

러프집합과 퍼지적분을 이용한 클래스 재사용도 측정

Measuring The Reusability of Class By Rough Sets and Fuzzy Integral

김영천* · 김혜경* · 최완규* · 김영식** · 이성주*
Young-Cheon Kim* · Hye-Kyoung Kim* · Wan-Kyoo
Choi* · YoughSik-Kim** · Sung-Joo Lee*

*조선대학교 전자계산학과

**서강정보대학 정보통신과

e-mail: yckim@stmail.chosun.ac.kr

요 약

컴포넌트의 재사용도 측정은 컴포넌트가 재사용되는 시점에서 컴포넌트의 이해와 적용을 위해 소요되는 노력의 정도를 측정한다. 여러 연구들이 컴포넌트의 재사용도 측정 방법을 제시하였지만 측정 속성(척도)들과 컴포넌트들의 삽입 삭제의 어려움, 가정된 지식의 요구, 각 측정 속성들에 대한 중요도 제시의 부재 등의 문제들이 있다.

따라서, 본 연구에서는 이러한 문제점들을 해결하기 위해서 실제로 재사용되고 있는 객체지향 컴포넌트들과 여러 연구에서 제시되고 있는 메트릭들을 종합하고, 퍼지 적분과 러프 집합을 이용하여 클래스의 재사용도를 측정한다.

I. 서론

재사용에서 컴포넌트가 재사용에 널리 이용되기 위해서는 먼저 컴포넌트의 재사용성(reusability)에 대한 측정이 수행되어야 한다. 컴포넌트의 재사용성 측정은 두 가지의 단계로 나눌 수 있다. 첫째, 컴포넌트의 재사용 가능성(the potentiality of reusability)이다. 재사용 가능성 측정은 컴포넌트의 획득 시점에서 컴포넌트의 품질을 평가하여 컴포넌트의 재사용 여부를 판정한다. 둘째, 컴포넌트의 재사용도(the degree of reusability)의 측정이다. 컴포넌트의 재사용도 측정은 컴포넌트가 재사용되는 시점

에서 컴포넌트의 이해와 적용을 위해 소요되는 노력의 정도를 측정한다[15].

여러 연구들[3, 4, 7, 10, 15]이 컴포넌트의 재사용도를 측정 메트릭을 제시하고 있지만 다음과 같은 요구조건들을 만족하지 못하고 있다.

- 1) 측정 속성(척도)들과 컴포넌트들의 삽입 삭제를 쉽게 할 수 있어야 한다.
- 2) 타당성에 근거하여 컴포넌트들을 정량적으로 평가할 수 있어야 한다.
- 3) 가정된 지식을 요구하지 않아야 한다.
- 4) 각 측정 속성들의 중요도를 객관적으로 산출할 수 있어야 한다.
- 5) 소프트웨어 재사용에 있어서 현재 실제적으로 가장 많은 재사용이 이루어지고 있는 것이 객체 지향 컴포넌트이므로, 이에 대한 재사용 측정이 이루어져야

한다.

따라서, 본 논문에서는 이러한 문제들을 해결하기 위해서 러프집합과 퍼지척도에 근거하여 객체지향 컴포넌트인 클래스의 재사용도를 측정할 수 있는 방안을 제시한다.

II. 컴포넌트의 재사용도 측정

러프 집합 이론은 애매한 범주들에 대한 새로운 수학적 접근 방법으로 자동 분류, 패턴 인식, 학습 알고리즘 등 다양한 분야에 적용될 수 있으며, 러프 집합의 지식 표현시스템은 각 측정 속성들이 동일한 강도(strength)를 갖는지, 또는 각 측정 속성들이 분류 능력(classification power) 측면에서 얼마나 다른가를 쉽게 산출한다[6, 14].

Sugeno의 퍼지 적분[13]은 어떤 대상이 여러 항목에 대해서 평가되고 각 평가 항목의 중요도에 차이가 있을 때 이들 평가치를 종합하는데 이용될 수 있으며, 특히 주관적인 판단이 개입되는 평가 문제에서도 유용하게 이용될 수 있다.

2.1 측정 인자들

객체지향 컴포넌트의 재사용도를 측정하기 위해서는 기존의 기능 중심 컴포넌트의 평가방법들 뿐만 아니라, 객체 지향 컴포넌트의 평가방법들을 도입하여야 한다[11, 12].

여러 연구들[1, 4, 5, 8, 9, 11]은 실험적인 연구를 통해서 코드 라인수(LOC)는 30 이하, Cyclomatic number는 10 이하, Effort는 1,000 이하가 적합하다고 제시하였고, 하나의 클래스에서의 멤버 함수의 개수는 7 ± 2 가 적합하고, 클래스에서의 LCOM은 2이하, DIT(클래스의 상속 깊이)는 3 이하가 적합하며, IDD(클래스의 상속성)는 0.5 이하가 적합하며, EXT(확장성)는 1이 적합하다고 제시하였다.

