

관계 유형에 기반한 온톨로지 모듈 평가 메트릭

Evaluation Metrics for Ontology Modules Based on the Relationship Type

오선주(Sunju Oh)*

초 록

온톨로지 모듈화에 대한 필요성이 증가하고 있으며 이에 따라 다양한 방법으로 온톨로지를 모듈화하려는 시도들이 이루어져 왔다. 그러나 모듈화 과정을 통하여 생성된 온톨로지 모듈의 품질과 모듈화 일치성을 평가할 수 있는 공통된 기준인 평가 메트릭에 대한 연구가 부족한 실정이다. 본 연구에서는 소프트웨어 메트릭 이론을 바탕으로 하여 온톨로지에 적합한 모듈 응집도와 결합도 메트릭을 제안한다. 새로이 제안하는 메트릭은 기존의 온톨로지 메트릭과 상이하게 온톨로지 모듈을 적용 대상으로 하며 모듈의 클래스가 다른 클래스와 형성하는 관계의 유형을 세분화시킴으로써 관계 유형에 따른 영향도를 각각 반영할 수 있도록 하였다. 또한 제안한 메트릭의 검증을 위하여 메트릭 검증 프레임워크를 이용하여 이론적으로 타당성을 검증하고 웹 상의 온톨로지를 모듈화하여 적용하는 실증적 실험을 하였다. 본 연구에서 제안하는 온톨로지 모듈 응집도와 결합도 메트릭은 온톨로지 엔지니어들이 온톨로지 모듈과 모듈화 기법을 선택하여 사용하는 기준으로 활용할 수 있을 것이다.

ABSTRACT

In response to an increased need, various methods for ontology modularization have been proposed. However, few studies have focused on evaluative methods for ontology modules. In this study, we devise novel metrics to measure ontology modularity. To evaluate the ontology modules, we introduce cohesion and coupling based on the theory of software metrics. A cohesion metric and two coupling metrics were used to measure cohesion and coupling for ontology modules. These metrics were also used to check consistency between the ontology modules and the original ontology. The new metrics support a more detailed relationship between classes in ontology modules. We validate the proposed metrics using the well known verification framework and perform the empirical experiments to complement previous investigations. This study offers ontology engineers valuable criteria with which to select and use ontology modules and modularization techniques.

키워드 : 온톨로지, 모듈, 모듈화, 메트릭, 결합도, 응집도

Ontology, Module, Modularization, Metrics, Coupling, Cohesion

* 서울대학교 컴퓨터공학과 박사 후 연구

2009년 11월 25일 접수, 2010년 04월 07일 심사완료 후 2010년 04월 12일 게재확정.

1. 서 론

온톨로지는 도메인 내 정보 공유와 정보 통합을 위한 핵심 기반 기술로서 그 필요성이 인식되고 점차 사용이 증가하고 있다. 다양한 분야에서 온톨로지들이 개발되어 사용되고 있다. 그러나 하나의 획일적이고 크기가 큰 온톨로지는 재사용, 유지 관리, 확장성 측면에서 활용에 어려운 점이 많아 점차 소규모의 모듈화된 온톨로지의 사용에 대한 관심이 증가하고 있다[6]. 예를 들면 병원간 진료 의무 기록 통합 전산화를 위해 필수라 할 수 있는 의료 온톨로지인 SNOMED-CT, UMLSC 등의 의료 온톨로지는 업계의 사실상 표준으로, 앞으로 의료 업계에서 사용될 예정이다. 그러나 이들 온톨로지들은 OWL, RDF와 같은 온톨로지 언어가 아닌 특수한 언어를 사용하여 제작되었으며, 오랜 기간 동안 축적된 지식의 축적물로 크기가 상당히 방대하다. 그러나 워낙 크기가 대응량이기 때문에 Protégé와 같이 일반적으로 사용되고 있는 온톨로지 에디터에서 조작 사용이 어려운 실정이다.

대규모 온톨로지를 소규모 단위로 분할하여 사용하면 여러 그룹의 사용자들끼리 자료를 공유하고 협력하는 데 용이할 것이다. 그러므로 대규모의 온톨로지로부터 필요한 일부분을 추출하여 모듈화 하는 기법의 실용화가 무엇보다도 요구된다고 할 수 있다.

또한 온톨로지를 처음부터 새로 구축하는 것보다 품질이 보증된 온톨로지 모듈을 도입하여 구축함으로써 새로이 구축하는 온톨로지의 품질을 보장 받을 수 있고 처음부터 새로 구축하는 것보다 비용측면에서 더욱 절감되며 유지 보수와 협업을 위한 공동 표준화를

위해서도 모듈의 재사용은 도움을 줄 수 있을 것이다.

현재까지 온톨로지를 모듈화하기 위한 다양한 기법들이 연구되어 왔으며 온톨로지 모듈화에 대한 필요성이 여러 연구에서 이미 언급되어 온 바 있다[4-7, 11, 14]. 그러나 모듈화 작업의 결과물인 온톨로지 모듈의 품질과 모듈화 방법에 대한 평가 기준과 평가 방법에 대한 연구는 미흡한 실정이다.

그러므로 본 연구에서는 온톨로지 모듈에 대한 모듈 평가 메트릭인 응집도 메트릭과 결합도 메트릭을 제시하고 이론적 실증적 타당성을 보이고자 한다. 제안하는 온톨로지 모듈 평가 메트릭은 응용, 도메인, 일반, 표현 온톨로지 등의 분류와 관계없이 OWL로 작성된 온톨로지를 대상으로 한다.

또한, 실증 연구를 위하여 현재까지 연구 발표된 대표적인 온톨로지 모듈화 방법인 SW OOP[4], PATO[15], Prompt[11]를 이용하여 온톨로지 모듈을 구하고 각각의 모듈을 대상으로 새로이 제안한 메트릭들을 적용해 본다. 그리고 전문가들이 온톨로지 모듈을 평가하게 한 뒤, 새로 제안하는 온톨로지 모듈 메트릭을 적용한 결과와 비교해 보고 메트릭의 효과를 입증한다.

본 연구의 구성은 다음과 같다. 제 2장에서는 소프트웨어 메트릭과 온톨로지의 평가 메트릭에 관련된 선행 연구들을 살펴보았으며 제 3장에서는 온톨로지 모듈의 평가를 위한 평가 메트릭을 제시하며, 제 4장에서는 본 연구에서 제안한 메트릭의 이론적 실증적 검증을 수행한다. 제 5장에서는 연구 성과에 대해 토의하며 향후 연구 방향에 대해 논의한다.

