

유사 정보 추출에 기반한 조정 에이전트 모델[†]

A Coordination Agent Model based on Extracting Similar Information

양 소 진* 이 현 숙** 오 경 환***
(So-Jin Yang) (Hyun-Suk Lee) (Kyung-Hwan Oh)

요 약 90년대 들어 각광받고 있는 에이전트 기술은, 동형 또는 이형의 에이전트가 각자 맡은 일을 해결하는 일종의 멀티 분산 시스템으로, 조정 에이전트가 여러 응용 에이전트의 통신과 제어를 담당하는 멀티 에이전트 시스템이 주류를 이룬다. 에이전트 기반 응용 기술로는 인터넷 정보 검색, 온라인 쇼핑, 메시징, 네트워크 관리등이 있는데, 포괄적으로 본다면 이러한 기술은 인터넷의 보편화에 따라 넘쳐나는 정보를 효율적으로 다루기 위한 기술이라고 볼 수 있다.

본 논문의 목적은 유사도와 강화 학습을 사용하여, 정보를 제공하는 에이전트와 정보를 요청하는 에이전트간의 연결을 매개하는 조정 에이전트(Coordination Agent, Middle Agent) 구현 방식을 제안하는데 있다. 정보 에이전트를 사용하는 멀티 에이전트 시스템에서, 정보 에이전트는 조정 에이전트에게 자신의 정보를 등록하고, 질의 에이전트는 조정 에이전트에게 원하는 질의를 요청함으로써, 서로에 대한 지식 없이도 정보의 교류가 가능하다. 그러나, 정확한 정보의 교류가 가능하기 위해서는, 정보 에이전트가 자신이 제공하는 정보를 조정 에이전트에 등록했을 때, 등록된 정보와 정확히 일치하지 않는 정보를 질의 에이전트가 요청했을 경우에 대한 적절한 대처 방안이 마련되어야 한다.

이를 위하여 본 논문에서는, 질의 에이전트의 질의와 가장 밀접한 정보를 제공하는 것으로 판단되는 정보 에이전트를 찾는 방안을 제안하고 가상의 실험공간에서 얻어진 결과에 대하여 타당성을 검증한다.

주제어 멀티에이전트시스템, 조정에이전트, 강화학습, 기본 온톨로지, 정보에이전트, 질의에이전트

Abstract Speaking generally, agent-based technology is a kind of technology to handle the flood of information resulted from the popularization of the internet. Agent system is a multi-distributed system which consists of both homogeneous and heterogeneous agents. Generally there is a coordination agent in between, which is in charge of control and message flow among the application agents.

The purpose of this thesis is to propose a coordination method among agents, some of which provide informations and some of which request them. In multi-agent system, the Information Providing Agent(IPA) registers its capabilities to Coordination Agent(CA), and the Information Requesting Agent(IRA) requests CA what it needs. To coordinate them with satisfactory results, the coordination agent ought to have an ability to return a relatively proper data to the requester, which is supposed to be similar, even though it is not so exact as was intended. For this, this thesis proposes a scheme for an coordination agent to find an IPA which provides an information which correlates most closely with that of IRA.

[†] 본 연구는 교육부 BK의 연구사업지원과 서강대학교 산업기술연구소의 부분자원에 의하여 이루어짐.
* 삼성전자 네트워크 사업부
** 동양공업전문대학 전산경영기술공학부
*** 서강대학교 컴퓨터학과

연구재부분야 : 인공지능, 에이전트시스템
주 소 : 121-742 마포구 신수동 1번지 서강대학교 컴퓨터학과
오경환 교수
전 화 : 02)705-8489 FAX : 02)704-6273
Email : kwoh@ccs.sogang.ac.kr

1. 이전 연구 결과에 대한 보완

90년대 중반 이후로 인공 지능 분야에서 에이전트에 관한 연구가 증대되고 있는데, 이것은 사용자로 하여금 넘쳐나는 지식의 홍수 속에서 좀 더 합리적으로 정보에 접근할 수 있도록 하기 위한 지능형 에이전트 간의 협력에 대한 연구의 필요성이 높아졌기 때문이다.

그러나 하나의 에이전트가 해결할 수 있는 문제의 양이나 범위에는 한계가 있고, 여러 개의 에이전트가 제한된 시간 내에 한정된 자원을 적절히 활용, 배분해야 한다는 문제가 발생한다. 이러한 문제를 해결하기 위하여 멀티 에이전트 시스템이 제안되었다(3). 멀티 에이전트 시스템은 각각의 에이전트가 전체의 균형을 유지하면서 각자의 목표를 달성하고, 그 달성된 목표들이 통합되어 전체의 목적을 이루게 된다.

조정 에이전트를 가지는 멀티 에이전트 시스템은 여러 응용 에이전트 사이에 조정 에이전트(coordination agent, middle agent)가 존재하여, 이 조정 에이전트를 통해 응용 에이전트 사이에 메시지 전달이 일어나게 된다. 멀티 에이전트 시스템에서는, 응용 에이전트는 자신의 처리 능력을 단지 조정 에이전트에게만 알리면 되고, 각각의 에이전트에 대한 제어와 통신은 조정 에이전트를 통해 이루어지게 된다.

