

# RIF 규칙을 이용한 지식 공유 서비스에서의 불확실성 문제 분석

김 재 훈\*

## An Analysis of Uncertainty Problems in Knowledge Sharing Service Using RIF Rules

Jaehoon Kim\*

---

이 논문은 2017년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임.  
(No. 2016R1D1A1B03930564).

---

### 요 약

본 논문에서는 시맨틱 웹과 RIF 규칙을 이용한 지식 공유 서비스의 개념을 소개한다. 사용자들은 자신의 지식을 RIF 규칙으로 표현하고 이를 공유함으로써 서로에게 유용한 정보를 추천한다. 이러한 서비스는 마치 커다란 규칙 기반 전문가 시스템을 시맨틱 웹상에 적용하는 것이라고도 볼 수 있다. 시맨틱 웹은 사용자들이 어떻게 온톨로지를 공유하고 또한 방대한 규모의 온톨로지를 정확히 숙지하는가 등의 근본적인 문제점을 가지지만, 본 논문에서는 제안하는 서비스에서 필연적으로 발생하는 불확실성의 문제를 자세히 분석한다. 특정 문제 영역을 고려하는 기존 전문가 시스템과는 달리 다양한 문제 영역 사이에서의 지식 공유이므로 불확실성의 문제는 발생할 수밖에 없다. 본 논문에서는 이를 자세히 분석하고 제안 서비스에서의 중요 문제점을 보인다.

### Abstract

In this paper, we introduce the concept of knowledge sharing service using Semantic Web and RIF rules. Users represent their knowledge with RIF rules and share the rules to infer information useful to each other. This service can be seen as applying a large rule-based expert system to Semantic Web. Although Semantic Web has fundamental problems such as how users can share ontologies and how users can exactly know the large ontology, we analyze in detail uncertainty problems inevitably occurring in the proposed service. Unlike existing expert systems that consider a specific problem domain, uncertainty problems arise in our service because knowledge sharing among various problem domains is considered. In this paper, we analyze it in detail and show that it is a significant problem in the proposed service.

### Keywords

semantic web, RIF rule, knowledge sharing, uncertainty problem

---

\* 서일대학교 정보통신공학과  
- ORCID: <https://orcid.org/0000-0002-0005-4506>

· Received: Jul. 04, 2018, Revised: Aug. 06, 2018, Accepted: Aug. 09, 2018  
· Corresponding Author: Jaehoon Kim  
Dept. of Information & Communication Engineering, Seoul University, 28,  
Yongmasan-ro 90 gil, Jungnang-gu, Seoul, 02192, Korea,  
Tel.: +82-2-490-7258, Email: [jhkimykg@gmail.com](mailto:jhkimykg@gmail.com)

## 1. 서 론

시맨틱 웹(Semantic Web)은 HTML 언어로 기술된 웹 문서의 내용에 RDF, OWL 온톨로지 언어를 이용하여 의미적 정보를 추가 기술함으로써 이를 이용한 보다 정확한 정보검색을 가능하도록 하려는 것이다. 이에 대한 현실적인 방법으로 링크드 데이터(Linked Data)[1]는 URI에 RDF로 표현한 의미 정보를 추가 부여함으로써 기계 스스로 관련 정보를 탐색해 나아갈 수 있도록 한다.

본 논문에서는 시맨틱 웹을 활용한 다른 응용 사례로서 사용자들이 자신의 지식을 온톨로지를 이용하여 표현된 규칙을 이용하여 서로 공유하는 서비스를 고려해 보고자 한다. 이러한 서비스에서 사용자들은 자신의 지식을 IF THEN의 규칙으로 표현하여 웹에 공개하고, 이러한 규칙은 다른 사용자에게 공유되어 유용한 정보를 추천하게 된다. 예로, Park이라는 사용자가 어떤 음식이 특정 질병에 안 좋다는 지식을 규칙으로 표현하여 웹상에 등록할 경우 해당 질병을 갖고 있는 Lee라는 사용자가 어떤 음식점을 방문할 경우 해당 규칙의 추천에 의하여 그 음식은 주문하지 말 것을 알려준다. 이러한 규칙은 상호간의 의미 해석을 위하여 시맨틱 웹에 기반한 온톨로지 언어로 기술된다.

이러한 서비스에서 어떻게 모든 사용자들이 온톨로지를 공유해 나갈 수 있으며 또한 방대한 규모의 온톨로지를 숙지하여 관련 규칙을 작성해 나갈 수 있는가 등의 시맨틱 웹의 근본적인 문제가 존재하지만, 본 논문에서는 제안 서비스에서 필연적으로 발생할 수밖에 없는 불확실성의 문제(Uncertainty Problem)에 대하여 자세히 살펴보고자 한다. 본 논문에서 고려하는 서비스는 마치 커다란 규모의 규칙 기반 전문가 시스템을 시맨틱 웹상에 적용하는 것이라고도 볼 수 있으며, 따라서 규칙 생성에 있어서의 불확실성 문제가 존재하게 된다. 기존 개념과의 차이는 특정 문제 영역이 아닌 다양한 문제 영역 사이에서의 자유로운 규칙 생성이므로 불확실성 문제가 필연적으로 발생할 수밖에 없다는 것이다.

본 논문은 다음과 같이 구성된다. 2장에서는 본 논문의 내용이 기존 연구와 어떻게 다른지 살펴보고, 3장에서는 기반 환경인 시맨틱 웹의 OWL DL

온톨로지 언어와 RIF(Rule Interchange Format) 표준 규칙 언어[2]에 대하여 간단히 정리한다. 4장에서는 본 논문에서 고려하는 RIF 규칙을 이용한 지식 공유 서비스를 소개하며, 5장에서는 이러한 서비스에서의 불확실성 문제를 자세히 살펴본다. 6장에서는 이러한 불확실성 문제를 해결하기 위한 방법들에 대하여 살펴보고, 7장에서는 제안 서비스가 기존의 지식 공유 환경과 비교하여 어떤 의미를 갖는지를 고찰한다. 8장에서는 결론 및 향후 연구내용을 언급한다.