2. 관련 연구

2.1 온톨로지 모듈화에 대한 연구

온톨로지 모듈화에 대한 연구들은 모듈화에 적용되는 기법의 공통적 특성을 기준으로 다음과 같이 분류되어 왔다. Seidenberg and Rector [14]는 모듈화 방법을 분할법, 추출법, 질의 기반 방식 등 세 가지 유형으로 정의하였다.

분할법에 대한 연구들은 Stuckenschmidt [15]와 Cuenca Grau[4]를 중심으로 이루어져 왔다. 이 방법은 온톨로지의 의미적인 측면 보다는 내부적인 구조로 온톨로지를 분할하였다. Cuenca Grau[4]은 모듈은 주어진 주제(topic)에 대해 자기 충족적으로 정의되어야 하며 (self-contained) 주어진 조건을 정확하고도 충분히 만족하는 공리의 최소 부분집합(minimal subset)이라고 정의하였다. 한편 Stuckenschmidt and Klein[15]은 모듈화를 위한 온톨로지의 구조적 측면에 대하여 강조하고 모듈의 요건을 다음과 같이 정의하였다. 첫째, 연결성(Connectedness), 둘째, 모듈의 크기 및 수, 셋째, 표현의 중복성(Redundancy of Representation)을 모듈의 중요한 요건으로 기술하였다.

한편, 추출법은 Noy[11], Seidenberg[14], D'Aquin[5], Doran[7]을 중심으로 연구되어 왔으며 이들은 온톨로지를 그래프 구조로 간주하고 그래프를 순회하는 방식으로 노드와 노드 사이의 관계를 따라 관련 노드들을 추출하였다. Noy[11]의 방식은 사용자와 상호작용을 통해 탐색의 시작 노드와 관련 속성(property), 탐색의 깊이(depth)를 사용자로부터 입력 받고 탐색 시작 노드에서 시작하여

지정된 탐색의 깊이에 도달할 때까지 클래스의 속성을 따라 반복적으로 계층 구조를 순회한다. 순회가 진행되는 과정에 사용자의 의도에 따라 특정 속성을 배제시킬 수 있다.

또한 Doran[7]은 모듈의 요건에 대해 다음과 같이 정의하고 있다. 첫째, 온톨로지 모듈은 자기 충족적(Self-contained)이어야 하며, 둘째, 개념 중심적(Concept centered)이어야 한다. 즉, 온톨로지 모듈은 시작 개념에 대해 충분한 정보를 포함하여야 한다. 셋째, 일치성(Consistency)이 만족되어야 한다. 즉, 온톨로지와 추출된 온톨로지 모듈은 서로 논리적으로 일치하여야 한다.

마지막 유형은 질의 기반 방법이다. 많은 연구자들이 데이터베이스의 필드로부터 영감을 얻어 SQL과 비슷한 온톨로지 질의를 정의하는데 질의의 결과가 하부 온톨로지 세그먼트(sub-ontology segment)를 형성한다. 질의 기반 방식은 SQL결과와 유사한 뷰(View)를 제공하는데, 이것은 매우 익숙한 형태의 인터페이스로 인식되지만 의미적(semantics) 측면은 매우 취약하여 뷰가 수정되는 경우 원본 온톨로지를 처리하는 방법에 대해서도 구체적으로 정의가 되어있지 않다. 질의 기반 방식은 소수의 개념만을 포함하는 적은 규모의 온톨로지나 단일 사용자 환경에서 적합한 방식이다.

2.2 메트릭에 대한 연구

2.2.1 소프트웨어 메트릭

소프트웨어의 품질을 평가하는 메트릭에 대하여 활발히 연구가 이루어졌다. 소프트웨

어 메트릭 중 대표적인 것으로 소프트웨어 복잡도 메트릭이 사용되어 왔는데 복잡도 메트릭은 소프트웨어의 이해도, 유지보수, 신뢰도, 테스트 능력에 영향을 주며 이와 같은 능력을 향상시키기 위해 소프트웨어 복잡도를 낮추는 것이 필요하다[10]. 소프트웨어 복잡도를 낮추기 위한 주요한 방법 중 하나는 모듈화를 하는 것이다. 이를 위하여 Page-Jones [13]는 분할과 정복(Divide and conquer) 원리를 이용하여 소프트웨어가 적당한 크기로 분할되고 독립적 기능을 가지는 것이 소프트웨어 재사용성(reusability), 이해도(understandability), 유지보수 용이성(maintainability), 확장성(modifiability)을 위해 필요하다고 정의하였다.

한편, Constantine and Yourdon[3]은 소프트웨어 모듈화는 결합도와 응집도에 의해 측정될 수 있다고 하였다. 소프트웨어 모듈화는 소프트웨어의 유지보수와 생산성을 높일 수 있는 수단으로 인식되어왔으며, 특히 객체 지향 개념은 추상화, 은닉화, 독립성을 제공함으로써 모듈화의 장점을 체계화하였다. 객체 지향 개념에 기반한 객체지향 메트릭에 대한 연구도 활발하게 이루어졌는데, Chidamber and Kemerer[2]는 객체 지향 설계를 위한 메트릭을 제안하였다. 그들은 객체지향 개념에 기반하여 Weighted Methods per Class(WMC), Response for a Class(RFC), Lack of cohesion(LCOM), Coupling between Object Class(CBO), Depth of Inheritance(DIT), Number of Children(NOC) 등 6가지 지표를 소프트웨어의 품질을 측정하는 주요한 메트릭으로 정의하였다. 그 중 LCOM은 응집도를 측정하고 CBO와 RFC는 결합도를 측정하는 목적으로

제안되었다. 객체 지향 소프트웨어에서 클래스는 인스턴스와 메소드로 구성되는데 LCOM은 메소드에서 공유되지 않는 인스턴스 수와 공유되는 인스턴스의 수의 차로 정의하였다. LCOM의 값이 0일때, 즉 인스턴스들이 메소드에 의해 공유될수록 높은 응집도를 가진다고 가정하였다. CBO는 결합되어 있는 다른 클래스의 수로 정의한다. 결합되어 있다는 것은 하나의 메소드가 다른 클래스의 메소드나 인스턴스를 사용한다는 것을 의미한다. 반면 RFC는 한 클래스에서 다른 클래스의 메소드를 호출하는 수를 계산한다. CBO나 RFC가 높으면 결합도가 높아 모듈의 재사용을 어렵게 만든다.

위에서 살펴본 바와 같이 소프트웨어 모듈화에서 모듈의 응집도와 결합도는 중요한 역할을 한다. 일반적으로 모듈내의 응집도가 높고, 즉 내부의 엔터티 간의 밀접한 관계를 가지고, 외부 모듈과의 결합도가 낮은 모듈이 우수한 모듈로 평가를 받게 된다.

