

영화 매트릭스와 화이트헤드의 시스템의 진화

전 철*

2004년 8월 22일

영화 매트릭스와 화이트헤드의 가상과 현실의 문제¹에서는, 가상의 매트릭스를 구성하는 인공지능과 그 가상의 세계를 깨고자 하는 세력 사이의 가상과 현실의 개념이 고려되었다. 여기에서는 영화 매트릭스 2 Reloaded에서 아주 인상적으로 제시된 인공지능의 진화의 문제를 성찰하고자 한다. 이 진화의 문제를 꼭 인공지능의 특징으로만 제한적으로 이해할 필요는 없다. 어찌피 지능과 생물과 생명체가 공유하는 어떠한 보편적 특징으로서 진화의 문제를 고려할 수 있기 때문이다. 만약 진화라는 표현이 생물학적인 뉘앙스로 매우 치중되었다 한다면, 패턴화 내지는 규칙화라는 표현으로 바꾸어도 무방할 것이다.

매트릭스 2의 결정적인 반전은 네오의 정체성에 대한 자각이다. 네오는 가상적인 매트릭스라는 시스템이 야기할 수 있는 오류가 형상화된 존재이다. 그의 역할은 말하자면 시스템의 프로그램적인 오류로 인한 버그를 잘 보존하고 기록함으로써, 그 이후 시스템을 더 완벽하게 구성하게 하는 하나의 구성요소일 뿐이다. 그는 더 이상 가상의 세계를 파괴하고 새로운 현실을 창조하는 구세주가 아니라, 그 가상의 시스템에 부속된 효과적인 장치일 뿐이다.

이 영화에서는 인간과 인공지능이 어떠한 이유로 대립각을긋게 되었는지를 분명히 제시하고 있지 않다. 그리고 그것이 영화의 핵심적인 해명거리가 되지도 않아 보인다. 그러나 인공지능이 인간을 대하는 태도와 그 정황을 유추해 볼 때 이와 같은 판단은 적절할 듯하다. 즉 인공지능은 더욱 더 완벽하게 인간을 장악할 수 있는 매트릭스의 세계를 끊임없이 구현하려고 한다는 것이다. 오류에 관한 소스를 갖고 있는 네오와 같은 매신저들은 더 완성도 높은 매트릭스 시스템의 구현을 위하여 이미 이전에 5번이나 매트릭스를 리로딩 하고

*전철의 신학동네 <http://theology.co.kr>

¹영화 매트릭스와 화이트헤드의 가상과 현실의 문제

제거되었다. 그러나 이러한 시스템의 진화는 꼭 인공지능에만 국한되는가? 우리는 여기에서 그 시스템이 현실의 생명체이든 아니면 가상의 매트릭스이든, 그 각각의 시스템은 어떻게 자신을 진화시키는가 하는 문제로 확대시켜 보도록 하자.

우선 이 시스템의 진화의 문제와 직접적으로 관련된 통찰은 괴델의 불완전성의 정리에서 어떠한 단서를 발견할 수 있다. 우선 매우 혁명적인 발견이라 할 수 있는 괴델의 불완전성의 정리란 다음과 같다. 하나, 어떠한 공리체계가 일관적이고 무모순적이라고 하더라도 그 체계로 증명 혹은 반증할 수 없는 명제가 반드시 존재한다는 것이다. 둘, 공리체계가 모순적이지 않다는 것을 그 체계 안에서는 증명할 수 없다는 것이다.² 즉 괴델이 언급한 공리체계를 시스템으로 이해해본다면, 그 시스템이 아무리 모순이 없고 완벽하게 구현한다 하더라도 거기에는 그 시스템에서 어떻게 진위여부를 가려낼 수 없는 정보가 새롭게 등장한다는 것이다. 그리고 그 시스템의 결함은 그 내부에서는 확인할 어떠한 방법이 없다는 것이다. 사실 이 정리는 수론에 관한 놀라운 발견이지만, 철학의 건물이 당한 피폭의 충격들은 훨씬 커 보이기도 한다. 시간이 상당히 지난 지금도 여전히 여러 곳에서 휘청거리고 있으니 말이다.

