

1990년대 이후 **온톨로지**(ontology)는 인공지능 연구에서 각광을 받고 있다. 연구자들마다 ‘온톨로지’란 용어를 조금씩 다른 의미로 사용하고 있지만, 널리 받아들여지는 정의는 “관심 영역 내 공유된 개념화에 대한 형식적이고 명시적인 명세”다. 여기서 ‘관심 영역’은 특정 영역 중심적이라는 것을, ‘공유된’은 관련된 사람들의 합의에 의한 것이라는 것을, ‘개념화’는 현실 세계에 대한 모형이라는 것을 뜻한다. 즉 특정 영역의 지식을 모델링하여 구성원들의 지식 공유 및 재사용을 가능하게 하는 것이 바로 온톨로지인 것이다. 또 ‘형식적’은 기계가 읽고 처리할 수 있는 형태로 온톨로지를 표현해야 한다는 것을 뜻한다. 그 결과로서 얻어지는 ‘명시적인 명세’는 일종의 공학적 구조물로서 다양한 용도로 사용된다.

온톨로지를 사전과 비교하면 ‘개념화’를 쉽게 이해할 수 있다. 사전에는 각각의 표제어에 대해 뜻풀이, 동의어, 반대어 등 언어적 특성들이 정리되어 있다. 온톨로지에는 표제어 대신 개념이, 그리고 언어적 특성들 대신 개념들 간 논리적 특성들이 기록된다. ‘개념(class)’은 어떤 공통된 속성들을 공유하는 ‘개체들(instances)’의 집합이고, 개체는 세상에 존재하는 구체적인 개별자이다. 온톨로지에서는 개념은 관계를 통해 다른 개념들과 연결된다. 필수적인 관계는 개념 간의 계층 구조를 형성하는 상속 관계이다. 상속 관계에서 하위 개념은 상위 개념의 모든 속성을 물려받는다. 예컨대 ‘스누피’라는 특정 개체가 속한 견종 ‘몰티즈’라는 개념은 ‘개’의 하위 개념이므로, ‘몰티즈’는 상위 개념인 ‘개’가 가진 모든 속성을 물려받는다. 널리 사용되는 또 다른 관계로 부분-전체 관계가 있다. 이외에도 온톨로지에는 관계를 포함한 다양한 논리적 특성들을 기록할 수 있다.

온톨로지 표현 언어는 대부분 일차 술어 논리에 기초를 두고 있다. 일차 술어 논리는 ‘모든’과 ‘어떤’을 변수와 함께 사용하는 언어로 표현력이 매우 뛰어나다. 예컨대 “진짜 이탈리아 피자는 오직 얇고 바삭한 베이스만을 갖는다.”를 일차 술어 논리로 옮기면 “모든 x에 대해, 만약 x가 진짜 이탈리아 피자라면, 얇고 바삭한 베이스인 어떤 y가 존재하고 x는 y를 베이스로 갖는다.”가 된다. 그런데 이것이 반드시 장점인 것은 아니다. 일차 술어 논리로 정교하고 복잡하게 표현된 온톨로지를 막상 기계는 효율적으로 다룰 수 없는 경우가 발생하기 때문이다. 따라서 온톨로지 표현 언어는 일차 술어 논리에 각종 제약을 두어 표현력을 줄이는 대신 취급을 용이하도록 한 것이 대부분이다. 예컨대 월드 와이드 웹 컨소시엄의 권고안인 ‘웹 온톨로지 언어’ OWL에는 Lite, DL, Full의 세 가지 버전이 있는데, 후자로 갈수록 표현력이 커진다. 즉 OWL DL은 OWL Lite의 확장이고 OWL Full은 OWL DL의 확장이다. OWL DL까지는 계산학적 완전성과 결정 가능성이 보장된다. 이는 OWL DL로 표현된 온톨로지에서는 추론 엔진이 유한한 시간 내에 항상 해를 찾을 수 있음을 뜻한다.

OWL을 쓰면 복잡하고 다양한 논리적 특성들을 표현할 수 있지만 논리학에 익숙하지 않은 사용자에게 OWL은 너무 어렵다. 이로 인해 그 이름과는 달리, 웹에서 OWL이 널리 쓰이는 것은 아직까지 요원해 보인다. 오히려 전문 지식에 대한 정교한 논리적

표현이 요구되는 영역에서는 OWL이 이용되는 경우가 있다. 예컨대 미국 국립암센터에서 개발한 의료 영역 온톨로지인 NCI 시소러스는 OWL 포맷으로도 제공되는데, 이것은 약 4만 개의 개념과 백 개 이상의 관계로 이루어져 있다. 이외에도 의료 영역은 일찍부터 여러 그룹에서 각기 목적에 맞는 온톨로지를 발전시켜 왔다. 대표적인 것으로는 UMLS, SNOMED-CT 등이 있다.