## II. 관련 연구

전문가 시스템과 시맨틱 웹에 관한 대부분의 연구는 참고문헌 [3]에서처럼 웹으로부터 수집된 RDF와 OWL 문서로부터 전문가 시스템을 이용하여 새로운 지식을 추천하는데 연구의 초점이 맞추어져 있다. 이러한 응용은 본 논문에서 고려하는 사용자들 사이에서의 RIF 규칙 및 온톨로지 추천을 이용한 지식 공유 서비스와는 다르다.

시맨틱 웹의 규칙 언어인 SWRL(Semantic Web Rule Language)[4] 혹은 RIF를 활용한 대부분의 연구는 특정 문제 영역에서의 지식 추천 활용에 초점이 맞추어져 있다. 예로, 참고 문헌 [5]에서는 의료 문제 영역에서의 온톨로지 및 SWRL 규칙 언어를 이용한 진단 추천에 대한 연구이다. 본 논문에서는 특정 문제 영역뿐만이 아닌 다양한 문제 영역 사이에서의 지식 공유 및 추천을 살펴본다. 이러한 다양한 문제 영역 사이에서의 자유로운 지식 공유의 형성은 불확실성의 문제를 필연적으로 발생시킨다. 만약 특정 문제 영역만을 고려한 SWRL과 RIF를 이용한 전문가 시스템의 구축이라면 이러한 불확실성 문제는 기존 방법대로 시스템 설계 및 개발 과정에서 해결되어질 것이다.

시맨틱 웹의 표준 규칙 언어인 RIF는 그 목적이 규칙 언어로서의 의미보다는 기존에 존재하는 다양한 규칙 언어와의 상호 운용성(Interoperability)를 위하여 디자인 되었다. 즉 기존의 규칙 언어인 Drools, Jess, CLIPS 등으로 작성된 이질적 규칙들은 시맨틱 웹 환경에서 RIF로 변환되어 함께 추천될 수 있다.

본 논문에서는 이러한 RIF를, 다른 응용 사례로

서, 시맨틱 웹 환경에서의 지식 공유를 위한 목적으로 활용해 보고자 한다.

자료 조사를 통하여 시맨틱 웹과 RIF를 이용한 지식 공유 서비스에서의 불확실성 문제에 관한 연구는 찾아 볼 수 없었으며, 본 논문에서는 기존 전문가 시스템에서의 불확실성 문제 분석[6]을 토대로 이를 체계적으로 분석한다.

### III. OWL DL과 RIF

#### 3.1 OWL DL

시맨틱 웹의 온톨로지 언어는 RDF[7]와 OWL[8]이 있으며, OWL은 다시 OWL Lite, OWL DL, OWL Full로 나누어진다. RDF는 subClassOf, subPropertyOf의 클래스, 속성간의 기본적인 포함 관계 (Subsumption Relationship)를 정의할 수 있으며, OWL Lite는 이러한 RDF 문법 외에 클래스, 속성간의 동등(Equivalent) 관계, 속성 제약(Property Restriction), 버전 정보(Versioning), 설명(Annotation) 등의 어휘를 추가한다. OWL DL은 OWL Lite의 어휘를 포함하며 서술 논리(Description Logic) 수준의 추론을 지원한다. OWL Full은 가장 많은 온톨로지 표현을 위한 문법을 갖지만 비결정적(Undecidable)이며 완전한 추론을 보장할 수 없다. 따라서 본 논문에서는 OWL DL 수준의 온톨로지 언어를 고려하도록 한다.

#### 3.2 RIF

지식 공유를 위하여 사용할 수 있는 기존의 다양한 규칙 언어가 존재하지만, 본 논문에서는 기반 환경인 시맨틱 웹을 고려하여 표준 규칙 언어인 RIF를 사용하도록 한다.

RIF는 Datalog[9]와 유사하며, horn 문절(Horn Clause) 형식을 갖는다. RDF와 OWL 온톨로지과 결합될 경우 규칙의 각 항목은 정의 1의 RDF 트리플 패턴으로 표현될 수 있다.

- 정의 1 (RDF 트리플 패턴)

그림 1과 같이 RDF 그래프  $G = (V, E)$ 를 고려하자.  $v, u \in V$ 는 노드를 나타내고,  $e \in E$ 는  $v$ 와  $u$  사이

의 에지를 나타낸다. RDF 트리플은  $G$ 에서의  $[v, e, u]$ 를 의미한다.  $v, e, u$ 에 대해 각각 임의의 노드와 에지를 매칭한다는 의미로  $\$x, \$y, \$z$ 를 사용할 수 있다.

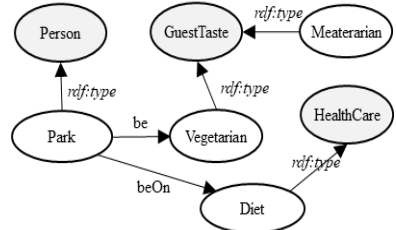


그림 1. RDF 데이터의 그래프 표현  
Fig. 1. Graph representation of RDF data

RIF 규칙의 예로 그림 4에서의 규칙 1을 살펴보자:  $[\$x, \text{be}, \text{Vegetarian}] \rightarrow [\$x, \text{like}, \text{Salad}]$ . 규칙 1은  $\$x$ 가 채식주의자이면 샐러드를 좋아한다는 정보를 추론하는 RIF 규칙으로 그림 1에서 Park은 Salad를 좋아한다는 사실을 추론할 수 있다.