멀티 에이전트 사회에서 정보를 탐색할 때 가장 문제가 되는 것은, 정보를 제공하는 에이전트와 정보를 요청하는 에이전트 사이에서 목적이 되는 정보가 정확히 일치하지 않을 경우에 대한 처리이다. 단지 키워드만을 등록하여 그 키워드를 대상으로 정보를 탐색하는 방법의 한계는, 등록되지 않은 키워드에 대한 요청이 접수되었을 경우에는 대처 방안이 없다는 데 있다. 이 문제에 대한 해결 방안으로, 정보를 제공하는 에이전트에 대한 선호도를 경험적 지식으로 삼아 에이전트를 선택하는 방법이 제안되었다(7). 그러나 이 방법은, 학습의 대상이 되는 정보 에이전트가 정보 제공을 중단하거나, 연결이 끊어지는 경우에는 학습 결과가 무용지물이 되는 결과를 초래한다. 본 논문은 이 문제에 대한 해결 방안으로, 정확한 내용의 정보 요청이 아니라더라도 그 정보와 유사한 개념을 가진 정보를 제공하는 정보 에이전트에게 그 정보에 대한 요청을 하는 방안을 제시하고자 한다. 즉, 요청된 정보와 유사도가 높은 개념을 가진 정보, 즉 같은 온톨로지에 속하는 정보를 제공하는 정보 에이전트에게 정보 요청을 하게 된다.

본 논문의 구성은 다음과 같다. 제 2장에서는 강화 학습과 유사도를 이용하여 정보를 찾아내는 알고리즘

을 제시한다. 3장에서는 실험을 통하여 연구의 타당성을 보이며, 4장에서는 결론과 앞으로의 연구 방향을 제시한다.

2. 유사 정보 추출을 수행하는 조정 에이전트 모델

2.1 기존 조정 에이전트 연구의 문제점

본 장에서는 유사도와 강화 학습을 기반으로 하여, 정보를 제공하는 에이전트와 정보를 요청하는 에이전트 간의 연결을 매개하는 조정 에이전트(Coordination Agent, Middle Agent)의 정보 검색 알고리즘을 제안하고자 한다. 에이전트가 방대한 인터넷 사회에서 정보를 구할 때, 구하고자 하는 정보와 정확히 일치하는 정보를 찾는다는 것은 매우 어렵고, 따라서 정확하지는 않더라도 유사하다고 판단되는 정보를 찾아내는 방안이 필요하다. 즉, 정보 에이전트가 자신이 제공하는 정보를 조정 에이전트에 등록했을 때, 등록된 정보와 정확히 일치하지 않는 정보를 질의 에이전트가 요청했을 경우에 대한 해결 방안이 마련되어야 한다.

문제 제기

이형 에이전트 사회에서 정보를 탐색할 때 가장 문제가 되는 것은, 정보를 제공하는 에이전트와 정보를 요청하는 에이전트 사이에서 목적이 되는 정보가 정확히 일치하지 않을 경우에 대한 처리이다. 단지 키워드만을 등록하여 그 키워드를 대상으로 정보를 탐색하는 방법의 한계는, 등록되지 않은 키워드에 대한 요청이 접수되었을 경우에는 대처 방안이 없다는 데 있다. 이 문제에 대한 해결 방안으로써, '경험적 지식에 의거한 조정 에이전트(7)'가 제안되었다. 일치하는 키워드가 없다는 이유로 'fail' 신호를 남발하는 경우를 막기 위하여, 질의 에이전트에 대한 경험적 지식을 이용하여 임의의 정보 에이전트에게 정보를 요청하는 것이다. 이 과정을 요약하면 다음과 같다.

- 정보 에이전트가 조정 에이전트에게 정보를 등록한다.
- 질의 에이전트는 연결된 정보 에이전트에 대한 선호도를 feedback 신호로써 조정 에이전트에 보낸다.
- 조정 에이전트는 각 정보 에이전트들에 대한 질의 에이전트의 선호도를 경험적 지식의 대상으로 삼아 강화 학습으로 학습한다.

- 질의가 계속되는 동안 강화 학습으로 경험적 지식을 갱신한다.

이 방법은 무작위로 추출하는 방법보다 성능이 우수하고, 강화 학습으로 생기는 부하는 고려 대상이 되지 않을 만큼 경미하다. 그러나 이 방법은, 연결의 대상이 되는 정보 에이전트에 대한 선호도를 학습함으로써, 학습된 정보 에이전트가 정보 제공을 중단하거나 연결이 끊어지는 경우, 그 에이전트에 대한 정보는 무용지물이 되어 질의 에이전트에게 적절한 정보를 제공하는 정보 에이전트가 사라지는 셈이 된다.

문제 해결

본 장에서는 이 문제에 대한 해결 방안으로, 정확한 내용의 정보 요청이 아니더라도 그 정보와 유사한 개념을 가진 정보를 제공하는 정보 에이전트에게 그 정보에 대한 요청을 하는 방안을 제시하고자 한다. 즉, 정보 에이전트 자체에 대한 선호도에 따라 연결하는 것이 아니라, 요청된 정보와 유사도가 높은 개념을 가진 정보, 즉 같은 온톨로지에 속하는 정보를 제공하는 정보 에이전트에게 정보 요청을 하게 된다. 이 경우, 임의의 정보 에이전트가 연결이 끊어질 경우에도, 그 정보 에이전트의 키워드가 등록되었던 온톨로지 내에, 등록된 다른 정보 제공 에이전트들이 남아있을 때에는, 질의 에이전트의 정보 요청에 또 다른 에이전트가 적절히 대응할 수 있게 되는 것이다. 예를 들어, 'car' 라는 키워드를 가지고 정보를 찾는 에이전트가 있을 때, 'car' 에 대한 정보를 제공하는 정보 에이전트는 없으나 'Automobile' 이라든가, 'Locomotive' 라는 정보를 제공하는 에이전트가 있다면, 같은 온톨로지 범주 내에 속한다는 것을 근거 삼아 정보를 요청할 수 있다.