2.2.2 온톨로지 메트릭

소프트웨어 모듈간의 결합도가 클 때와 같이 응용 시스템에서 사용하는 온톨로지간의 결합도가 크면 시스템을 이해하고, 수정하기 어려울 뿐 아니라 온톨로지를 사용하는 시스템의 복잡도에도 영향을 줄 수 있다. 그러므로 온톨로지의 품질을 측정할 수 있는 메트릭이 필요하다. Yao[17]와 Orme[12]은 온톨로지가 외부에서 정의된 개념을 레퍼런스하지 않고 자체적으로 정의한 개념들만을 이용하는 것이 바람직하다고 간주하고 온톨로지 결합도와 응집도를 측정하는 메트릭인 NEC, REC, RI, NoR, NoL, ADIT_LN을 <표 1>과

같이 정의하였다.

또한 Ma[9]는 동적으로 변화하는 환경하에
서 온톨로지 시스템의 응집도를 측정하는 NOP,

NMIS, AVAI 메트릭을 제안하였다(<표 1>
참조). 이들 메트릭은 온톨로지의 구조적인 측
면보다 의미적인 측면을 고려하여 온톨로지

<표 1> 온톨로지 메트릭 비교

저자	평가대상	메트릭	특 성
Yao et al.[17]	결합도	NEC	(Number of External Classes) 온톨로지 외부에서 정의된 클래스의 수
		REC	(Reference to External Classes) 온톨로지내에서 외부 클래스를 레퍼런스하는 수
		RI	(Reference Includes) 온톨로지 정의 파일의 상단에 명시된 외부 클래스 포함 수
Orme et al.[12]	응집도	NoR	(Number of Root class) 온톨로지 내 정의된 뿌리 노드의 수
		NoL	(Number of Leaf class) 온톨로지 내 정의된 종단 노드의 수
		ADIT-LN	(Average Depth of Inheritance Tree of all Leaf Nodes) 경로의 평균 길이 : 온톨로지 내 모든 경로 길이의 합계를 경로 수로 나눈 값
Ma et al.[9]	응집도	NOP	(Number Of Partitions)
		NMIS	(Number of Minimally Inconsistent Subsets)
		AVAI	(Average Value of Axiom Inconsistencies)
Yang et al.[15]	복잡도	μ	개념당 평균 관계 수. $\mu = TNOR/TNOC$.
		ρ	개념당 평균 경로 수. $\rho = TNOP/TNOC$
		σ	최대 경로 길이와 평균 경로 길이의 비율. $\sigma = \Lambda/\nabla$
	기본사항	TNOC	(Total Number of Class) : 총 클래스 수
		TNOR	(Total Number of Relation) : 총 관계 수
		TNOP	(Total Number of Paths) : 총 경로 수
		λ_i	클래스 c_i 가 속한 경로의 최장 길이
		Λ	온톨로지의 최장 경로 길이
		∇	온톨로지의 평균 경로 길이
Burton -Jones et al.[1]	구문(Syntactic Quality)	SL	Lawfulness 구문의 정확성
		SR	Richness 구문의 포괄성
	의미(Semantic Quality)	EI	Interoperability 용어의 의미
		EC	Consistency 용어 의미의 일치성
		EA	Clarity 단어 의미의 평균 가짓수
	실질(Pragmatic Quality)	PO	Comprehensiveness 클래스와 속성의 수
		PU	Accuracy 정보의 정확성
		PR	Relevance 정보의 관련성
	사회(Social Quality)	OT	Authority 다른 온톨로지가 의존하는 정도
		OH	History 온톨로지가 사용된 횟수

가 진화하는 과정에서 온톨로지 partition간의 서로 모순된(inconsistent) 정도를 측정하여 응집도를 산출한다.

한편 온톨로지의 복잡도를 측정하기 위한 메트릭에 대한 연구도 이루어졌다. Yang[15]은 온톨로지 복잡도를 측정하는 메트릭을 제안하고 GO(Gene Ontology)에 적용하였다. 이들이 제시한 평가 메트릭은 온톨로지의 복잡도뿐 아니라 개발자들이 온톨로지의 크기를 알아볼 수 있도록 정보를 제공하고 메트릭의 결과를 바탕으로 필요 시 스케일링 할 수 있도록 하였다. 복잡도 메트릭(Complexity Metrics)은 기본 메트릭(Primitive Metrics)을 바탕으로 정의된다(<표 1> 참조).

또한 Burton-Jones[1]는 기호학 이론에 근거하여 온톨로지의 질을 평가하는 프레임워크를 제시하였다. 이들은 온톨로지를 구분적, 의미적, 실질적, 그리고 사회적 기준으로 평가하였다.

2.2.3 메트릭 검증 프레임워크(Metrics Validation Framework)

Kitchenham[8]의 검증 프레임워크는 메트릭의 검증을 위해 널리 사용되어 왔으며 측정 방법으로서 관측되어야 할 항목과 속성들을 정의하였다. 즉, 메트릭의 관측 대상인 엔티티(entity)와 엔티티의 속성(attribute)을 측정하는 단위(unit)와 척도(scale) 등을 정의하고 메트릭이 타당하기 위해서는 다음과 같이 속성, 단위, 도구, 프로토콜 타당성을 만족하여야 한다고 정의하였다.

- 속성 타당성 : 분석 대상인 엔티티가 속성을 가지는가?

- 단위 타당성 : 정의된 단위가 속성에 적합한가?
- 도구적 타당성 : 자체적으로 개발된 도구가 중복 카운팅하지 않고 측정을 정확하게 하는가?
- 프로토콜 타당성 : 모든 경우에 대해 측정이 일관성이 있고 모호하지 않는가?

또한 간접 척도인 경우 다음을 추가적으로 제시하여야 한다.

- 속성간의 관계 모델
- 차원 사용의 적절성
- 척도의 불연속성 없음
- 스케일 단위의 적절성

3. 연구 모델

본 장에서는 모듈을 평가하는 기준을 제시하고 더 나아가 모듈의 평가를 통하여 모듈화 일치성을 평가하고자 한다. 즉, 모듈이 갖추어야 할 일반적 특성을 기반으로 온톨로지 모듈을 정의하고 온톨로지 모듈을 평가하는 방법을 제시한다. 생성된 모듈이 모듈 특성 및 기준을 만족하는지 평가함으로써 모듈화 시 원본 온톨로지가 보유한 정보가 손실이 없는지 일치성을 판단할 수 있다.