- 정의 2 (RIF 규칙에서의 증거와 추론 결과)

본 논문에서는 RIF 규칙에서 왼편의 조건에 해당하는 각 RDF 트리플 항목을 증거(Evidence)라고 부르며, 오른편 RDF 트리플 항목을 추론 결과라고 부르도록 하겠다.

### IV. RIF 규칙을 이용한 지식 공유 서비스

본 논문에서 고려하는 RIF 규칙을 이용한 지식 공유 서비스는 다음과 같다. 먼저 관련 연구에서도 언급하였지만 제안하는 서비스가 기존 연구와 다른 점은 기존 연구는 특정 문제 영역에서의 지식 추론이지만 본 연구는 다양한 문제 영역 사이에서의 지식 공유 및 추론을 고려한다는 점이다.

- 정의 3 (문제 영역)

문제 영역(Problem Domain)은 전문가 시스템에서 전문가의 지식이 적용되는 영역을 말하며 본 논문에서도 규칙 기반의 지식 공유 서비스가 적용되는 영역을 의미한다.

4 RIF 규칙을 이용한 지식 공유 서비스에서의 불확실성 문제 분석

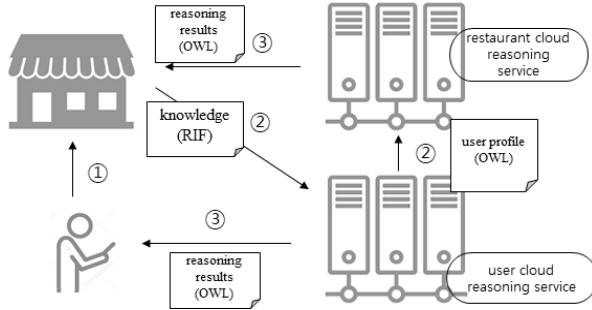


그림 2. RIF 규칙을 이용한 지식 공유의 개념  
Fig. 2. Concept of knowledge sharing using RIF rules

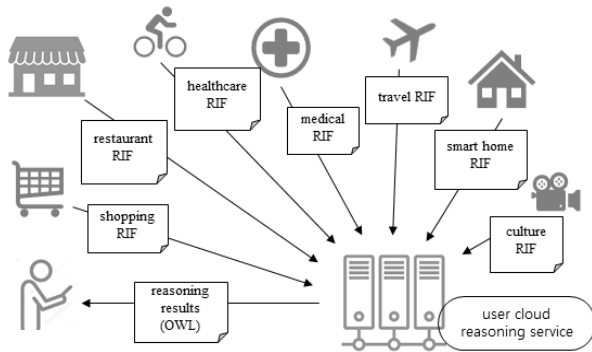


그림 3. 다양한 문제 영역 사이에서의 지식 공유  
Fig. 3. Knowledge sharing among various problem domains

Al-Hamadani와 Alwan[5]는 심혈관 진단 문제 영역에서의 온톨로지 및 규칙 언어를 이용한 진단 추론에 대한 연구를 수행하였다. 하지만 그림 2에서는 음식점 문제 영역과 개인 건강관리 문제 영역 사이의 지식 공유 서비스를 보여준다. 그림 3은 보다 다양한 문제 영역 사이에서의 지식 공유 및 추론 서비스가 이루어 질 수 있음을 보여준다.

지식 공유 서비스의 소개에 앞서 어떻게 모든 사용자가 온톨로지를 공유해 나갈 수 있으며, 또한 그러한 방대한 규모의 온톨로지를 숙지하여 관련 규칙을 작성해 나갈 수 있는가에 관한 두 가지 주요한 가정을 둔다.

▪ 가정 1 (온톨로지 공유)

각 문제 영역에 대한 표준화된 온톨로지의 구축이 활발하게 이루어지며 이것은 웹을 통하여 공개되어 진다. 이렇게 널리 알려진 온톨로지를 이용하여 사용자들은 RIF 규칙을 만들거나 새로운 영역의 온톨로지를 구축할 수 있다.

예로, 의료 문제 영역의 GALEN, Med, SNOMED CT, Foundational Model of Anatomy, UMLS와 같은 기존의 널리 알려진 온톨로지를 찾아볼 수 있으며 [10], 이외에도 다양한 문제 영역의 온톨로지 구축을 위한 노력을 찾아 볼 수 있다[11]-[13]. 그림 3에서 스마트폰 온톨로지와 의료 온톨로지를 이용하여 새로운 RIF 규칙을 등록할 수도 있고, 필요한 경우

스마트홈 온톨로지와 의료 온톨로지 사이의 새로운 온톨로지를 만들어 웹에 공개할 수도 있다.

▪ 가정 2 (온톨로지 엔지니어)

각 문제 영역별 온톨로지를 구축하고 공개된 온톨로지를 이용하여 새로운 영역의 온톨로지를 구축하거나 RIF 규칙을 생성하는 사용자는 훈련된 온톨로지 엔지니어이어야 한다.

온톨로지 및 RIF 지식 구축에 있어 일반 사용자들은 온톨로지의 문법 이해 및 효율적 표현에 있어 어려움이 있다. 따라서 훈련된 온톨로지 엔지니어에 의하여 본 논문에서 고려하는 지식 공유 서비스가 구축 확장되어 나아감을 가정한다.

그림 2는 RIF 규칙을 이용한 지식 공유의 간단한 개념을 보여준다. 이러한 개념은 실제 그림 3과 같이 더욱 다양한 문제 영역 사이에서 발생한다.

음식점 문제 영역의 RIF 지식은 그림 4(a)와 같다. 그림 5(a)와 5(b)는 해당 RIF 지식이 사용한 웹상에 공개된 음식점 및 의료 온톨로지의 일부이다.