2.2 가중치 분류

2.2.1 기본 가정 및 유사집단 분류

정보 에이전트와 정보 요청 에이전트 사이의 매개를 담당하는 온톨로지의 구축과, 구축된 온톨로지와 각각의 에이전트가 등록하는 가중치 벡터를 기반으로 하여 매칭 하기 위한 기본 가정은 다음과 같다.

조정 에이전트가 가지는 기본 온톨로지는 영어 명사로 이루어지며 정보 제공 에이전트는 조정 에이전트에 정보(키워드)를 등록한다. 키워드는 단일 명사, 또는 단일 명사로 취급되는 복합어 등으로 이루어져 있는데, 본 논문에서는 영어 단어로 이루어진 키워드를 이

용한다. 키워드는 조정 에이전트가 중요하다고 판단되는 키워드 순으로 제공되는데, 즉 다음과 같다.

- 조정 에이전트는 기본 온톨로지(Base Ontology)를 가진다.
- 정보 에이전트는 키워드 등록시, 기본 온톨로지에 대한 가중치를 퍼지 벡터로 등록한다.
- 키워드 집합은 영어 단일 명사 여러 개로 이루어져 있다.
- 정보 에이전트가 등록한 가중치는 다음 레벨과 같이 분류된다.
 VS : Very Specific S : Specific
 G : General R : Relative
 L : Little VL : Very Little

(온톨로지 구축의 기본 가정)

keyword	water, lake, Tahoe, California, America...
---------	--

(미국의 Tahoe 호수에 대한 정보 등록)

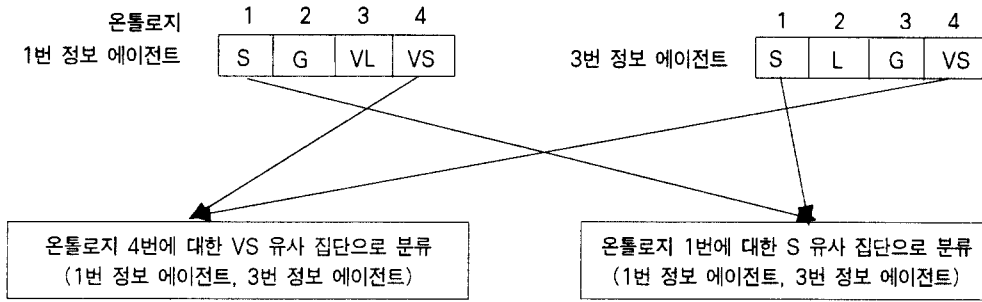
이 때, 단지 키워드만을 등록하여 그 자료를 요청하는 에이전트가 있을 때 넘겨주는 식의 단순 작업을 배제하기 위하여, 조정 에이전트가 이미 가지고 있는 기본 온톨로지에 근거하여 자신이 제공하는 자료의 가중치를 구한다. 가중치의 의미는, 정보가 어느 부류에 속하되 어느 정도로 구체적인 정보인지 나타내는 것이다. 초기 단계의 등록 과정은 다음과 같다.

A	B	C	D	Keyword
0.93	0.32	0.56	0.21	water, lake, Tahoe...

(기본 온톨로지에 대해 에이전트가 부여하는 가중치)

입력된 정보는 다음과 같이 분류된다.

0.9 이상	VS(Very Specific)
0.8 ~ 0.9	S(Specific)
0.65 ~ 0.8	G(General)
0.5 ~ 0.65	R(Relative)
0.35 ~ 0.5	L(Little)
0.35 이하	VL(Very Little)



(그림 1) 유사집단분류

서로 다른 정보 에이전트가 같은 항목에 대하여 S 이상(Specific)의 가중치를 등록했을 때, 이들을 VS(Very Specific) 집단, S(Specific) 집단의 유사 집단으로 분류한다. 이렇게 함으로써 질의 에이전트가 모호한 정보를 요청했을 때, 정보는 비록 불확실하지만 질의 에이전트가 중요하다고 선택한 항목에 대하여 밀접한 정보를 제공하는 여러 개의 정보 에이전트에게 도움을 요청할 수 있게 된다.

2.2.2 질의 에이전트의 정보 요청

본 연구에서는 정보 에이전트와 질의 에이전트간의 유사도를 구하기 위해 각 에이전트의 가중치가 필요하다. 정보 에이전트의 각 정보에 대한 가중치는 정보를 제공하는 에이전트가 부여하고, 질의 에이전트의 정보에 대한 가중치는 질의를 요청하는 사용자가 직접 부여하거나, 조정 에이전트가 강화 학습으로 얻어낸 값으로 부여할 수 있는데, 실험의 편의상 가상의 사용자를 대신하여 질의 에이전트가 부여하는 것으로 가정한다. 즉 조정 에이전트가 좀 더 밀접한 정보를 구하는데 도움이 되는 자료를 구하는 것이다. 가중치를 부여한다는 것은, 질의하는 내용이 어떤 항목(온톨로지)에 대하여 가장 관련이 있는가를 나타내는 것이다.