온톨로지 모듈을 평가하기 위해 소프트웨어 모듈의 품질을 측정하는 방법에 대한 기존 연구들을 참조하였으며, 이를 온톨로지 연구 분야에 적용하였다. 즉, 제 2.2.1절에서 살펴본 바와 같이 소프트웨어 모듈의 클래스는 인스턴스와 메소드로 구성되고 클래스내 인

스턴스들이 메소드들에 의해 공유될 때 응집도가 높으며 메소드가 다른 클래스의 메소드나 인스턴스를 사용하는 경우 결합도가 높아진다. 그러나 온톨로지는 모듈에는 메소드 개념이 없으므로 온톨로지 분야에 그대로 적용할 수 없다. 그러므로 온톨로지 모듈의 클래스와 관련 속성에 맞게 메트릭을 새로이 정의할 필요가 있다. 본 장에서는 온톨로지 모듈의 특성을 정리하고 모듈을 평가하기 위한 기준을 제시한다.

3.1 온톨로지 모듈

온톨로지 모듈의 평가를 위해서 온톨로지 모듈에 대해 우선 정의되어야 한다. 본 장에서는 제 2.1절 온톨로지 모듈화에 관한 연구에서 살펴본 온톨로지 모듈의 요건을 바탕으로 본 연구에서의 모듈, 모듈 응집도, 모듈 결합도에 대해 정의한다.

정리 1 : 온톨로지 모듈(Ontology module)

온톨로지 모듈은 내용량 혹은 복잡한 온톨로지로부터 분리한 재사용 가능한 구성요소를 지칭한다. 모듈은 원본 온톨로지 혹은 다른 온톨로지 모듈과 명확(definite)한 관계를 유지하며 자기 충족적(self-contained)이어야 한다. 자기 충족적이란 온톨로지 모듈에서 정의한 모든 개념들이 모듈 내 개념들에 의해 정의되며 모듈 외부의 다른 개념을 참조하지 않는 것을 의미한다. 그러므로 온톨로지 모듈은 자체적으로 정확하고 완전하게 정의된 클래스와 관련 사항들의 그룹이다. 온톨로지 모듈의 재사용과 관리의 효율을 위해서 모듈 응집도가 높고 모듈 결합도가 낮게 모듈을

구성하는 것이 바람직하다.

정리 2 : 온톨로지 모듈 응집도(cohesion)

온톨로지 모듈을 구성하는 클래스들간의 관련성을 나타낸다. 이 때, 클래스간의 관련성은 클래스가 구조적으로 모듈내의 다른 클래스들과 관계된 정도를 나타낸다. 모듈 내 클래스간 계층 관계나 연관 관계가 많을수록 응집도가 높다고 정의한다.

정리 3 : 온톨로지 모듈 결합도(coupling)

온톨로지 모듈이 다른 모듈과 가지는 관련성의 정도를 나타낸다. 즉, 한 모듈의 클래스들이 다른 모듈의 클래스들과 관계가 많을수록 높은 결합도를 가진다고 정의한다.

정리 4 : 관계 단절(Relation disconnection)

원본 온톨로지의 클래스들이 가진 계층 관계 혹은 연관 관계가 모듈화 과정을 통해 보존되지 않고 상실되는 경우를 관계의 단절상태로 정의한다.

정리 5 : 모듈화 일치성(Consistency)

원본 온톨로지의 공리(axiom)가 모듈화 이후 생성된 모듈에 보존된 경우 원본 온톨로지 와 생성 모듈이 일치(consistent)한다고 정의한다. 그러므로 관계의 단절이 심하게 일어날수록 일치성(consistency)은 떨어지게 된다.

3.2 온톨로지 모듈 메트릭

본 절에서는 OWL 온톨로지 모듈의 결합도와 응집도를 측정하기 위한 메트릭을 제안한다.

3.2.1 결합도 메트릭(Coupling metrics)

온톨로지 모듈의 결합도를 정의하기 위해 제 2.2.1절에서 소프트웨어 메트릭에 대하여 살펴보았다. 객체지향 소프트웨어에서는 결합도를 정적 결합도와 동적 결합도로 구분한다. 정적 결합도는 클래스와 클래스의 정적인 관계를 근간으로 하며 동적인 결합도는 클래스의 메소드의 호출이나 메소드 간의 인스턴스 변수의 공유 정도를 결합도의 척도로 측정한다[2]. 그러나 온톨로지에서는 메소드 개념이 없으므로 클래스와 다른 모듈의 클래스와의 정적 관계의 조밀도를 모듈 결합도로 간주하고 메소드 호출에 의한 동적인 측면은 고려하지 않기로 한다.

온톨로지를 구성하는 관계는 계층 구조를 구성하는 계층 관계와 연관 관계로 구분할 수 있다. 그런데 제 2장 관련 연구에서 살펴본 바와 같이 Orme[12]의 결합도 메트릭 NEC, REC, RI 메트릭에서는 온톨로지에서 참조하는 외부 클래스의 수로 결합도를 평가하므로 결과적으로 외부 클래스와 맺고 있는 관계의 유형에 상관없이 모든 관계를 외부 클래스 수로만 평가한다. 그러나 온톨로지의 계층 구조는 관계의 유형에 따라 서로 다른 강도의 결합력을 나타낸다. 즉, 계층 구조상에서 계층 관계는 상속의 특성에 의해 상속 정보를 공유하게 되며 모듈화에 의한 계층관계의 단절은 계층관계 이외의 관계의 단절에서 발생하는 공리(axiom)의 손실과는 상이하게 상대적으로 많은 양의 정보를 잃게 된다. 예를 들면 계층 관계는 상속에 의해 많은 양의 속성을 공유하게 되므로 강한 결합력을 나타낸다고 볼 수 있다. 그러므로 관계의 유형에 따라 서

로 구분하여 계산하는 것이 필요하며 이를 반영하기 위한 온톨로지 모듈의 결합도 메트릭이 요구된다.

그러므로 본 연구에서는 소프트웨어 메트릭과 온톨로지 메트릭에 대한 선행 연구들을 바탕으로 관계 유형에 따라 온톨로지 모듈 결합도를 측정하는 메트릭인 NSHL과 NSRL을 새롭게 정의한다.

- NSHL(Number of Separated Hierarchical Links)

NSHL은 원래 원본 온톨로지에 존재했던 계층 관계가 모듈화에 의해 단절되는 수를 의미한다. 계층 관계는 계층 구조상의 상속에 의해 부모 혹은 조상 노드들의 속성을 상속받아 공유하므로 연관 관계 보다 강력한 결합력을 가진다고 가정한다. 본 연구에서는 OWL 온톨로지를 대상으로 하며 OWL 구문에서 subclass 구문으로 명확히 정의된 경우 계층관계로 간주한다.