그림 2에서 ①사용자가 음식점에 들어설 경우 음식점에서 일어날 수 있는 지식을 정의한 ②RIF 규칙이 사용자 추론 서버로 전송된다. ③동시에 클라우드 상의 사용자 추론 서버로부터 일반에게 공개가 허용된 사용자 프로파일의 일부분이 음식점 추론 서버로 전송된다.

그림 5(c)는 공개가 허용된 사용자의 프로파일의 일부를 보여주며 이 또한 OWL 언어로 기술된다.

③양쪽의 추론 서버에서는 제공받은 프로파일 혹은 규칙을 기준에 저장하고 있는 사실 및 RIF 규칙과 함께 추론하여 새로운 사실을 알아내고 이를 사용자와 음식점 내의 스마트 기기로 전송한다. 스마트 기기에 전송되는 정보 또한 OWL 언어로 표현되며 이러한 정보는 에이전트 프로그램에 의해 해석되어 처리된다.

예로 그림 5(c)의 프로파일이 음식점 추론 서버로 전송될 경우 Park이라는 고객이 다이어트 중이며 채식주의자이므로 규칙 2에 의해 닭가슴살 샐러드를 추천할 수 있다. 또한 Park의 추론 서비스에서는 Park이 당뇨병을 앓고 있으므로 규칙 3에 의해 닭고기는 추천할 수 없도록 할 수 있다.

```

규칙 1:
[$x, be, Vegetarian] → [$x, like, Salad]

규칙 2:
[$x, beOn, Diet] ∧ [$x, be, Vegetarian] → [$x, like, ChickenBreastSalad]
    
```

(a) 음식점 RIF  
(a) Restaurant RIF

```

규칙 3:
[$x, have, Diabetes] → [$x, dislike, SugaryFood]

규칙 4:
[$x, have, RenalDisease] ∧ [$x, eat, Vegetables] → [Vegetables, beBadFor, $x]
    
```

(b) 사용자 RIF  
(b) User RIF  
그림 4. RIF 규칙의 예  
Fig. 4. Sample RIF rules

```

<rdf:Description rdf:ID = "Salad">
<rdfs:subClassOf rdf:resource = "Food"/>
</rdf:Description>

...

<rdf:Description rdf:ID = "Meatearian">
<rdf:type rdf:resource = "GuestTaste"/>
</rdf:Description>

<rdf:Description rdf:ID = "Vegetarian">
<rdf:type rdf:resource = "GuestTaste"/>
</rdf:Description>
    
```

(a) 음식점 온톨로지  
(a) Restaurant ontology

```

<rdf:Description rdf:ID = "Diabets">
<rdfs:subClassOf rdf:resource = "Disease"/>
</rdf:Description>

<rdf:Description rdf:ID = "RenalDisease">
<rdfs:subClassOf rdf:resource = "Disease"/>
</rdf:Description>

...

<rdf:Description rdf:ID = "Diet">
<rdf:type rdf:resource = "HealthCare"/>
</rdf:Description>
    
```

(b) 의료 온톨로지  
(b) Medical ontology

```

<Person rdf:ID="Park">
<be rdf:resource="Vegetarian"/>
<beOn>Diet</beOn>
</Person>
    
```

(c) 사용자 프로파일  
(c) User profile  
그림 5. 온톨로지 데이터  
Fig. 5. Ontology data

## 6 RIF 규칙을 이용한 지식 공유 서비스에서의 불확실성 문제 분석

사용자 프로파일을 포함한 사실 정보는 다음과 같은 다양한 경로를 통하여 수집될 수 있으며 OWL 온톨로지 언어로 기술된다.

- 사용자와 관련한 명백한 사실 정보를 제공할 수 있는 기관으로부터의 수집: 예로 병원 진료 기록 및 스마트 헬스케어 관련 개인 건강 기록 정보
- 사용자가 직접 작성한 프로파일 정보: 사용자는 자신의 추론 서비스에 자신이 채식주의자임을 직접 입력할 수 있음. 음식점의 경우도 자신의 추론 서비스에 음식 관련 정보를 직접 입력
- 센터로부터의 수집 정보: 예로 날씨, 온도에 따른 적당한 음식 메뉴의 추천
- RIF 추론 결과: RIF 추론 결과로부터 의미 있는 사실 정보를 저장. 예로, 규칙 4의 추론 결과 [Vegetables, beBadFor, Park]를 사용자 프로파일 정보에 추가 저장

### V. 불확실성 문제 분석

과거 규칙 기반 전문가 시스템[6]에서 체계적으로 정립된 불확실성 문제는 다음과 같으며 본 논문의 제안 서비스에서도 동일하게 적용된다.

- 개별 규칙에서의 오류
- 규칙의 양립성에 의한 오류
  - 모순, 포함 관계, 중복, 누락

이러한 오류는 기존의 경우 시스템 설계 및 개발 단계에서 해당 문제를 발견하고 수정해 나가는 과정을 거친다. 하지만 본 논문에서 고려하는 다양한 문제 영역 사이에서의 규칙 오류는 수정될 수 있는 문제가 아니며 필연적으로 발생할 수밖에 없다. 중복의 경우는 기존과 달리 본 논문에서 오히려 충돌 해결을 위한 암시적 우선순위 값을 제공해 줄 수 있으며, 또한 온톨로지 및 규칙 수정에 의한 불확실성 문제가 존재한다.

#### ▪ 정리 1 (여러 문제 영역 사이의 지식 표현)

본 논문에서 고려하는 지식 공유 서비스는 하나의 문제 영역 뿐만이 아닌 여러 문제 영역에 걸친 지식을 표현하게 된다.

예로, 의료 문제 영역에서의 [Park, have, Renal Disease] 사실과 음식점 문제 영역에서의 [park, eat, Vegetables] 사실은 두 문제 영역 사이에서의 지식  $[Sx, have, RenalDisease] \wedge [Sx, eat, Vegetables] \rightarrow [Vegetables, beBadFor, Sx]$ 에 의해 해당 신부전 환자에게 날것의 채소를 먹는 것은 해롭다는 사실을 알려줄 수 있다.