가중치 부여

- ▲ 사용자 측 가중치 : 가상의 사용자를 대신하여 질의 에이전트가 부여
- ▲ 정보 에이전트의 가중치 : 실험 중 임의의 데이터를 사용하여 부여

그러나, 사용자가 질의의 내용에 관해 전혀 알지 못하는 경우가 있다. 즉 사용자 자신도 질의가 무엇인지 모를 때, 또는 사용자의 질의가 질의 에이전트의 온톨로지

내에 존재하지 않을 때, 조정 에이전트는 그 사용자(질의 에이전트)에 대한 기존의 정보를 바탕으로 현재 질의에 대한 가중치를 적용하게 된다. 즉 사용자(질의 에이전트)가 평소 관심을 보이던 항목을 파악하여 적합한 정보 에이전트를 찾는다. 따라서 어떤 종류의 질의인지 알 수 없는 상황에서도 적절히 대처할 수 있는 조정 에이전트의 구축이 가능하게 된다.

2.3 가중치 벡터를 이용한 유사도 매칭

먼저 질의 에이전트가 요청한 키워드와 동일 키워드를 제공하는 에이전트를 찾는다. 이때 일치하는 정보가 있다면 그 키워드를 제공하는 정보 에이전트를 질의 에이전트와 이어줌으로써 조정 에이전트의 작업은 끝난다. 그러나 현실 세계에서 이렇게 일치하는 정보를 찾기란 쉽지 않다. 따라서 일치하지 않는 정보의 처리를 위한 해결 방안이 필수적이다.

만일 동일 정보를 제공하는 정보 에이전트가 존재하지 않을 때에는, 질의 에이전트의 각 온톨로지에 대한 가중치와, 질의 에이전트가 가장 높은 가중치를 부여한 온톨로지를 중심으로 유사 집단으로 분류된 정보 에이전트들의 유사도를 비교하여 최대의 값을 가지는 정보 에이전트를 구한다. 이때 정보 에이전트와 질의 에이전트의 유사도를 구하는 방법으로는 Costas P.Pappis가 제안한 다음의 3가지 유사도를 사용한다(8).

- (1) 교집합과 합집합의 연산을 이용한 유사도 (Measure based on the operations of union and intersection)

$$S_{A,B} = \frac{\sum (a_i \wedge b_i)}{\sum (a_i \vee b_i)}$$

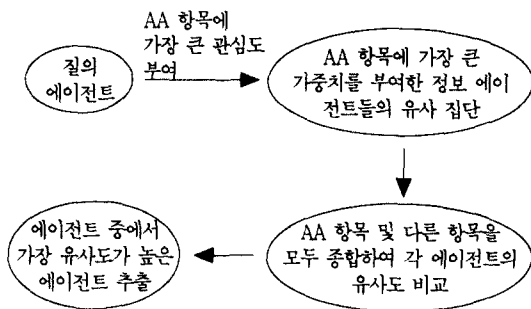
(2) 최대차를 이용한 유사도
(Measure based on the maximum difference)

$$S_{A,B} = 1 - \max_i (|a_i - b_i|)$$

(3) 차와 합을 이용한 유사도
(Measure based on the difference and the sum of grades of membership)

$$S_{A,B} = 1 - \frac{\sum_i |a_i - b_i|}{\sum_i (a_i + b_i)}$$

이렇게 하는 이유는, 질의 에이전트가 정보를 요청할 때 하나의 정보에 대해서만 특별한 질의를 할 수도 있지만, 여러 개의 정보에 대하여 다각적인 요청을 할 경우도 있기 때문이다. 따라서 가장 높은 관심도의 내용만을 찾아 주는 단순한 방법보다는, 여러 면으로 질의 에이전트의 만족도를 높여주는 정보를 제공하는 에이전트를 찾고자 한다. 그러므로, 질의 에이전트가 비교적 높은 관심도를 보였다고 판단되는 항목을 중심으로 분류되어 있는 정보 에이전트의 유사 집단을 찾아서, 여러 항목에 대한 관심도를 서로 비교하여 여러 면에서 가장 유사한 정보를 제공한다고 판단되는 정보 에이전트를 찾도록 한다. 이 방법의 타당성은 실험을 통하여 검증하도록 한다.



(그림 2) 유사도 기반 매칭

2.4 강화 학습을 이용한 특성값 구축

임의의 질의 에이전트가 어느 순간에, 어느 온톨로지에도 속하지 않는다는 신호와 함께 정보를 요청할 경우, 정보 에이전트와 비교 대상이 되는 가중치 벡터

값의 부재가 생긴다. 이런 경우는 사용자가 질의한 정보가 질의 에이전트의 온톨로지 내에 존재하지 않는 경우, 즉 질의 에이전트가 사용자가 의도한 정보가 어떤 의미를 가지는지 전혀 모르는 경우에 발생한다. 이런 경우를 대비하여, 조정 에이전트는 평상시에 각 질의 에이전트가 전송하는 온톨로지에 대한 가중치 벡터에 근거하여 강화 학습을 수행한다. 즉 그 이전에 임의의 질의 에이전트가 질의하고자 하는 키워드와 함께, 조정 에이전트의 온톨로지와 그 키워드의 관련성, 즉 가중치를 함께 전송한 경우가 있었던 경우, 그 가중치 자료는 결과로 연결된 유사도와 함께 그 질의 에이전트의 특성 자료로 분류된다. 유사도와 가중치는 각각 강화 학습의 보상값(r)과 현재 상태(s)로써, 다음의 식을 이용하여 새로운 상태(s)로 전이하는 것이다.