정의 1 : 계층 관계 단절도(NSHL)

모듈 M의 계층관계 단절도(NSHL)를 다음과 같이 정의한다.

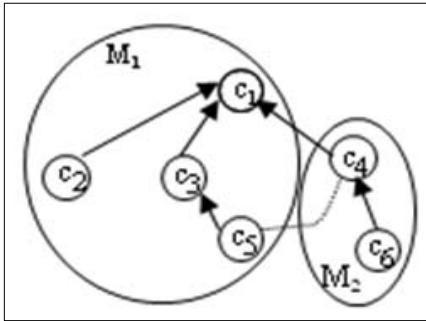
$$NSHL(M) = \sum_i \sum_j nshr(c_i, c_j),$$

$nshr(c_i, c_j)$: 클래스 c_i 와 c_j 사이의 단절되는 계층 관계의 수.

$c_i \in M, c_j \in \{O-M\}$. O : 원본 온톨로지, M : 모듈, - : 차집합 연산자.

예를 들면 아래 <그림 1>에서 NSHL을 계산하면 모듈 M_1 과 M_2 사이를 연결하는 계층

관계의 수는 1개이므로 $NSHL(M_1) = NSHL(M_2) = 1$ 이 된다.



<그림 1> 온톨로지 모듈 예

• NSRL(Number of Separated Relational Links)

NSRL은 원본 온톨로지에 존재했던 여러 형태의 관계 중에서 모듈화에 의해 단절되는 연관 관계 수를 의미한다. 연관 관계는 계층 관계가 아닌 일반적 관계를 의미하며 계층구조를 구성하는 관계보다 결합력이 약하다고 간주한다. 본 연구에서는 OWL 온톨로지를 대상으로 하며 OWL 구문에서 ObjectProperty 등으로 명확히 정의된 경우 연관 관계로 간주한다.

정의 2 : 연관 관계 단절도(NSRL)

$$NSRL(M) = \sum_i \sum_j nsnr(c_i, c_j)$$

$nsnr(c_i, c_j) = 1$ 클래스 c_i 와 c_j 사이의 단절되는 연관 관계의 수.

$c_i \in M, c_j \in \{O-M\}$. O : 원본 온톨로지, M : 모듈, - : 차집합 연산자.

예를 들면 <그림 1>에서 NSRL을 계산하면 모듈 M_1 과 M_2 사이를 연결하는 연관관계

의 수는 1개이므로 $NSRL(M_1) = NSRL(M_2) = 1$ 이 된다.

새로 정의하는 NSHL과 NSRL 메트릭은 선행 연구와 비교하여 다음과 같은 측면에서 구별된다. 첫째, 온톨로지 모듈 결합도는 단순히 외부 클래스에 대한 참조 수가 아니라 관계의 유형에 따른 중요도를 결합도의 영향 요인으로 새롭게 반영할 수 있다. 즉, 모듈 결합도는 모듈화에 의해 단절되는 계층관계인 IS-A 관계의 수인 NSHL과 연관 관계의 수인 NSRL에 의해 가변적으로 영향을 받는다. 이때, 클래스와 클래스간의 계층 관계는 연관 관계보다 강력한 연관성을 가지고 있다고 간주한다. 그러므로 모듈에서 계층구조의 계층 관계가 단절되는 경우 결합도가 더 높아진다. 둘째, NSHL과 NSRL 메트릭은 모듈화 과정의 일치성을 평가할 수 있다. 즉, 기존의 메트릭들은 모듈화 과정의 품질 측정과 직접적인 관련성이 없는 반면 본 연구에서 제안하는 메트릭은 모듈화 과정에 손실되는 클래스, 관계, 공리 등이 모듈화 과정에 어느 정도 단절되는 지 측정하고 단절 정도가 큰 경우 일치성이 떨어진다고 간주한다.

3.2.2 응집도 메트릭(Cohesion metrics)

소프트웨어 응집도 메트릭은 한 클래스 내에서의 메소드 간의 동일한 속성을 공유하는 정도나 지역 변수 혹은 메소드를 사용하는 정도로 측정한다. 그러나 온톨로지 모듈의 응집도는 한 모듈 내에서 클래스간에 가지는 관계 수에 따라 응집도를 측정한다.

또한 선행 연구에서 살펴본 바와 같이 Yao [17]는 NoR, NoL, ADIT-LN을 응집도 메트릭으로 정의하였지만 ADIT-LN은 온톨로지

를 구성하는 계층 구조의 깊이를 위주로 응집도를 측정하며, 계층관계 이외의 일반적인 관계에 대해서는 반영하지 못하는 한계가 있다. 또한 NoR, NoL 역시 계층 구조의 최상단과 최하위 클래스의 수를 기반으로 한다. 그러므로 계층관계 이외의 연관 관계를 응집도에 반영할 수 있는 새로운 메트릭의 필요성에 따라 메트릭 NR를 제안한다.

•NR(Number of Relations)

온톨로지 모듈에 정의된 클래스가 보유한 관계(relation)의 수를 의미하며 모듈의 내부에서 클래스들이 가지는 계층 관계뿐 아니라 연관 관계의 조밀도를 응집도의 척도로 정의한다.

정의 3 : 클래스당 관계 수(NR)

$$NR(M) = \sum_i \sum_j \frac{sr(c_i, c_j)}{|M|}$$

$sr(c_i, c_j)$

$$= \begin{cases} \frac{1}{distance(c_i, c_j)} & c_i \text{와 } c_j \text{가 서로 연결된 경우} \\ 0 & c_i \text{와 } c_j \text{가 서로 연결되지 않은 경우} \end{cases}$$

$distance(c_i, c_j)$ = 클래스 c_i 와 c_j 사이의 최단 거리

클래스 $c_i \in M, c_j \in M, M$: 모듈, $||$: 카디널리티

예를 들면 <그림 1>에서 모듈 M_1 의 NR을 계산하면 $(1+1+1+1/2)/4 = 7/8$ 이 된다.

4. 검증(Validation)

본 연구에서는 메트릭 표준 검증 프레임워

크를 사용하여 새로이 제안하는 메트릭의 이론적인 타당성을 검증하고 웹 상에서 개방된 임의의 온톨로지를 모듈화하여 응집도와 결합도를 측정하는 실증적 실험을 수행하였다.