#### ▪ 정리 2 (규칙의 세밀함)

본 논문에서 고려하는 지식 공유 서비스에서 규칙의 세밀함은 제어될 수 없다.

기존 규칙 기반 시스템을 설계할 경우 규칙의 세밀함은 중요한 요소이다. 만약 너무 큰 세밀함은 여러 지식이 하나의 규칙에 함축되어 있으므로 수정하기 어렵고, 만약 너무 작은 세밀함은 다른 규칙들을 참조할 없이 이해하기란 어려우므로 복잡할 것이다. 단일 문제 영역에서의 기존 규칙 기반 시스템은 이러한 세밀함을 신중히 고려하여 설계한다. 하지만 본 논문의 경우 여러 문제 영역에 걸쳐 불특정 다수의 사용자들에 의해 임의로 생성되는 규칙의 세밀함은 제어되기 쉽지 않다.

#### ▪ 정리 3 (단조 및 비단조 추론)

본 논문의 지식 공유 서비스에서는 단조 추론(Monotonic Reasoning) 및 비단조 추론(Nonmonotonic Reasoning)이 함께 고려된다.

그림 2에서 음식점의 RIF 규칙에 따라 사용자의 프로파일이 변경되었음을 가정하자. 그리고 이 RIF 규칙에 오류가 있어 이후 RIF 규칙이 다시 작성됨을 가정하자. 이러한 경우 이전의 RIF 규칙에 의해 변경된 프로파일을 삭제하고 다시 수정된 RIF 규칙을 적용하여 프로파일을 갱신하는 것은 어려운 일이다. 왜냐하면 음식점의 추론 서비스가 그 음식점을 이용한 모든 사용자에게 RIF 규칙이 변경되었음을 진송할 수는 없는 일이기 때문이다. 반대로 그 음식점을 이용한 사용자가 주기적으로 그 음식점의 RIF 규칙이 변경됨을 모니터링할 수도 없는 문제이기 때문이다. 만약 비단조 추론을 고려한다면 사용자는 음식점뿐만 아니라 자신이 이용한 모든 장소의 RIF 규칙의 변경됨을 주기적으로 모니터링 해야

하며 이것은 무척 큰 시스템 자원의 소모 및 성능 저하를 가져오기 때문이다. 다른 방법으로 사용자가 그 음식점을 다시 방문할 때 RIF 규칙의 변경 여부를 확인할 수 있지만, 사용자가 그 음식점을 언제 방문할지 모르는 상황에서 이전 RIF 규칙에 의한 추론 과정을 모두 기록하는 것은 불가능한 일이기 때문이다. 그 기간이 1년 혹은 그 이상 일 수도 있다. 이 기간 동안 다른 규칙에 의해 갱신된 프로파일은 1년이 지나서 다시 갱신하는 것 또한 어려운 문제일 것이다. 따라서 본 논문의 지식 공유 서비스에서는 단조 추론이 고려되어야 한다.

하지만 사용자의 프로파일의 위의 예에서처럼 음식점 RIF 규칙의 액션에 의하여 변경되는 일은 흔치 않다. 대부분은 신뢰할 만한 기관으로부터의 RIF 규칙에 의하여 변경될 것이다. 예로, 사용자 질병 및 건강 기록 정보는 의료 기관의 RIF 규칙에 의하여 변경될 것이다. 따라서 이러한 경우 비단조 추론 시스템이 고려될 수 있으며, 사용자의 추론 서비스는 해당 온톨로지 및 RIF 규칙의 변경을 주기적으로 모니터링하여 이를 반영할 수 있다. 이러한 신뢰할 만한 기관의 수는 앞서의 사용자가 방문한 모든 장소의 수만큼 크지 않을 것이기 때문이다.

정리 1과 2, 그리고 정리 3의 단조 추론은 본 논문에서 고려하는 지식 공유 서비스에서 불확실성의 문제가 필연적으로 발생할 수밖에 없는 이유를 설명해 준다. 이러한 이유로 개별 규칙에서의 오류 그리고 여러 규칙 사이에서의 포함 관계, 중복 관계, 모순 관계에 의한 오류들이 쉽게 발생할 수 있다.

▪ 정리 4 (개별 규칙에서의 불확실성)

그림 6과 같은 개별 규칙의 오류에 의한 불확실성의 문제가 존재한다.

그림 6과 같이 기존 전문가 시스템에서와 같은 개별 규칙에서의 불확실성의 문제가 존재한다[6]. 모호한 오류(Ambiguous Errors)는 규칙이 하나 이상

의 방법으로 해석될 수 있는 경우이고, 불완전 오류(Incomplete Errors)는 몇 가지 증거 정보가 빠져있는 경우이고, 부정확한 오류(Incorrect Errors)는 규칙의 정보가 틀린 경우이고, 측정 오류(Measurement Errors)는 규칙에 사용된 값이 잘못 측정된 경우이다. 랜덤 오류(Random Errors)는 규칙이 랜덤 특성을 갖는 정보를 사용하기 때문에 발생하는 오류이며, 시스템 오류(Systematic Errors)는 시스템의 편향(Bias)에 의해 발생하는 오류이다. 만약 규칙에 확신과 불확신의 척도 값을 적용할 경우 이러한 측정값은 인간의 주관적 견해이므로 증거의 가능성(Likelihood of Evidence) 오류가 존재한다. 또한 규칙의 각 증거를 연결할 시(Combining Evidence)의 AND, OR, NOT의 연결 오류가 존재할 수 있다.

