

인공지능

# 컴퓨터 추론 및 지식 공유능력 크게 확대 철자·문법은 물론 문장내용도 수정 가능

피로 모르고 같은 작업 무한 되풀이 인간으로서는 매우 어렵다고 생각하는 작업을 컴퓨터는 간단히 해치운다는 것을 우리는 그동안 경험해 왔다. 1944년의 맨허튼계획에서는 수십명이 여러 달에 걸쳐 수치계산에 매달렸으나 오늘날의 기술은 몇분 안들이고도 같은 일을 할 수 있다. 그러나 1956년 여름 미국 다트머스대학에 모인 인공지능 전문가들 중에서 40여년 뒤 그들이 설정한 인공지능의 목표가 오늘날 이토록 미달하리라고 예측한 사람은 아무도 없다.

문제는 컴퓨터의 빈약한 추론능력이 다. 컴퓨터가 강력한 추론능력을 발휘하는 것은 다만 적용분야를 좁게 한정하는 경우에 한한다. 예컨대 1965년 미국 스탠퍼드대학의 덴트럴계획에서는 고도의 추론으로 화합물의 화학구조를 추정하는 시스템을 개발하고 있다. 또 1975년 개발된 의료진단시스템인 마이신도 뇌막염으로 한정하는 한 평균적인 내과의보다 높은 정도로 진단할 수 있다. 이렇게 작업을 간단하게 규칙화 할 수 있는 한 컴퓨터의 능력은 인간을 훨씬 웃돈다. 컴퓨터는 피로를 모르고 같은 작업을 무한히 되

풀이 할 수 있기 때문이다. 한편 인간으로서는 대수롭지 않은 작업인 예컨대 친구의 얼굴을 식별한다거나 이해하기 어려운 낱말을 회화의 문맥으로 보충하는 일은 자동화하기 어렵다는 것이 오늘의 현실이다.

그동안 컴퓨터 네트워크, 사용자 인터페이스(다른 기능을 가진 기계장치 간의 경계에서 공용되는 부분 또는 2자를 중개하는 장치) 에이전트, 하드웨어에 관한 기술이 비약적인 발전을 이룩했는데도 불구하고 인공지능 연구 개발에는 진전을 거의 보지 못했다. 1970년대부터 80년대 초에 걸친 초기의 인공지능 연구는 어느 정도의 성과를 올렸으나 그 뒤 이렇다 할 발전은 없었고 오늘날 인공지능의 꿈을 포기하는 과학자들도 있다.

전문가시스템이라고 불리는 프로그램은 전문적인 특정한 영역에서는 그런대로 기능을 발휘하지만 질문이 조금이라도 전문영역을 벗어나는 순간부터 손을 들고 만다. 더욱이 이런 시스템은 질문이 자기의 능력을 넘어서는 것인지 아닌지조차 판단할 수 없다. 이들 프로그램은 또 서로 지식을 공유할 수도 없다는 결점을 지니고 있다.

예컨대 의료진단시스템인 마이신은 폐를 검진하는 프로그램이나 암의 화학요법을 자문하는 프로그램과 '대화'를 할 수 없다. 이런 현상은 지식을 되도록 간단하게 기술(記述)하기 위해 시스템개발자가 영역(領域) 특유의 가정에 바탕을 두고 프로그램을 작성했기 때문이다.

## 애매성·비유 넘치는 문장도 이해

인간은 많은 지식을 다른 사람과 함께 갖고 있다. 지적으로 행동하기 위해 필요한 지식의 많은 부분은 수천년에 걸친 문화적·생물학적 진화를 통하거나 또는 많은 사람들이 어울리던 유아기의 체험을 통해 공유하기 때문에 평상시에 의식하는 일이 없다. 인간처럼 컴퓨터끼리도 지식을 공유하게 만들자면 의식조차 하지 않는 이런 암묵의 지식을 모두 끄집어내어 계산할 수 있는 형식으로 고쳐 써야 한다.

미국 텍사스주 오스틴 소재의 사이코프사 CYC 프로젝트 연구자들은 10여년 전부터 이런 작업을 계속하고 있다. 이들은 먼저 신문기사나 소설·광고 등의 단편을 모아 개개의 글을 이해하는데 필요한 공유지식을 이은 작



玄源福

〈과학저널리스트/본지 편집위원〉

아무리 단순한 작업에도 우리는 상식적 지식을 사용한다. 사진의 주석 리스트를 보고 피부가 젖은 사람의 사진을 찾고 있다. 주석에는 ‘젖은’이라는 말의 유의어(類義語)는 없으나 상식을 사용하면 남자의 피부가 땀으로 젖어 있다는 것을 알 수 있다.

업부터 착수했다. 이를테면 글을 쓰는 사람은 독자가 갖고 있는 지식을 어떻게 가정하여 글을 쓰는가를 분석하는 작업이다. 그 결과 정의를 내린 개념의 수는 약 10만 그리고 분석한 상식적 지식은 약 1백만개에 이르렀다. 이 상식적 지식은 신문에 적을 수 있는 사실도 아니고 사전에서 보여 주는 정의도 아니다. 그것은 사람들이 수프를 어떻게 마신다거나 어린이는 어떤 동물을 무서워한다든가 따위의 보통 누구나 보고 많은 사람이 공통으로 믿고 있는 그런 따위의 지식이다.

CYC가 완벽하다고 할 수준에 이르자면 아직도 더 많은 노력이 필요하지

만 적어도 지식의 공유를 추진하는 출발점에는 접근하고 있다고 평가되고 있다. CYC의 지식베이스를 사용하면 애매성이나 비유에 넘치는 문장도 이해할 수 있는 프로그램을 작성할 수 있게 될 것이다. 이렇게 되면 CYC는 문장에서 정보를 끄집어냄으로써 이미 갖고 있는 개념을 보강하거나 새로운 개념을 획득할 수 있게 된다. 또 새로운 규칙을 발견하거나 세계에 관한 그럴싸한 가설을 세워 이것을 검증할 수도 있을 것이다. CYC 프로젝트는 2005년까지 지식베이스를 더욱 확장할 계획이며 응용시스템도 개발할 계획이다. 스프레드시트나 데이터베이스, 문

서편집시스템, 온라인검색시스템 등 우리가 흔히 사용하는 응용시스템 속에 상식적 지식을 꾸며 넣을 계획이다. 그래서 앞으로는 워드프로세서가 철자나 문법을 점검할 뿐 아니라 문장의 내용까지 점검할 수 있게 된다. 예컨대 ‘이 문제는 뒤에 의논한다’라고 말한 뒤 전혀 의논을 하지 않고 있으면 기계는 이런 사실을 경고해 주는 기능을 생각할 수 있다. 문서검색프로그램에서도 사용자의 질문내용을 잘 음미하여 문서가 키워드를 포함하지 않는 경우에도 상식적 지식으로 충분한 관련성이 있으면 키워드를 선택할 수 있게 된다.

이미 눈부신 발전을 거듭하고 있는 하드웨어나 네트워크와 이런 지적 프로그램이 잘 공명한다면 컴퓨터가 하는 온갖 서비스는 오늘날보다 훨씬 비용이 덜 들고 널리 보급될 것이다. 또 사용자의 정보를 훌륭하게 관리하거나 보다 지적인 에이전트 소프트웨어를 작성할 수 있게 되어 가상적인 환경 속에 깊숙이 젖어들 수 있다. 인공지능의 목표는 코 앞으로 다가오기 시작했으며 21세기에는 획기적인 변화가 전개될 것으로 기대된다.