이 인공지능의 시스템은 매트릭스라는 가상세계를 통하여 인간의 인격을 감쪽같이 통제해야 하는 과제가 있다. 즉 이 인공지능 시스템의 조건은 현실세계에 충분히 충족적이어야 한다. 이런 면에선 매트릭스는 현실의 인간 개체들이 완벽하게 활동할 수 있도록 하는 하나의 시스템이 되어야 하기에 그 충족을 위하여 끊임없는 자기보완이 필요하다. 심지어 완벽한 공리계 안에서도 해결해 낼 수 없는 새로운 정보와 명제가 등장한다고 할 때, 가상의 매트릭스에서 현실을 충실히 구현하고자 하는 막대한 과제를 지니고 있는 매트릭스라는 공리계에서 얼마나 많은 낯선 정보와 명제가 등장할 지는 어렵지 않게 추측할 수 있을 것이다.

실로 인간 조건의 공리계를 충족하고 포섭할 수 있는 매트릭스의 공리계는 더 얼마나 구현되기 어려운 일인가. 네오와 같은 디버깅 소스가 계속 시스템에 새롭게 구현된다 하더라도, 그 리로딩의 과정이 인간 조건들의 복잡성을 공리계 안에 수렴할 수 있을 지는 회의적으로 보인다. 어쩌면 인공지능 시스템의 수렴도가 인간 복잡성의 발산도를 극복할 수 없을런지도 모른다. 사실 네오는 인간의 공리계와 인공지능의 공리계를 동시에 묘하게 공유하고 있기 때문에 시스템에 있어서는 불치의 바이러스로 기능될 소지가 논리적으로 다분

²Kurt Gödel, Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme (1931)

히 크다. 그러므로 괴델의 불완전성의 정리를 네오의 이중적 공리계 문제로 직접 연결시키는 것은 과한 판단인지도 모른다.

그러나 아주 흥미롭게도 시스템 내부에서 순수하게 등장한 스미스의 바이러스로의 변식은 공리계의 문제와 직결되는 대목이다. 스미스는 시스템 내부에서 필요에 의해 발생한 존재이다. 그러나 시스템의 컨트롤을 벗어나는 바이러스이며 오히려 시스템을 위협하는 바이러스이다. 현실에서 출현한 가상이지만 그 가상이 오히려 현실장악력을 더 지니는 경우라고 할 수 있다. 매트릭스에서는 인간세계에서 우발적으로 출현한 인공지능, 그리고 인공지능에서 우발적으로 출현한 스미스를 통해서 이러한 시스템의 진화가 지니는 부정적인 측면을 선명하게 강조하였다. 그러나 이러한 한 시스템에서 등장했지만 시스템이 감당하지 못하는, 소위 새로운 명제의 등장은 그렇게 꼭 암울한 요소만을 지니지는 않는다.

그것은 생명의 진화를 구조적으로 설명할 수 있는 적절한 해석일 수도 있으며, 소위 튜링 이후의 주제인 인공지능의 진화 매커니즘을 이해할 수 있는 중요한 요소일 수도 있다. 분명한 것은 오늘날 그것이 생명이든 인공지능이든 대상에 대한 자신의 규칙을 어떻게 새롭게 구성해 나아가는지 완벽히 밝혀내고 있지 못하였다는 점이다. 그렇기 때문에 이러한 시스템의 진화라는 화두가 공각기동대, 터미네이터, 매트릭스, 아이로봇에 이르기까지 영화를 매개로 유사한 방식으로 표현되어온 것이다.