$$Q^*(s,a) + r + \gamma \max_a Q^*(s',a)$$

(r: 현재 상태 s'에서 새로운 상태 s로 전이되었을 경우의 보상 신호 값(유사도 값))

(γ: 현재 환경에 미치는 영향을 더욱 비중 있게 나타내기 위한 상수)

Q*(s',a) : 현재의 상태 값

Q*(s,a) : 갱신된 상태 값

조정 에이전트는 평소에 이러한 강화 학습을 수행하여 질의 에이전트의 특성을 파악하고 있으며, 어떤 시점에서 뜻밖의 질의가 요청되었을 경우 조정 에이전트는 강화 학습을 통하여 얻어진 자료를 바탕으로 그 질의 에이전트의 질의 키워드에 나름대로의 가중치를 부여하게 된다.

3. 구현 및 실험

3.1 실험 방법

정보 에이전트는 총 100개, 질의 에이전트는 총 20개가 존재하며, 등록하는 키워드는 각각의 에이전트가 가지는 가중치에 따라 다르다. 실험은 다음의 2단계로 나누어 진행한다.

먼저 조정 에이전트는 기본 12개의 기본 온톨로지를 가진다. 조정 에이전트의 온톨로지는 조정 에이전트의 관리자가 적절히 구현하는데, 본 실험에서는 '지리'에서 12개의 온톨로지를 구했다. 정보 에이전트가 키워드와 함께 등록되면 각각의 가중치 벡터에 따라 VS(Very Specific), S(Specific)의 유사 그룹으로 분

류된다. 본 실험은 이와 같이 나누어 유사도를 구하는 방법의 타당성에 관한 실험이다. 질의 에이전트가 질의 내용과 함께 각각의 온톨로지의 항목에 대하여 가중치를 등록한다. 질의 에이전트가 등록하는 가중치 중에서 가장 높은 순위의 항목을 중심으로, 이미 구성된 VS, S 그룹과 유사도를 비교하여 가장 유사한 정보를 제공한다. 그리고 판단되는 정보 에이전트에게 우선적으로 질의 하게 된다. 또한, 모든 정보 에이전트의 유사도를 구하고, 그 중에서 가장 큰 유사도를 보이는 정보 에이전트를 찾아, 유사 그룹으로 나누어 구한 경우의 정보 에이전트와, 가장 큰 유사도를 보이는 정보 에이전트가 제공하는 정보의 질을 비교하여 보도록 한다.

3.2 실험 결과

질의 에이전트, 즉 사용자가 등록한 가중치에 따라 집단별 유사도 비교를 한 결과이다. 정보 에이전트가 자신의 정보를 등록할 때 기본 온톨로지에 근거하여 가중치를 등록하고, 이것을 각 범위에 따라 VS(Very Specific), S(Specific) 집단으로 나눈다. 정보 에이전트는 총 100개, 질의 에이전트는 총 20개가 존재하며, 등록하는 키워드의 개수는 정보 에이전트가 가지는 각 항목에 대한 가중치에 따라 다르다. 사용자가 등록하는 가중치는, 실험의 편의상 가상의 사용자를 대신하여 질의 에이전트가 등록한다.

가중치 부여

- ▲ 사용자 측 가중치 : 가상의 사용자를 대신하여 질의 에이전트가 부여
- ▲ 정보 에이전트의 가중치 : 실험 중 임의의 데이터를 사용하여 부여

가상 사용자 측 가중치 추출

각 질의 에이전트는 온톨로지 상에서 무작위로 키워드를 추출하고, 그 키워드를 중심으로 하는 정규 분포에 따라 각 온톨로지에 대한 가중치가 얻어진다. 실험에 사용되는 변수를 정리하면 다음과 같다.

- ▲ 기본 온톨로지 : 0,1,2,3,4,5,6,7의 8개가 존재한다. 만일 질의 에이전트가 찾고자 하는 질의의 가중치가 3.8이라고 하면, 이 질의는 기본 온톨로지 3과 4사이 에 있으며, 또한 3보다는 4에 가깝다는 의미로 해석할 수 있다.
- ▲ 정보 에이전트가 질의 에이전트를 대신하여 가중치를 만들어 내는데, 질의 에이전트가 선호하는 온톨로

지는 8개의 온톨로지 중 어느 것이나 될 수 있다. 즉, 2.5번 온톨로지를 선호할 수도 있고, 1번만을 선호할 수도 있다. 가정할 수 있는 질의 에이전트의 예를 그래프로 나타내보면 다음과 같다(그림 3-1,3-2).

▲ 가중치 추출 과정은 다음과 같다.