4.1 메트릭의 이론적 검증

본 연구에서는 제 2.2.3절에 기술된 Kitchenham[8]의 검증 프레임워크를 이용하여 새로이 제안하는 모듈 평가 메트릭을 검증한다

4.1.1 결합도 이론적 타당성 검증

본 연구에서 새로 정의한 모듈화 결합도 측정 메트릭 NSHL(Number of Separated Hierarchical Links)과 NSRL(Number of Separated Relational Links)은 모듈화에 의해 모듈 외부와 연결이 단절된 관계의 수를 세는 직접 척도이다. Kitchenham의 메트릭 측정 타당성 프레임워크에서 NSHL과 NSRL은 직접 메트릭으로서 타당성을 검증하기 위해, 다음과 같이 관련 속성을 정의하였다.

NSHL

- 엔터티는 메트릭 적용 대상인 온톨로지 모듈이다.
- 측정될 속성은 단절되는 계층 관계의 수이다.
- 단위는 계층 관계이다.
- 데이터 척도는 등간 척도이다.

NSRL

- 엔터티는 메트릭 적용 대상인 온톨로지 모듈이다.
- 측정될 속성은 단절되는 연관 관계 수

이다.

- 단위는 연관 관계이다.
- 데이터 척도는 등간 척도이다.

위의 기준이 속성, 단위, 도구, 프로토콜 타당성을 검증하는 기준이 되며 NSHL과 NSRL은 각각의 메트릭에 대한 타당성을 다음과 같이 만족한다.

- 모든 경우에 대해 속성 타당성을 가진다. 즉, NSHL의 속성은 외부 클래스와 가지는 IS-A관계이며 NSRL의 속성은 모듈 외부 클래스와 가지는 연관관계의 수이다. 온톨로지 모듈은 명시적으로 정의된 계층관계와 연관 관계를 모두 포함한다.
- 모든 경우에 대해 단위 타당성을 가진다. 즉, NSHL의 측정 단위는 IS-A 관계이며 NSRL의 경우는 연관 관계이다.
- 모든 경우에 대해 도구적 타당성을 가진다. 즉, 속성인 관계의 수를 계산할 때 측정 알고리즘은 중복 없이 정확하게 측정한다.
- 모든 경우에 대해 프로토콜 타당성을 가진다. 즉, 제안된 측정 도구들은 일관성이 있으며 모호하지 않다.

이와 같이 본 연구에서 제안한 NSHL과 NSRL의 메트릭은 Kitchenham 검증 프레임워크의 타당성을 모두 만족한다.

4.1.2 응집도 이론적 타당성 검증

새로이 정의한 모듈 응집도 측정 메트릭 NR (Number of Relation)은 모듈 내부의 클래스당 관계의 수를 계산하는 간접 척도이다. Kitchenham의 검증 프레임워크의 관련 속성들

을 다음과 같이 정의하였다.

- 엔터티는 온톨로지 모듈이다.
- 측정될 속성은 모듈 내부의 관계의 수이다.
- 단위는 클래스간 관계이다.
- 데이터 척도는 비율 척도이다.

NR 메트릭은 간접 척도로서 Kitchenham의 검증 프레임워크의 타당성을 다음과 같이 만족한다.

- NR은 속성 타당성을 가진다. 즉, 엔터티인 온톨로지 모듈에서 측정 속성인 관계가 명확히 정의되어 있다.
- NR은 단위 타당성을 가진다. 즉, NR의 측정 단위는 클래스가 보유한 관계이다.
- NR은 도구적 타당성을 가진다. 즉, NR은 측정 알고리즘을 통하여 중복 없이 정확하게 측정된다.
- NR은 프로토콜 타당성을 가진다. 즉, 본 연구에서 정의한 NR의 정규 수식은 모호하지 않으며 일관성이 있다.

또한 NR은 간접 척도이므로 다음을 추가적으로 만족함을 보인다.

- 엔터티와 속성간의 관계 모델 : 엔터티인 온톨로지 모듈은 모듈 내 관계가 존재하기 위해 필요하다.
- 차원 사용의 적절성 : 클래스당 관계 수를 구하기 위해 모듈의 클래스 수로 나누는 것은 적합하다.
- 관계의 수 계산의 불연속성이 없다.
- 스케일 단위 : 모듈 내 정의된 관계는 모듈의 속성을 측정하기에 적합한 단위이다.

4.2 실증적 검증

개발된 메트릭을 실증적으로 검증하기 위해서 메트릭을 적용할 온톨로지 모듈을 다음과 같이 준비하였다. 즉, 실험 대상이 될 모듈화된 온톨로지를 교육 도메인에서 얻기 위해 SWOOGLE 검색엔진을 이용하여 웹 상에서 “Student”, “University” 검색어로 검색하여 Koala.owl[18], Iswc.owl[19], Univ-bench.owl[20] 등 세 개의 온톨로지를 얻었다. 이들 온톨로지에 모듈화 도구인 Prompt factor, PATO와 SWOOP를 적용하여 모듈화하였다. 그런데 모듈화 도구에 따라 OWL 파일대신 뷰(View)를 제공하는 경우도 있는데 이러한 경우 뷰를 실증적 검증에 사용하였다. 즉, SWOOP은 온톨로지 모듈을 독립된 파일 형태로 추출해 낼 수 있지만 PATO와 Prompt factor는 모듈의 뷰만을 제공하므로 모듈 뷰를 실증적 검증에 사용하였다. 모듈화 결과 <부표 1>과 같이 30개의 모듈이 추출되었으며 추출된 모듈을 대상으로 모듈 응집도, 결합도 메트릭을 적용하였다.

또한 메트릭의 검증을 위해 전문가의 판단을 이끌어내는 실험실 실험 방식이 이용되었다. 6명의 소프트웨어 전문가들을 선정하여 모듈의 응집도와 결합도를 평가하도록 하였다. 실험에 참여한 전문가들은 10년 이상의 소프트웨어를 개발한 경험이 있으며 모듈화에 대한 사전 지식이 있는 사람들로, 온톨로지 모듈의 특성에 대한 설명을 추가적으로 교육하였다. 모듈화 도구를 적용하여 나온 30개의 모듈들을 전문가들에게 모듈의 내부 구조를 알 수 있는 그래프 형태의 자료로 제시하고 모듈 응집도와 결합도를 다음의 척도로

평가하도록 하였다.

- 0 = Low
- 0.25 = Moderately
- 0.5 = Average
- 0.75 = High
- 1.0 = Excellent

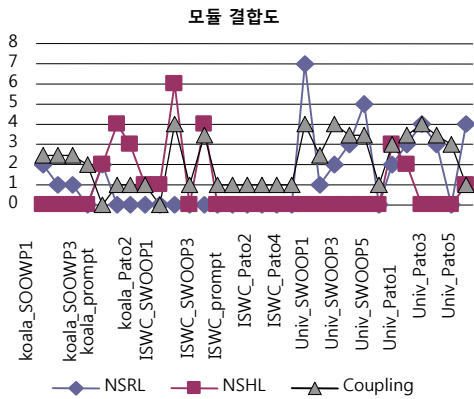
메트릭을 적용한 결과값과 평가자들의 모듈 평가 결과의 평균값을 비교하여 두 결과값이 관련성이 있다는 것을 보임으로써 메트릭의 타당성을 보였다.