그러므로, 본 연구는 기존 연구들이 제안한 분류 기준들을 수용, 종합하여 표 1과 같이 클래스의 재사용 분류기준을 설정하였다.

표 1. 분류 기준 표

측정인자		분류			
		0	1	2	3
멤버함수의 개수		≤ 5	$5 < \sim \leq 7$	$7 < \sim \leq 9$	> 9
LOC	평균	≤ 30	> 30		
	위반	no	yes		
CYC	평균	≤ 10	> 10		
	위반	no	yes		
effort	평균	≤ 10000	> 10000		
	위반	no	yes		
LCOM		≤ 2	$2 < \sim \leq 9$	$9 < \sim \leq 200$	> 200
CAMC		≤ 0.5	> 0.5		
EXT		0	$0 < \sim \leq 0.5$	$0.5 < \sim \leq 1$	> 1
IDD		≤ 0.5	> 0.5		
DIT		≤ 3	$3 < \sim \leq 6$	> 6	

소프트웨어 품질 측정에서 각 품질 평가 척도들에 대한 통합된 하나의 척도가 제시되려면, 소속 등급(membership grade)과 관련되어 특징 지을 수 있는 퍼지집합을 이용할 수 있다[12].

따라서, 본 논문에서는 식(1)과 같은 퍼지 소속함수를 이용하여 각 측정인자들의 측정값을 [0,1]구간으로 정규화한다. 식(1)에서 x_0 와 x_1 의 값들은 표 1에서 제시된 값들 중 최소값과 최대값이다.

$$\mu(x) = \begin{cases} 1 & , x \leq x_0 \\ 1 - 2 \frac{x - x_0}{x_1 - x_0} & , x_0 < x \leq \frac{x_0 + x_1}{2} \\ 2 \frac{x_1 - x}{x_1 - x_0} & , \frac{x_0 + x_1}{2} < x \leq x_1 \\ 0 & , x > x_1 \end{cases} \quad (1)$$

2.2 척도의 중요도

재사용 컴포넌트 품질 평가에서, 모든 품질 특성(측정 속성)들이 동일한 중요성을 가지지는 않기 때문에 객관적으로 가중치를 부여할 수 있는 방법이 사용되어야 한다[2].

러프집합의 지식 표현 시스템에서 조건 속성 집합(C)과 의사결정 속성(D) 집합간의 종속성은 식(2)와 같이 정의되고, C의 부분 집합 C'의 중요도는 식(3)과 정의되며, 전체 속성의 중요도에서 특정 속성의 중요도가 점유하는 비율로 표현되는 속성의 상대적 중요도는 식(4)와 같다[6, 14].

$$k = r_c(D) = \frac{|POS_c(D)|}{|U|} \quad (2)$$

$|POS_c(D)|$: POS_c(D)의 원소의 개수

$|U|$: 전체 집합의 원소의 개수

$$Sc = r_c(D) - r_{c'}(D) \quad (3)$$

$r_c(D)$: C의 D에 대한 종속성

$r_{c'}(D)$: C-C'의 D에 대한 종속성

$$RS_i = \frac{S_i}{\sum_{j=1}^n S_j}, \quad (i, j, \dots, n) \quad (4)$$

RS_i : 속성 i의 상대적 중요도

S_i, S_j : 속성 i, j의 중요도

2.3 재사용도 측정

본 논문에서는 각 척도들이 재사용도에 미치는 영향의 정도를 결합시키기 위해서 Sugeno의 퍼지 적분을 이용한다.

집합 X가 유한 집합이고, P(X)는 X의 멱집합이고, 함수 h(x)는 평가항목 x(x ∈ X)에 대한 평가치이고, E ∈ P(X)에 대해 정의되는 퍼지 척도 g(E)는 전체적인 평가에 대한 평가항목 E의 평가치에 기여하는 정도라 하자. 이 때 $x_i \in X(i=1, 2, \dots, n)$ 에 대해서 $h(x_i) \leq h(x_{i+1})$ 이고, $E_i = \{x_k \mid k=i, i+1, \dots, n\}$ 이면, 평가함수 h의 상대적 중요도 함수 g에 대한 Sugeno의 퍼지적분은 식(5)와 같다.

$$\int_x h(x) \cdot g(x) = \text{Max}_{i=1}^n \text{Min}[h(x_i), g(E_i)] \quad (5)$$

본 논문에서 재사용도는 다음과 같은 측정된다. 1단계는 재사용 컴포넌트 분류를 위해 선택된 측정 속성들에 대한 측정값을 퍼지 소속함수를 이용하여 [0,1] 구간으로 정규화한다. 2단계는 분류 정보에서 러프 집합을 이용하여 각 측정 속성들의 상대적 중요도를 구한다. 3단계는 Sugeno의 퍼지적분을 적용하여 각 인자들의 중요도와 소속 등급을 종합하여 컴포넌트의 재사용도를 측정한다.