온톨로지는 일반적으로 특정 영역 종사자들의 관심과 필요에 의해 구축되나 반드시 그런 것은 아니다. 1984년 개발이 시작된 Cyc는 인간의 모든 지식을 담고자 하는 대규모 온톨로지다. 지식 공학자 소와(Sowa)는 철학의 연구 성과를 적극적으로 수용한 상위 수준 온톨로지를 제시한 바 있다. 세상에 존재하는 모든 것을 분류하려면 시간, 공간과 같은 일반적인 개념들을 다루어야만 하는데, 이는 철학자들이 이런 개념들에 대해 가장 오랫동안 깊이 사유했기 때문이다.

22. 온톨로지<sup>1)</sup>에 대한 설명으로 적절하지 않은 것은?

- ① 지식의 공유와 재사용을 위해 설계된 인공물이다.
- ② 대상 체계의 개념 구조를 명시적으로 드러내고자 한다.
- ③ 실제 사용하려면 기계가 처리할 수 있는 형태로 표현되어야 한다.
- ④ 개념과 그 개념에 속한 개체들은 상속 관계에 의해 서로 연결된다.
- ⑤ 동일한 영역에서도 종사자들의 관심과 필요에 따라 서로 다른 온톨로지가 구축될 수 있다.

23. 온톨로지 표현 언어에 대해 추론한 내용으로 적절 <보기>에서 있는 대로 고른 것은?

<보 기>

ㄱ. 동일한 온톨로지를 서로 다른 두 개의 언어로 각각 표현하기 위해서는 이들 언어의 표현력이 동등해야 한다.

ㄴ. 일차 술어 논리 표현 “모든 x에 대해, x가 빵이면 x는 장미이다.”는 ‘빵’이 상위 개념, ‘장미’가 하위 개념인 상속 관계를 나타낸다.

ㄷ. 계산학적 완전성에 대한 보장 없이 최대의 표현력을 활용하여 온톨로지 구축을 원하는 사용자는 OWL Lite보다는 OWL Full을 사용할 것이다.

- ① ㄱ                      ② ㄴ                      ③ ㄷ
- ④ ㄱ, ㄴ                ⑤ ㄴ, ㄷ                한 것만을

24. 윗글과 <보기>를 바탕으로 소와의 상위 수준 온톨로지에 대해 이해한 것으로 적절하지 않은 것은?

<보 기>

소와의 상위 수준 온톨로지를 그림으로 나타내면 다음과 같다.

T는 세상에 존재하는 모든 것들의 집합을, ⊥는 공집합을 뜻한다. T 바로 아래 원초적 개념으로 ‘Independent’와 ‘Relative’와 ‘Mediating’, ‘Physical’과 ‘Abstract’, ‘Continuant’와 ‘Occurrent’ 이렇게 7가지가 있다. 하나의 선으로 연결된 두 개념 중 위쪽이 상위 개념, 아래쪽이 하위 개념이다.

한편 상속 관계는 추이성(transitivity)을 갖는 대표적인 관계다. 즉 A, B, C가 각각 개념이라 할 때, 하위 개념 A가 상위 개념 B와 상속 관계를 맺고 하위 개념 B가 상위 개념 C와 상속 관계를 맺으면, 하위 개념 A는 상위 개념 C와 상속 관계를 맺는다.

- ① 상위 개념으로 원초적 개념을 단 한 개만 갖는 개념은 없고, 오직 2개의 원초적 개념을 갖는 개념은 모두 6개다.
- ② T는 세상에 존재하는 모든 것들이므로 이 개념은 존재하는 모든 속성을 다 가지고 있고, ⊥에는 어떠한 개체도 속하지 않으므로 이 개념은 어떠한 속성도 갖지 않는다.
- ③ ‘Continuant’와 ‘Occurrent’의 공통 하위 개념은 오직 ⊥뿐이므로, ‘Continuant’의 속성과 ‘Occurrent’의 속성을 모두 갖는 개체는 존재하지 않는다.
- ④ ‘Object’는 ‘Actuality’의 하위 개념이고 또한 ‘Continuant’의 하위 개념이기도 하므로, ‘Actuality’의 속성과 ‘Continuant’의 속성을 모두 물려받는다.
- ⑤ ‘Process’는 ‘Actuality’의 하위 개념이고 ‘Actuality’는 ‘Physical’의 하위 개념인데, 상속 관계는 추이성을 가지므로, ‘Process’는 ‘Physical’의 하위 개념이다.