이러한 규칙의 개별적 오류는 발생될 수 있고 이러한 틀린 규칙의 결과에 따른 불확실성 문제는 쉽게 전파될 수 있다. 또한 규칙을 수정할 경우 정리 3에서의 단조 추론에 따라 이를 수정한 규칙과 이전 규칙이 공존하므로 불확실성의 문제가 발생한다.

▪ 정리 5 (여러 규칙들 사이의 불확실성)

그림 7과 같은 여러 규칙들 사이에서의 불확실성 문제가 존재한다.

본 논문에서 고려하는 지식 공유 서비스의 경우 다양한 추론 서비스로부터의 동일/유사 규칙을 공유할 수 있으므로 여러 규칙 사이의 불확실성 문제가 존재하게 된다.

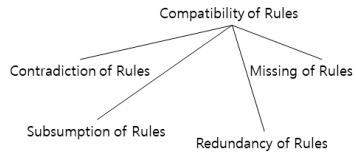


그림 7. 규칙의 양립성에 의한 오류  
Fig. 7. Errors by compatibility of rules



그림 6. 개별 규칙에서의 오류  
Fig. 6. Errors in individual rules

그림 7은 여러 규칙들 사이에서의 기존 전문가 시스템의 불확실성 문제를 보여준다[6]. 규칙의 증거는 같으나 추론 결과가 다른 모순(Contradiction of Rules)이 존재할 수 있으며, 그림 4에서의 규칙 1과 2와 같이 규칙사이의 포함 관계(Subsumption of Rules)에 의한 불확실성의 문제가 존재할 수 있다. [Park, beOn, Diet]라던 규칙 1을 추론하겠지만, 만약 [Park, beOn, Diet]와 [Park, be, Vegetarian]이라던 규칙 1과 2를 동시에 추론할 것이다. 따라서 어떤 결과를 적용해야 할지에 대한 충돌 해결의 방법이 필요하다.

규칙의 중복은 같은 규칙이 중복되는 경우이고, 규칙의 누락은 규칙의 집합에서 존재해야 할 규칙이 누락된 경우이다. 규칙의 중복은 기존 규칙 기반 시스템과 달리 본 연구에서는 충돌 해결을 위한 효과적인 방법으로 사용될 수 있다. 이것은 다음 장에서 설명하도록 한다. 규칙의 누락에 의해 필요한 추론이 이루어 질 수 없으므로 이로 인한 불확실성의 문제를 발생시킨다.

기존 규칙 기반 시스템은 단일 문제영역의 시스템을 구축하는 것이므로 이러한 불확실성의 문제를 설계 및 개발 단계에서 해결할 수 있지만, 여러 문제 영역 사이의 대규모의 지식 공유 서비스를 구축하는 경우 이러한 불확실성의 문제는 필연적으로 발생하고 이를 해결하기 위한 방법이 필요하다.

다음 장에서는 이러한 불확실성의 문제를 해결하기 위한 충돌 해결의 방법들을 살펴보기로 한다.

▪ 정리 6 (온톨로지의 변경)

기존 온톨로지가 변경될 경우 이를 사용하는 사실 정보 및 규칙의 변경이 고려되어야 하며, 변경이

반영되지 못할 경우 이로 인한 불확실성의 문제가 존재한다.

정리 3의 비단조 추론에 의하여 기존 온톨로지의 변경이 있을 경우 이로 인한 사실 정보 및 규칙의 변경, 그리고 RIF 추론이 제 수행되어 반영되어야 한다. 하지만 이러한 변경이 반영되지 못하면 불확실성의 문제가 발생하게 된다.

VI. 불확실성의 고려 및 충돌 해결 방법

앞장에서는 제안하는 지식 공유 서비스에서 불확실성의 문제가 필연적으로 발생함을 설명하였다. 또한 이러한 문제는 다양한 문제 영역 사이에서의 임의적 규칙 생성이므로 단일 문제 영역에서의 규칙 기반 시스템 개발과는 달리 통제될 수 없는 성격의 문제임을 설명하였다.

기존 규칙 기반 전문가 시스템에서는 이러한 불확실성 문제를 위해 그림 8과 같이 다양한 충돌 해결(Conflict Resolution) 방법들이 깊이 있게 연구되었으며[6], 본 서비스에서도 이러한 방법들을 활용할 수 있다.

먼저 명시적 우선순위 값을 부여하는 방법(Explicit Priority of Rules)을 고려할 수 있다. 하지만 이 방법은 다양한 환경에서 다수의 사용자가 규칙을 생성하므로 통일된 기준에 의한 우선순위 값을 부여하는 것이 불가능하므로 실현하기 어려운 방법이다. 다음으로 암시적 우선순위 값(Implicit Priority of Rules)으로 패턴의 특이성(Speciality of Patterns)을 고려할 수 있다. 예로, 보다 많은 증거를 갖는 규칙에 의한 추론 결과를 고려하는 것이다.

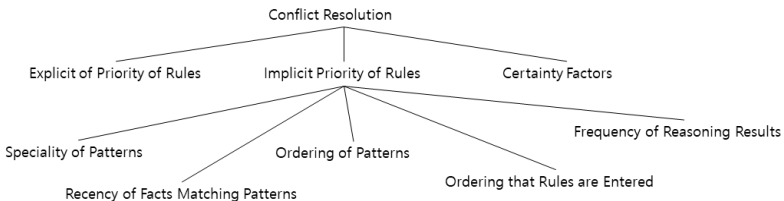


그림 8. 충돌 해결 방법  
Fig. 8. Conflict resolution methods



다음으로 패턴과 매칭되는 최근 사실의 중요도 (Recency of Facts Matching Patterns)를 고려할 수 있으며, 사전 순서(Lexicographic Order)와 같은 패턴 순서(Ordering of Patterns)를 고려할 수 있다. 또한 입력된 패턴의 순서(Ordering that Rules are Entered)에 따라 최근에 입력된 규칙을 먼저 적용하는 방법을 고려할 수 있다.