사실 폭넓게 이 생명의 진화, 생명의 생존 전략의 차원에서 시스템의 진화는 그것은 단순히 지엽적인 문제가 아니라 존재 자체의 일반적 특징의 중요한 면이기도 하다. 이러한 괴델의 수론이나 사이버네틱스의 인공지능, 루만의 시스템이론, 그리고 샤르맹의 복잡성의 원칙과 어떠한 일반적 특징을 공유하고 있는 면을 화이트헤드의 범주이론에서도 발견할 수 있다. 사실 화이트헤드의 집합론에 대한 고민을 괴델이 직접 받아 고민하여 수론에서 새롭게 발견한 것이 불완정성의 원리이다. 그만큼 이 묘한 범주와 대상의 이율배반은 화이트헤드에게도 끊임없는 고민거리였다. 그의 수학원리에서의 집합론이 어떻게 후기 형이상학에서 직접 검토되고 적용되었는지를 확인하는 것 자체가 하나의 큰 과제로 보여진다. 하지만 간단하게 우리가 매트릭스에서 제시된 시스템의 진화와 관련된 화이트헤드의 사고들을 유추해보자.

그는 소위 범주론자라 불릴 수 있다. 그의 형이상학을 통하여 드러나는 세계는 이러한 복잡한 범주들을 전제로 하여 아주 정교하게 결합되고 구성된 세계들이다. 존재하는 것에 관한 범주, 설명될 수 있는 것의 범주, 이러한 범주로 인하여 제시되는 범주 등 3조의 큰

범주가 있고 이 각 조에는 또 각각의 범주 내용이 있다. 그리고 이 3조의 기반이 되는 범주로 궁극자의 범주 The Category of the Ultimate가 있다. 사실 화이트헤드가 그 엄청난 수십개 하부의 범주들과 그 상위의 범주들의 수직적 층위간에 어떠한 정합성을 확보해내고 있는지, 나는 여전히 화이트헤드의 머리를 들어가보지 않아서 오리무중이다. 그러나 화이트헤드는 이 궁극자의 범주에 대해서는 짧게나마 언급하면서, 모든 수십개의 범주들을 작동시키게 하는 궁극적 요인이 된다고 말을 한다. 또한 이 궁극자의 범주는 다른 범주에 기여하고 기능할 때에만 의미가 있지 그 자체는 아무 의미가 없는 범주라고 말을 한다. 이는 말하자면, 그 자체는 아무것도 아니지만 어떠한 사물의 궁극적인 새로움을 설명해내는 범주이다.

괴델은 공리계에서 등장했지만 공리계로 환원이 안되는 새로운 정보의 출현 패턴을 수론에서 증명해냈다면, 화이트헤드는 이러한 하부단위로 환원이 안되는 새로움을 하나의 일반적 패턴으로 이해하고, 과감하게 그 새로움의 출현이 자연의 궁극적 특성이라는 가설을 제시하였다. 즉 화이트헤드는 이러한 새로움의 문제를 존재 일반의 궁극적인 특성으로 확대시켰다.

네오는 매트릭스 프로그램의 공리계로 포섭되지 않는 새로운 오류들의 정보 소스를 시스템에 첨가하는 진화된 시스템적 존재이다. 시스템의 진화는 환경의 민감성에 상응한다. 매트릭스에서 인공지능 시스템은 인간의 장악을 위하여 효과적인 공리계 구성을 끊임없이 시도한다. 그 시도의 과정에서 프로그램 외부로부터 일부 생성된 네오라는 이질적 정보와 순수하게 프로그램 내부에서 생성된 스미스라는 이질적 정보를 수렴하고 새로운 시스템의 내적 요소로 구현해야 한다. 인공지능 시스템은 두 요소를 어떻게 포섭하여 진화할 수 있는지, 그리고 매트릭스의 공리계와 현실세계의 공리계는 어떻게 서로 관계하며 그로 인하여 진화할 수 있는지는 매트릭스 3편 Revolution에나 가서야 선명하게 제시될 듯 하다.