본 실증적 연구를 위해 메트릭 값을 계산하기 위한 프로토타입 evaluator : ModEval를 개발하여 실험에서 적용하였다. 프로토타입은 윈도우 XP 그리고 Visual C++ 6.0 환경에서 개발되었다.

4.2.1 메트릭을 적용한 결과

본 절에서 새롭게 제안한 결합도 메트릭 NSHL과 NSRL을 온톨로지 모듈에 적용하여 보았다. 다음 <부표 1>은 Koala.owl, Iswc.owl, Univ-bench.owl 등의 온톨로지에 SWOOP, Prompt와 PATO를 적용하여 생성한 모듈에 대하여 모듈화 결합도 응집도 메트릭을 적용한 결과이다. NR의 경우 모듈화 기법에 독립적으로 끌고루 분포하였다. 또한 NSHL은 Koala, Univ-bench 온톨로지에 SWOOP을 사용하여 모듈화한 대부분의 모듈에서, PATO를 사용하였을 보다 상대적으로 낮은 결합도를 나타내었다. 이와는 반대로 NSRL은 대부분의 모듈에서 SWOOP을 사용하였을 경우가 PATO를 사용하였을 경우보다 상대적으로 높은 결합도를 보여주었다. 이와 같은 결과가 나오게 된 원인은 모듈화 도구 SWOOP

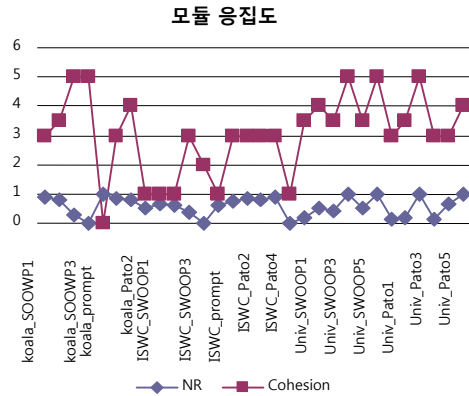
이 계층관계의 보존을 중요시하여 모듈화 시 계층관계를 대부분 보존시켜서 NSHL 값이 낮 으며 반하여 PATO의 경우에는 상대적으로 연관관계가 보존이 잘 되기 때문에 NSRL 값이 낮게 나타난 것으로 추정된다(<그림 2> 참조).



<그림 2> 모듈 결합도 결과값

또한 선행 연구에서 제시되었던 NEC, REC, RI는 모듈화를 고려하지 않고 외부 클래스의 참조 수만을 반영함으로써 모듈화 과정 중에 잃게 되는 관계 정보들을 반영하지 못하는 문제점이 있지만 메트릭 NSHL, NSRL을 적용한 결과 모듈화 과정 중에 단절되는 관계들을 관계 유형별로 구분하여 메트릭 측정에 반영하여 모듈의 결합도 측정의 정확도를 개선시킬 수 있다. 또한 모듈화 이전의 원본 온톨로지와 모듈화로 생성된 모듈들이 보유한 관계 정보를 서로 비교함으로써 손실된 관계의 수를 모듈화의 일치성을 평가하는 용도로 사용할 수 있다. 즉, 결합도가 높은 모듈은 손실된 정보가 높은 것으로 간주되고 모듈화 일치성이 떨어지는 것으로 평가할 수 있다.

온톨로지 모듈의 결합도를 측정하는 실험과 동일한 방법으로 SWOOP, PATO, Prompt 등 세 가지 모듈화 방식을 적용하여 나온 모듈에 응집도 메트릭 NR을 적용하여 <그림 3>과 같은 결과가 나왔다. SWOOP로 분할된 모듈의 경우 응집도의 차이가 비교적 크게 나는 것을 알 수 있다. 이와 같은 현상은 일부 모듈에 많은 관계들이 밀집되게 분할되기 때문인 것으로 해석할 수 있다. Prompt로 추출된 모듈의 경우 비교적 높은 응집도를 보였는데 이는 대부분의 관련 노드와 관계들이 모듈 추출 시 보존되었기 때문이다. PATO의 경우 다양한 크기의 응집도를 나타내었다.



<그림 3> 모듈 응집도 결과값

4.2.2 전문가 평가 결과와의 비교 분석

전문가 평가의 신뢰도 검증을 위해 Inter-rater 신뢰도를 측정하여 0.91을 얻었다. 이는 전문가들 평가 결과 사이에 신뢰도가 있음을 나타낸다. 피어슨 상관 분석을 통하여 전문가들의 평가 결과와 메트릭의 계산 값 사이에 <표 2>, <표 3>과 같이 유의한 상관 관계가 있음을 알 수 있었다.

〈표 2〉 응집도 측정값과 전문가 평가와의 피어슨 상관 분석

Metrics	Correlation Coefficient ρ
NR	0.540**

주) ** significant at the 0.01 level.

〈표 3〉 결합도 측정값과 전문가 평가와의 피어슨 상관 분석

Metrics	Correlation Coefficient ρ
NSHL	0.385*
NSRL	0.481**

주) ** significant at the 0.01 level.

* significant at the 0.05 level.

5. 결 론

본 논문은 온톨로지의 모듈을 평가할 수 있는 새로운 결합도 메트릭과 응집도 메트릭을 제안하고 제안한 온톨로지 모듈 응집도 메트릭과 결합도 메트릭에 대하여 이론적 실증적 검증을 수행하여 검증하는 속성들이 만족됨을 보였다.

본 논문에서 제안한 응집도와 결합도 메트릭은 이전 연구들의 한계를 보완하여 첫째, 계층관계와 연관관계를 모두 고려한 응집도를 산출할 수 있고 둘째, 서로 다른 모듈의 클래스간 관계의 유형을 고려하여 좀 더 세분화된 결합도에 대한 분석이 가능하고 셋째, 온톨로지 모듈화에 따라 단절되는 관계를 고려함으로써 모듈화 일치성을 판단할 수 있다는 측면에서 의의가 있다. 본 연구에서 제안된 메트릭은 향후 온톨로지 모듈화 기법들과 모듈화된 온톨로지를 평가할 수 있는 척도로 활용될 수 있을 것이다. 향후 과제로는

본 논문에서 정의한 결합도와 응집도 지표를 지속적으로 발전시키는 것이며 개발된 메트릭을 실무에 적용하여 그 유용성을 입증하는 것이다.