5장에서 규칙의 중복 및 동일 결과를 추론하는 규칙에 의한 추론 결과의 빈도수(Frequency of Reasoning Results)는 해당 규칙의 중요도를 관별하는 주요 기준이 될 수 있다. 예로, 아래와 같이 사용자의 추론 서비스에 등록된 중복 규칙에 대하여 [Vegetables, beBadFor, \$x] 결과를 갖는 규칙의 빈도수를 고려하여 해당 결과를 채택할 수 있다.

[ $\$x$ , have, RenalDisease]  $\wedge$  [ $\$x$ , eat, Vegetables]  $\rightarrow$  [Vegetables, beBadFor,  $\$x$ ]  
 [ $\$x$ , have, RenalDisease]  $\wedge$  [ $\$x$ , eat, Vegetables]  $\rightarrow$  [Vegetables, beBadFor,  $\$x$ ]  
 [ $\$x$ , have, RenalDisease]  $\wedge$  [ $\$x$ , eat, Vegetables]  $\rightarrow$  [Vegetables, beGoodFor,  $\$x$ ]

마지막으로 확률 값(Probability) 대신 사람이 자신의 주관적 판단 수치를 규칙에 부여하는 확신도 (Certainty Factors)의 값을 고려할 수 있다. 단일 문제 영역이 아닌 본 논문의 경우 정확한 확률 값을 측정하는 것은 불가능하지만, 1~10 사이의 척도와 같은 일반적인 사람들의 확신/불확신의 주관적 수치는 객관적으로 고려될 수 있다. 왜냐하면 그림 6에서 증거의 가능성 오류(Likelihood of Evidence)가 존재하지만 다수를 고려할 경우 전체의 평균 기대치를 예상할 수 있기 때문이다.

MYCIN[6][14]에서는 표 1과 같이 하나의 규칙에서의 증거  $E$ 가 AND, OR, NOT으로 연결될 경우 확신도의 값을 정의하고 있다. 예로, 그림 4의 규칙 2에 대하여 [Park, beOn, Diet]의  $CF$  값이 6이고 [Park, be, Vegetarian]의  $CF$  값이 9인 경우 추론 결과 [Park, like, ChickenBreastSalad]의  $CF$  값은 6이다.

[Park, beOn, Diet]  $\wedge$  [ $\$x$ , be, Vegetarian]  $\rightarrow$  [Park, like, ChickenBreastSalad]

또한 같은 추론 결과를 가져오는 두 규칙이 다른 확신도 값을 가질 경우 확신도 값의 계산은 표 2와 같다. 양쪽  $CF$  값이 모두 0보다 큰 경우 첫 번째 조건을 적용하며, 한쪽의  $CF$  값이 0보다 작은 경우 두 번째 조건을 적용하며, 모두 0보다 작은 경우는 세 번째 조건을 적용한다.

만약 두 개 이상의 확신도 값을 조합하여야 하는 경우에는 먼저 두 개의 조합 결과 값을 계산한 후 이를 다시 표 2에 의해 세 번째  $CF$ 와 조합하여 계산한다.

표 1. MYCIN에서의 추론 결과의 확신도 값 계산  
 Table 1. Certainty factor calculation of reasoning results in MYCIN

Evidence, $E$	Certainty Factor, $CF$
$E_1$ AND $E_2$	$\min(CF(E_1), CF(E_2))$
$E_1$ OR $E_2$	$\max(CF(E_1), CF(E_2))$
NOT $E$	$-CF(E)$

표 2. 동일 추론 결과가 다른 확신도 값을 가질 경우의 조합 계산  
 Table 2. Combination calculation when the same reasoning results have different certainty factors

When	$CF_{combine}(CF_1, CF_2)$
both $CF > 0$	$CF_1 + CF_2(1 - CF_1)$
one $CF < 0$	$\frac{CF_1 + CF_2}{1 - \min( CF_1 ,  CF_2 )}$
both $CF < 0$	$CF_1 + CF_2(1 + CF_1)$

## VII. 시맨틱 웹 및 RIF 규칙을 이용한 지식 공유 서비스에 대한 고찰

기존의 지식 공유를 위한 방법으로 웹을 통한 정보검색 및 SNS 서비스를 살펴볼 수 있다. 하지만 본 논문에서 제안하는 지식 공유는 규칙을 이용한 정보 공유이다. 사용자간의 상호 의미 해석을 위하여 시맨틱 웹의 온톨로지 언어인 OWL DL과 표준 규칙 언어인 RIF를 활용한다. 이러한 온톨로지 언어 및 규칙에 의한 정보 공유는 앞서 그림 3과 같은 유비쿼터스 컴퓨팅 환경과 아래의 스마트 기기에서의 에이전트와의 대화를 통한 사용자 인터페이스에 있어 적합한 방법이라고 사료된다.

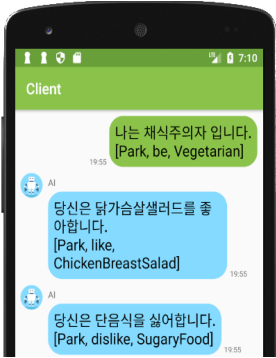


그림 9. 스마트 기기에 표현된 OWL 추론 결과  
Fig. 9. OWL reasoning results displayed on a smart device

제안하는 지식 공유를 위하여 기존의 방법대로 특정 목적의 웹 애플리케이션 혹은 앱(App)을 개발하는 것을 고려할 수 있다. 하지만 이러한 방법은 다양한 문제 영역에서 각각의 개발이 이루어질 경우 저마다 다른 방식의 데이터 모델과 프로그래밍 도구로 개발하는 노력과 비용이 필요하다. 만약 이러한 각각의 문제 영역에서 저마다 개발된 지식 공유 서비스를 서로 연결하려고 한다면 이 또한 어려운 문제가 된다.