III. 실험 및 결과

본 논문에서는 179개에 해당하는 Boland의 OWL 클래스들을 대상으로하여 재사용도를 측정하였다. 실험 대상인 클래스들을 각 측정인자들에 대하여 측정하고, 그 결과값들을 표 1의 분류 기준에 따라 속성값으로 분류하여 정보표를 구성한 후, 식(4)를 적용한 각 척도들의 상대적 중요도는 표 2와 같다. 주어진 컴포넌트들에 식(1)을 적용한 퍼지 소속 등급과 식(5)를 적용한 재사용도 측정값은 표 6과 같다.

표 2. 각 척도들의 상대적 중요도

척도	상대적 중요도
LOC	0
CYC	0
EFF	0
NOM	0.142857
LCOM	0.21822
CAMC	0.059211
EXT	0.182266
VoL	0.028818
VoC	0.019702
VoE	0.088671
IDD	0.009853
DIT	0.206895

표 3. 소속등급과 재사용도 측정값

클래스	재사용도
1	0.5
2	0.758378
3	0.665023
.	.
.	.
.	.
.	.
177	0.6302
178	0.82
179	0.82

본 논문에서 제안한 재사용도 측정은 재사용 컴포넌트들에서 추출된 정보에 근거하여 재사용도를 측정하므로 메트릭의 삽입과 삭제가 용이하다. 또한 러프 집합을 이용하여 재사용 결정에 미치는 각 메트릭들의 중요도를 산출하였으며, 본 연구의 측정 방법은 판단지식의 정확성에 관한 문제를 어느 정도 완화하기 위해서 어떤 가정된 지식에 기반하는 것이 아니라 현업에서의 실험과 연구를 통해 타당함이 인정된 컴포넌트들에 관한 지식에 근거하여 재사용도를 측정다.

IV. 결론

본 논문에서는 러프집합과 퍼지 적분을 이용

한 객체지향 컴포넌트인 클래스의 재사용도 측정을 제안하였다.

179개의 클래스에 대한 재사용도를 측정한 결과 제시한 측정방법은 서론에서 제기된 문제들을 해결할 수 있음을 알 수 있었다.

V. 참고 문헌

- [1] L.J.Arther, *Measuring Programmer Productivity and Software Quality*, John Wiley & Sons, New York, pp.138-142, 1985.
- [2] Balog, A. Trifu, R. Raduta, A.. "Quality Evaluation of Public Administration Software Products: A Practical Study.", *The 5th European Conference on Software Quality*, 1996.
- [3] B.A. Burton, et. al, "The Reusable Software Library", *Software Reuse: Emerging Technology*, IEEE CS Press, pp. 129-137, 1990.
- [4] Caldiera G. and V.R.Basili, "Identifying and Qualifying Reusable Software Components", *IEEE Computer*, pp.61-70, Feb, 1991.
- [5] S.R.Chidamber and C.F.Kemerer, "A Metrics Suite for Object-Oriented Design", *IEEE Trans.Software Eng.*, vol.20, no.6, pp.476-493, June, 1994.
- [6] Pawlak Z., *Rough Sets-Theoretical Aspects of Reasoning about Data*, London: Kluwer Academic Publishers, 1991.
- [7] Prieto-Diaz, R. Freeman, P.. "A Software Classification Scheme for reliability.", *IEEE Software*, Vol.4, pp.6-16, 1987.
- [8] Robert C. Sharble and Samuel S. Cohen, "The Object-Oriented brewery: A Comparison of Two Object-Oriented Development Methods", *ACM SIGSOFT. SE. Note*, Vol.18, No.2, pp.60-73, 1993.
- [9] Szentes.J., Gras j., "Some Practical Views of Software-Complexity metrics and a Universal Measurement Tool", *First Austrlian Software Engineering Conference*, 'Canberra, pp.14-16, May, 1986.
- [10] 권기현 · 이경환. "학습알고리즘을 이용한 재사용 의사결정 지원.", *한국정보과학회지*, Vol. 21, pp.2235-2242, 1994.
- [11] 김재생 · 송영재, "클래스의 재사용을 위한 정보 분석 및 품질척도", *정보과학회 논문지*, Vol.26, No.3, pp.393-400, 1999.
- [12] 이경환 등. "소프트웨어 재이용을 위한 연구.", *연구보고서*, 과학기술처, 1989.
- [13] 이광형, 오길록, *퍼지이론 및 응용*, pp. 1.1-1.40, pp. 9.1-9.36, 홍릉과학출판사, 1991.
- [13] 이성주 · 정환목 · 최완규, *러프 집합과 응용*, 조선대학교출판국, 1998.
- [14] 최완규 외, "컴포넌트의 재사용성 결정 및 측정시스템", *한국정보처리학회:소프트웨어 공학연구회지*, Vol.2, No.4, pp.~, 1999.