① 질의 에이전트와 정보 에이전트가 선호하는 온톨로지를 임의로 추출하고, 그 온톨로지에 따른 그래프를 얻는다. 선호 온톨로지는 여러 개가 있을 수 있으므로 그래프의 모양도 다양할 수 있다.(그림 3, 그림 3-1, 그림 3-2)

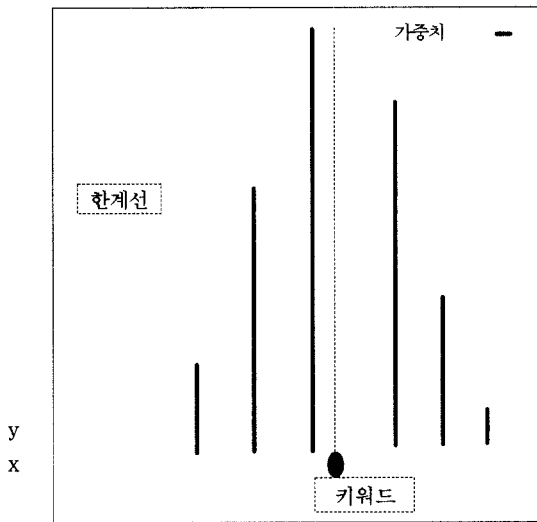
② 이제 질의 에이전트가 질의하려는 키워드를 임의로 얻는다. 그리고, 이 키워드와 질의 에이전트의 선호도 온톨로지 그래프가 만나는 점을 기준으로 정규 분포 그래프를 추출한다. 이렇게 추출한 그래프에서, 각 온톨로지에 해당하는 값이 각 온톨로지에 대한 질의 에이전트의 가중치가 된다. 질의 에이전트의 온톨로지의 의존도에 따라서, 키워드는 한 온톨로지 안에서만 추출될 수도 있고, 여러 곳에서 추출될 수도 있다.

③ 그러나 모든 가중치가 모두 채택되는 것은 아니고, 임의로 추출한 한계선보다 클 경우에만 채택된다. 키워드와 한계선은 무작위로 추출되므로, 선호하는 온톨로지가 단 한 개 존재한다 하더라도 얼마든지 임의의 키워드가 추출될 수 있다. 한계선이 높을 경우는 질의 에이전트가 선호하는 온톨로지에 대한 의존도가 높다는 것을 의미하고, 한계선이 낮을 경우는 질의 에이전트가 선호하는 온톨로지에 대한 의존도가 낮다는 것을 의미한다.

④ 질의 에이전트가 선호하는 온톨로지에 대해서만 일방적으로 많은 키워드를 추출하는 경우를 배제하기 위하여, <그림 3.3-1,3-2>의 한계선을 자주 변경하도록 한다. 한계선이 너무 높으면 하나의 온톨로지에 대한 키워드만을 추출하는 경우가 발생하므로, 10회의 질의마다 한계선을 변경하여, 무작위로 추출되는 키워드와 질의 에이전트의 선호도에 의존하는 키워드를 고루 추출하도록 한다.

⑤ 이렇게 하여 얻은 질의 에이전트의 각 온톨로지에 대한 가중치 값과, 정보 에이전트의 각 온톨로지에 대한 가중치 값을 비교하여 가장 적절한 정보를 제공하는 정보 에이전트를 구한다.

본 실험의 궁극적인 목적은 <그림 3.3-1,3-2>의 어떤 경우에도, 즉 질의 에이전트가 어떤 온톨로지에 어떤 선호도를 보인다 하더라도, 조정 에이전트는 질의 에이전트가 찾는 궁극적인 정보를 제공하는 정보 에이전트를 찾는 것이다.



(그림 3) 온톨로지에 대한 가중치 부여

(y축: 각 온톨로지에 대해서 질의 에이전트가 보이는 선호도 값

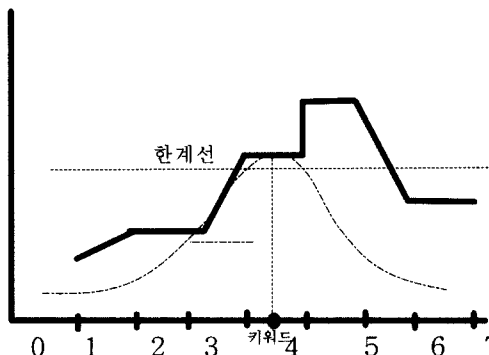
x축: 각 온톨로지(0~7)

키워드: 질의 에이전트가 질의한 값

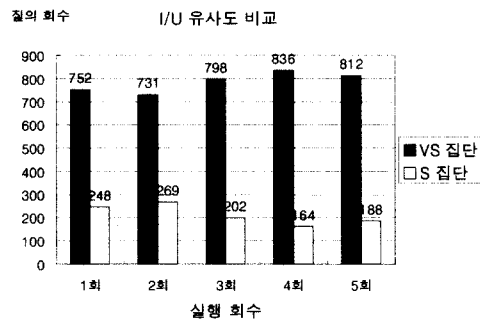
그래프: 질의 키워드와 질의 에이전트의 선호도 온톨로지 그래프가 만나는 점을 기준으로 한 정규 분포 그래프(각 온톨로지에 해당하는 값이 각 온톨로지에 대한 질의 에이전트의 가중치가 됨)

한계선: 키워드에 대한 선호도(y축)를 추출하는 한계값(한계선은 계속 변화하며, 한계선 이상의 값만 추출됨))

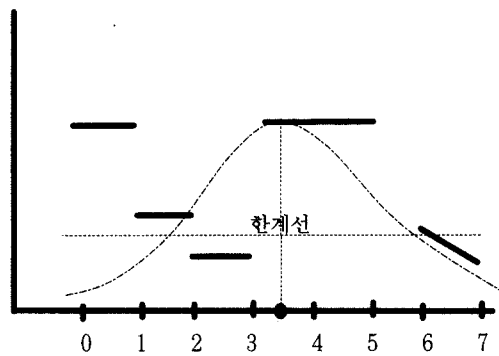
본 실험에서는 질의 에이전트의 가중치와 조정 에이전트에 등록된 정보 에이전트들의 가중치를 바탕으로 3가지의 유사도를 사용하여 성능을 비교한다. 다음은 각각의 유사도를 사용하여 얻은 결과의 그래프이다. 20개의 질의 에이전트가 100개의 정보 에이전트에게 1000번의 정보 요청을 하여 결과로 얻어진 유사도를 바탕으로 작성되었다. 그래프는 각각 1000번의 수행 후 VS 그룹의 유사도가 S 그룹의 유사도보다 큰 경우와 그 반대의 경우를 나타낸 것이다.