참 고 문 헌

- [1] Burton-Jones, A., Storey, V., and Sugumaran, V., and Ahluwalia, P., "A semiotic metrics suite for assessing the quality of ontologies," *Data Knowl. Eng.*, Vol. 55, No. 1, 2005, pp. 84-102.
- [2] Chidamber, S. R. and Kemerer, C. F., "A Metric Suite for Object Oriented Design," *IEEE Transactions on Software Engineering*, Vol. 17, No. 6, 1994, pp. 36-638.
- [3] Constantine and Yourdon, *Structured Design*, Prentice-Hall, Englewood Cliffs, NJ, 1979.
- [4] Cuenca Grau, B., Parsia, B., Sirin, E., and Kalyanpur, A., *Automatic Partitioning of OWL Ontologies Using E-Connections*, In : *Proc. of Description Logic Workshop*, 2005.
- [5] d'Aquin, M., Sabou, M., Motta, E., *Modularization : a Key for the Dynamic Selection of Relevant Knowledge Components*, In : *Proc. of the ISWC 2006 Workshop on Modular Ontologies*, 2006.
- [6] d'Aquin, M., Doran, P., Motta, E., Tamma, V., *Towards a Parametric Ontolo-*

- gy Modularization Framework Based on Graph Transformation, Workshop : International Workshop on Modular Ontologies, 2007.
- [7] Doran, P., Tamma, V., and Iannone, L., Ontology module extraction for ontology reuse : an ontology engineering perspective, CIKM 2007.
- [8] Kitchenham, B., Pfleeger, S. and Fenton, N., "Towards a Framework for Software Measurement Validation," IEEE Trans. Software Eng., Vol. 21, No. 12, 1995.
- [9] Ma, Y., Jin, B., and Feng, Y., "Semantic oriented ontology cohesion metrics for ontology-based systems," Journal of Systems and Software, Vol. 83, No. 1, 2010.
- [10] McCabe, T. J., "A complexity measure," IEEE Transactions on software engineering, Vol. 2, No. 4, 1976, pp. 308-320.
- [11] Noy, N. and Musen, M., Specifying Ontology Views by Traversal, In : Proc. of the International Semantic Web Conference (ISWC), 2004.
- [12] Orme, A. M., Yao, H. and Etkorn, L. H., "Coupling Metrics for Ontology-Based Systems," IEEE Software, Vol. 23, No. 2, 2006, pp. 102-108.
- [13] Page-Jones, M., Practical Guide to Structured Systems Design, Yourdon, 1980.
- [14] Seidenberg, J. and Rector, A., Web ontology segmentation : analysis, classification and use, Proceedings of the 15th international conference on World Wide Web, 2006.
- [15] Stuckenschmidt, J. and Klein, M., Structure-Based Partitioning of Large Concept Hierarchies, In : Proc. of the International Semantic Web Conference (ISWC), 2004.
- [16] Yang, Z., Zhang, D., and Ye, C., Evaluation Metrics for Ontology Complexity and Evolution Analysis, IEEE International Conference on E-Business Engineering, 2006.
- [17] Yao, H., Orme, A. M., and Etkorn, L., "Cohesion Metrics for Ontology Design and Application," Journal of Computer science, Vol. 1, No. 1, 2005, pp. 107-113.
- [18] <http://protege.stanford.edu/plugins/owl/owl-library/koala.owl>.
- [19] <http://annotation.semanticweb.org/ontologies/iswc.owl>.
- [20] <http://www.lehigh.edu/~zhp2/2004/0401/univ-bench.owl>.

〈부표 1〉 메트릭 계산결과

Ontology	Modular Approaches	Modules	NR	NSRL	NSHL	cohesion	coupling
Koala	SWOOP	Module 1	Low(0.92)	Medium(2)	Low(0)	3	2.5
		Module 2	Low(0.8)	Low(1)	Low(0)	3.5	2.5
		Module 3	Low(0.3)	Low(1)	Low(0)	5	2.5
		Module 4	Low(0)	Low(0)	Low(0)	1	2
	Prompt	Module α	Medium(1)	Low(0)	Medium(2)	4	1
	PATO	Module A	Low(0.86)	Medium(1)	Medium(4)	3	1
		Module B	Low(0.8)	Low(1)	Medium(3)	4	1
		Module C	Low(0.5)	Low(1)	Low(1)	1	1
Iswc	SWOOP	Module 1	Low(0.67)	Medium(6)	Low(0)	1	4
		Module 2	Low(0.63)	Low(0)	Low(0)	3	1
		Module 3	Low(0.36)	Medium(4)	Low(0)	2	3.5
		Module 4	Low(0)	Medium(10)	Low(0)	1	5
	Prompt	Module α	Low(0.63)	Low(0)	Low(0)	1	1
	PATO	Module A	Low(0.75)	Low(0)	Low(0)	3	1
		Module B	Low(0.86)	Low(1)	Low(0)	3	1
		Module C	Low(0.8)	Low(1)	Low(0)	3	1
		Module D	Low(0.89)	Medium(2)	Low(0)	3	1
		Module E	Low(0)	Medium(2)	Low(0)	1	1
Univ-bench	SWOOP	Module A	Low(0.2)	High(10)	Low(0)	3.5	4
		Module B	Low(0.5)	Medium(4)	Low(0)	4	2.5
		Module C	Low(0.41)	High(10)	Low(0)	3.5	4
		Module D	Medium(1)	Low(1)	Low(0)	5	3.5
		Module E	Low(0.5)	Medium(7)	Low(0)	3.5	3.5
	Prompt	Module α	Medium(1)	High(21)	Low(0)	5	1
	PATO	Module A	Low(0.15)	Medium(5)	Low(0)	3	3
		Module B	Low(0.2)	Medium(3)	Low(0)	3.5	3.5
		Module C	Medium(1)	Low(0)	Low(0)	5	4
		Module D	Low(0.14)	Medium(2)	Low(0)	3	3.5
		Module E	Low(0.67)	Low(0)	Low(0)	3	3
Module F		Medium(1)	Medium(4)	Low(0)	4	1	

주) * cohesion : Average of experts' evaluation, coupling : Average of experts' evaluation.

저 자 소개



오선주

1986년

1993년

2009년

2009년~현재

관심분야

(E-mail : ohsunju7@snu.ac.kr)

서울대학교 계산통계학과 (학사)

서울대학교 계산통계학과 전산과학전공 (이학 석사)

서울대학교 경영학과 MIS전공 (경영학 박사)

서울대학교 컴퓨터공학과 박사후연구원

온톨로지, e-비즈니스, Information Technology
Architecture 등