따라서 다양한 문제 영역 사이의 서로 연결된 지식 공유를 위하여 시맨틱 웹을 이용하고 RIF 규칙 언어를 사용하는 방법은 의미 있다고 볼 수 있다. 이러한 서비스가 다양한 문제 영역에서 구축되고 서로 연결될 경우 우리의 일상생활 속에서의 지능적인 서비스를 제공하는 또 하나의 방법이 될 것으로 사료된다.

### Ⅷ. 결론 및 향후 연구내용

본 논문에서는 시맨틱 웹 및 RIF 규칙을 이용한 사용자 상호간의 지식 공유 서비스의 개념에 대하여 소개하였다. 이것은 마치 시맨틱 웹에 커다란 규모의 규칙 기반 전문가 시스템을 구축해 나아가는 것으로 비유될 수 있다. 이러한 경우 임의적 규칙 작성에 의한 규칙의 오류, 모순, 포함 관계, 누락,

변경에 따른 불확실성의 문제가 필연적으로 발생하게 된다.

본 논문에서는 참고문헌[6]의 기존의 체계적으로 정립된 불확실성의 오류를 토대로 이에 해당 하는 사례를 분석하였으며, 또한 이에 해당하지 않는 본 연구에 특화된 불확실성의 경우를 분석 정리하였다. 하지만 시맨틱 웹의 근본적인 어려움인 온톨로지 구축 및 공유로 인하여 실제 다양한 문제 영역 사이의 RIF 규칙을 이용한 지식 공유 서비스를 충분히 구현 실험하지는 못하였다. 현재까지는 사용자 프로파일을 MySQL에 RDF 트리플 형태로 저장하고 이를 Drools[15]를 이용하여 규칙 추론하고 추론 결과를 간단히 구현한 메시지 앱에 전송하여 보여줌으로서 제안하는 RIF 규칙을 이용한 지식 공유 서비스의 기능을 확인하였다. 또한 사례 분석을 통하여 이러한 서비스의 구축에 있어 불확실성의 문제가 중요한 문제임을 확인하였다.

앞으로의 계획은 본 논문의 분석을 토대로 다양한 영역 사이의 지식 공유 서비스를 실제 구축해 나아가는 것이며, 이를 통해 불확실성 문제를 실제 검증하고 제안하는 지식 공유 서비스의 실현 가능성을 확인하는 것이다.

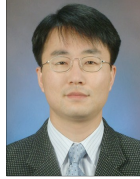
### References

- [1] Linked Data, <http://www.w3.org/standards/semanticweb/data>, [accessed: Jul. 02, 2018]
- [2] Rule Interchange Format Current Status, [http://www.w3.org/standards/techs/rif#w3c\\_all](http://www.w3.org/standards/techs/rif#w3c_all). [accessed: Jul. 02, 2018]
- [3] O. Verhodubs and J. Grundspenkis, "Towards the Semantic Web Expert System", Applied Computer Systems, Vol. 44, No. 1, pp. 116-123, Jan. 2012.
- [4] SWRL: A Semantic Web Rule Language Combining OWL and RuleML, <http://www.w3.org/Submission/SWRL/>. [accessed: Jul. 02, 2018]
- [5] B. T. Al-Hamadani and R. F. Alwan, "An Ontology-Based Expert System for General Practitioners to Diagnose Cardiovascular Diseases", Advances in Computational Sciences and Technology, Vol. 8, No. 1, pp. 53-65, Aug. 2015.

- [6] J. G. Giarratano and G. D. Riley, "Expert Systems Principles and Programming", Thomson Course Technology, 2005.
- [7] Resource Description Framework (RDF), <https://www.w3.org/RDF/>. [accessed: Jul. 02, 2018]
- [8] Web Ontology Language (OWL), <https://www.w3.org/OWL/>. [accessed: July. 02, 2018]
- [9] S. Ceri, G. Gottlob, and L. Tanca, "What you Always Wanted to Know About Datalog (And Never Dared to Ask)", IEEE Transactions on Knowledge and Data Engineering, Vol. 1, No. 1, pp. 146-166, Mar. 1989.
- [10] R. A. Greenes, "Clinical Decision Support: The Road Ahead", Elsevier, Dec. 2006.
- [11] Ontologies used for oneM2M, <http://www.onem2m.org/technical/onem2m-ontologies>, [accessed: Jul. 02, 2018]
- [12] H. Chung and J. Kim, "Design of Integrated Learning Ontology for Improving Learning Effects in e-Learning Environment", Journal of KIIT, Vol. 10, No. 8, pp. 105-116, Aug. 2012.
- [13] J. Cho, "Design of a Context Prediction Support System for U-City Life Safety", Journal of KIIT, Vol. 8, No. 8, pp. 123-131, Aug. 2010.
- [14] D. Heckerman, "Probabilistic Interpretations for Mycin's Certainty Factors", Machine Intelligence and Pattern Recognition, Vol. 4, pp. 167-196, 1986.
- [15] Drools, <https://www.drools.org/>. [accessed: Jul. 31, 2018]

## 저자소개

김 재 훈 (Jaehoon Kim)



1997년 2월 : 건국대학교

전자계산학과(공학사)

1999년 2월 : 건국대학교

컴퓨터·정보통신공학과(공학석사)

2005년 2월 : 서강대학교

컴퓨터공학과(공학박사)

2005년 3월 ~ 2006년 9월 :

삼성전자 정보통신총괄 통신연구소 책임연구원

2006년 9월 ~ 2009년 2월 : 서강대학교 컴퓨터공학과  
연구교수

2009년 3월 ~ 현재 : 서일대학교 정보통신공학과 부교수  
관심분야 : 시맨틱 웹, 데이터 프라이버시