(그림 3-1) 온톨로지 4,5에 동일한 선호도를 보이는 질의 에이전트



(그림 4) I/U 유사도 비교

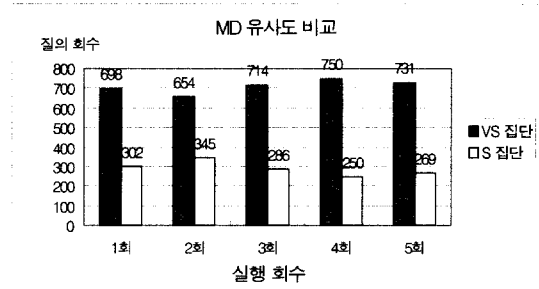


(그림 3-2) 일정하지 않은 선호도의 질의 에이전트

이 그래프는 교집합과 합집합의 연산을 이용한 유사도(Measure based on the operation of union and intersection)를 적용하여 얻은 결과이다. 1000번씩 5번의 수행 후 VS 그룹에서 얻어지는 유사도가 큰 회수가 약 780번(78%), S 그룹의 유사도가 큰 경우는 약 220(22%)번으로, VS 그룹의 유사도가 큰 경우가 월등히 많음을 알 수 있다.

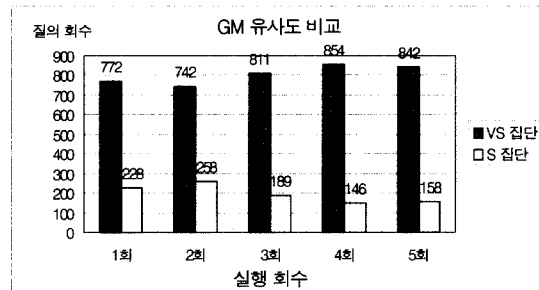
다음의 그래프는 최대차를 이용한 유사도(Measure based on the maximum difference)를 적용하여 얻어지는 결과이다. 이 경우 역시 VS 그룹의 유사도가 큰 경우는 약 710회(71%), S 그룹의 유사도가 큰 경우

가 약 290회(29%)로써 교집합과 합집합의 연산을 이용한 유사도보다 성능은 약간 떨어지지만 결과는 같다.

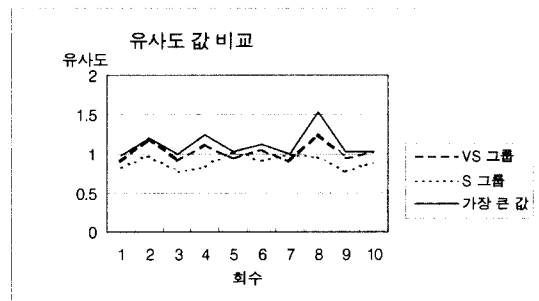


(그림 5) MD 유사도 비교

다음은 합과 차를 이용한 유사도(Measure based on the difference and the sum of grades of membership)을 적용한 결과이다. 결과는 앞의 두 유사도와 같고, VS 그룹의 유사도가 큰 경우가 약 800번(80%), 그 반대의 경우가 약 200번(20%)으로써 성능은 가장 우수하다.



(그림 6) GM 유사도 비교



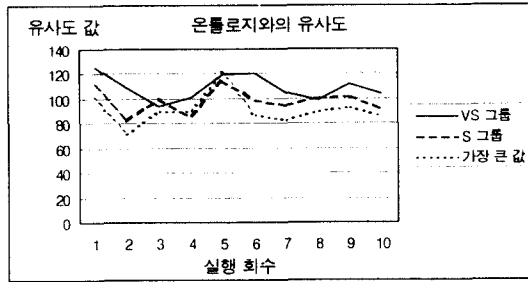
(그림 7) 집단별 유사도 값 비교

위 그래프는 100번의 질의를 10번 수행하여 얻어진 값으로, 가장 높은 유사도를 구한 경우의 유사도 값이 당연히 가장 높거나, 적어도 VS 집단, 또는 S 집단 중에서 큰 유사도 값과 동일하다. 그러나 이 경우, 유사도 값은 높으나 결과로 구한 정보 에이전트를 보면, 질의 에이전트가 질의한 내용과는 아무런 상관이 없는 경우를 추출한 경우가 상당히 많음을 확인할 수 있다. 즉 다만 가중치의 유사도만을 고려하여 결과 값이 높은 에이전트만을 연결한다고 하면, 질의 에이전트가 찾고자 하는 정보와는 전혀 무관한 정보를 제공하는 에이전트를 찾을 확률이 높아지게 된다고 볼 수 있다. 다음의 그래프는 각 집단이 구한 정보 에이전트의 정보가 속한 온톨로지와, 각각의 질의 에이전트의 질의 키워드가 속한 온톨로지를 비교하여 수치화한 것이다. 정보 에이전트는 유사도에 따라 여러 개가 구해지는데 (유사도에 따라 차례대로 연결하기 위하여), 대상이 되는 에이전트는 상위 30%의 에이전트로 한다. 계산식은 다음과 같다.

```

(if 질의 에이전트의 키워드가 속한 온톨로지(A) 정보 ⊆ 에이전트의 정보가 속한 온톨로지(B)
then
    유사도=유사도+3
    if 같다면 유사도=유사도+A가 요청하는 정보 수 + B가 제공하는 정보 수
    end if
else
    유사도=유사도-2-A가 요청하는 정보 수
end if) / 대상 정보 에이전트의 개수
    
```

