계들로 구성된다. 따라서 도메인 유스케이스와 도메인 엑터 두 요소를 가지고 있다. 도메인 유스케이스는 도메인에서 작업흐름을 표시하고, 유스케이스는 시스템이 무엇을 하는지 설명하고 방법은 정의하지 않는다. 유스케이스 작업흐름은 엑티비티 다이어그램, 시퀀스 다이어그램, 협동적 다이어그램, 상태차트 다이어그램을 이용하여 세세하게 묘사할수 있다[1]. 만약 타스크 온톨로지가 UML-OCL을 사용한다면 프로세스는 엑티비티 다이어그램으로 표현한다. 따라서 그들의 대부분은 도메인 유스케이스 모델로 직접 전이될 수 있다.

도메인 유스케이스 모델에서 도메인 엑터는 도메인에서 상호 작용하는 외부 엔티티를 정의 한다. 도메인 엑터는 유스케이스를 검사함으로써 확인된다. 일반적으로 엑터는 사람, 하드웨어 장치, 시스템에서 작동할 수 있는 다른 시스템의 역할을 표시한다[1].

두 번째, 클래스 온톨로지 모델은 도메인 객체모델로 매핑 된다. 도메인 객체모델은 어플리케이션의 정적 구조를 묘사하고 클래스와 관계에 대한 정보를 포함한다. 재활용분리수거 시스템 모델은 도메인 유스케이스 모델에서 케이스 사용에 따라 온톨로지 모델의 클래스로부터 생성된다. 도메인 유스케이스 모델에서 클래스와 객체는 클래스 온톨로지 모델의 하위 레벨 분류에 의해 생성되는데 이는 클래스가 인스턴스의 어디에 관계되는가하는 레벨이다.

3.3 도메인 모델을 시스템 모델로 매핑

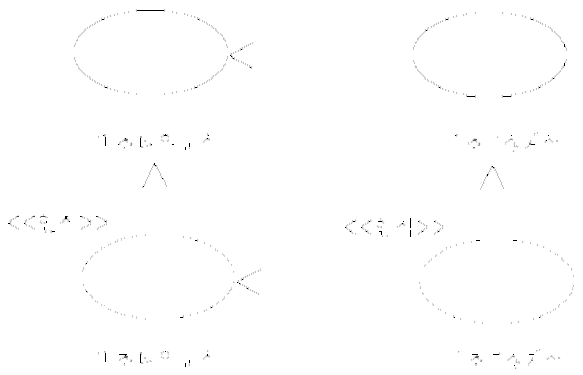
시스템 모델은 시스템 유스케이스 모델, 분석모델, 디자인 모델 세부분으로 구성된다. 온톨로지와 도메인 모델에서 기능적 요구와 시스템의 정적 객체 구조에 대한 정보는 분리되지만, 시스템 모델에서는 다시 결합된다. 시스템 유스케이스 모델은 시스템이 사용자와 전문가 사이에 어떤 관계가 있는지 설명하고, 엑터와 시스템사이의 상호작용을 포함하는 시스템의 행위을 나타낸다[1].

도메인 유스케이스 모델은 시스템 유스케이스 모델의 첫 번째 프로토타입 모형으로 사용된다.

분석모델은 유스케이스의 인식을 설명하고, 디자인모델의

추상개념을 제공하는 객체 모델이다. 분석 모델의 목적은 요구된 행동의 매핑을 제공하는 일이다. 세부적인 사항을 추가함으로서 분석 모델은 디자인모델로 변경되고, 분석 클래스는 직접 디자인 모델로 변경된다. 정확히 분석 모델은 사용자 인터페이스 정보와 데이터 구조로 결합된 유스케이스 실현을 나타낸다. [그림3]의 유스케이스 실현은 디자인 뷰에서의 유스케이스를 나타낸다.

[그림3]에서 "비철타입분석"에서 전체적으로 점으로 표시된 타원이 유스케이스 실현이다. 이는 디자인 레벨로부터 "비철타입분석"이라는 유스케이스를 인식한다.



유스케이스에 의존적인(스테레오타입 <<인식>>) 디자인 모델에서 유스케이스 실현은 유스케이스 모델에서 각각의 유스케이스를 위해서이다. 각각에 대한 유스케이스 실현을 위해 하나 또는 그 이상의 상호적 다이어그램은 객체 묘사와 그들의 상호작용을 이용해야한다[7].

4. 결론

본 논문에서는 일반 소프트웨어 개발자가 쉽게 이해할 수 있도록 UML을 온톨로지 모델 언어로 사용하여 UML 기반 온톨로지 모델로부터 시스템 모델을 추출하기위해 온톨로지 도메인 시스템 방법을 제안하였다. 이 방법을 통하여 도메인 모델은 온톨로지 모델과 시스템 모델사이에서 생기는 갭의 문제를 해결 하였다. 여기서 제시된 방법은 지식기반 시스템 뿐만 아니라, 복합지식모델 시스템에서도 적용할 수 있다.

향후 과제로는 UML-OCL 표현방법이 온톨로지 표현 형식론에 동등한 표현력을 가지는지 연구의 진행이 필요할 것이다.

참고문헌

[1] 채홍석, 객체지향CBD 개발 Bible, 한빛미디어, 2005

[2] 최은만, 객체지향 소프트웨어공학, 사이택미디어, 2006

[3] 임춘봉, 신인철, UML 입문, 홍릉과학출판사, 1999

[4] 이윤수, 정인정, UML 및 OCL을 이용한 서비스 온톨로지 설계 방안에 관한연구, 제23회 한국정보처리학회 춘계학술대회 논문집 제12권 제1호(2005)

[5] OCL : http://www.uml.org/

[6] Protege 2000 : http://protege.stanford.edu/

[7] Chen, Ontology Construction from Knowledge Acquisition system, PKAW 2000.