1번 온톨로지에 3개의 키워드, 2번 온톨로지에 2개의 키워드를 가지고 질의하는 질의 에이전트와, 1번 온톨로지에 2개의 키워드, 3번 온톨로지에 1개의 정보를 제공하는 정보 에이전트를 예를 들어 보기로 한다. 이 경우 이 두 개의 에이전트의 유사도는, 3(1번 온톨로지가 동일하므로)+5(질의 에이전트의 키워드 3개와 정보 에이전트의 키워드 2개)-2(정보 에이전트는 질의 에이전트가 요구하는 2번 온톨로지를 제공하지 않는다)-2(질의 에이전트가 2번 온톨로지에 대해 요구하는 정보)=4, 즉 4 만큼의 유사도를 가지는 셈이 된다. 위 식을 적용하여 100번씩 10번 수행된 결과는 다음과 같다.



(그림 8) 온톨로지와 유사도

그래프에서 보듯이, 가장 큰 유사도 값을 구하는 경우는 몇 번의 경우를 제외하고는 VS, S 집단보다 온톨로지와 유사도가 낮은 것을 알 수 있다. 즉, VS, S 집단의 경우 유사도 값은 조금 낮을 수 있지만, 그보다 질의 에이전트의 요청에 대하여 여러 가지 면을 고려하여 가장 적절한 정보를 제공하는 정보 에이전트를 찾아 준다고 볼 수 있다. 유사도 값이 큰 에이전트만을 찾아 연결하여 주는 것은 질의 에이전트의 정보 요청에 대한 적절한 대응이 되지 못하는 것이다. 즉, 무조건 유사도가 높은 경우만을 뽑는 것이 아니라, 유사 집단으로 일단 분류한 뒤, 여러 사항을 염두에 두고서 가장 적절히 정보를 제공한다고 판단되는 정보 에이전트를 구하는 방법이 시간상, 효율상 유리한 것임을 알 수 있다.

4. 결론

본 논문에서는 정보 에이전트가 자신이 제공하는 정보를 조정 에이전트에 등록했을 때, 등록된 정보와 정확히 일치하지 않는 정보를 질의 에이전트가 요청했을 경우에 대한 해결 방안으로 유사도와 강화 학습을 기반한 조정 방법을 제안하였다. 본 논문에서 조정 에이전트의 역할은, 정보를 요청하는 에이전트가 최대한 목적을 달성할 수 있도록, 에이전트의 특성을 파악하여 가장 유사한 정보를 제공한다고 판단되는 정보 에이전트를 추출하는 것이다. 조정 에이전트는 각각의 에이전트가 제공하는 가중치 벡터를 기반으로, 유사도와 강화 학습을 사용하여 정보 에이전트와 질의 에이전트를 연결하였고, 실험을 통하여 타당성을 검증하였다.

본 논문에서는 가상의 사용자를 대신하여 질의 에이전트가 가중치 벡터를 구하였으나, 현실에 적용하기 위해서는 보다 구체적인 사용자 인터페이스가 필요하다는 것이 향후 과제로 남는다. 또한 조정 에이전트가 가지는 온톨로지는 실험을 위하여 임의로 추출하였으

나, 역시 실생활에 적용 가능하기 위해서는 체계적으로 짜여진 온톨로지의 구축이 필요하며, 유사 정보의 교환이 가능할 수 있도록 온톨로지끼리의 교환, 공유에 관한 연구도 필요할 것이다.

참고 문헌

- [1] Katia P. Sycara (1998), "Multiagents Systems," *AI magazine*, summer, pp 79-92.
- [2] Pdraig Cunningham and Richard Evans (1997), "Software Agents : A Review," <http://www.cs.tcd.ie/Breanda.Nangle/iag.html>, May 27.
- [3] 최중민 (1997), "에이전트의 개요와 연구 방향," *정보과학회지 제 15권 제 3호*, pp 7-16.
- [4] Thomas R.Gruber (1991), "The role of common ontology in achieving sharable and reusable knowledge bases," *Proc. of the Second International Conference*, pp 601-602, Cambridge, MA.
- [5] Davis and Smith (1983), "Negotiation as a metaphor for distributed problem solving," *Artificial Intelligence*, 20(1) : pp 63-109, Jan.
- [6] Schwartz D. (1993), "Cooperating heterogeneous system : A blackboard based meta approach," *Technical Report 93-112, Center for Automation and Intelligent Systems Research*, Case Western Reserve University.
- [7] SoJeong Youn, HoSub Seo, KyungWhan Oh (1998), "A Coordination Method for Middle Agent Based an Empirical Knowledge with Fuzzy Rules", SoGang University.
- [8] Costas P. Pappis, Nikos I. Karacapilidis (1993), "A Comparative assessment of measures of similarity of fuzzy values," *Fuzzy Sets and Systems* 56, P171-174, North-Holland.